Decadal prediction of interannual tropical and North Pacific sea surface temperature

Fabian Lienert¹ and Francisco J. Doblas-Reyes¹,²

Received 18 October 2012; revised 30 April 2013; accepted 4 May 2013.

The performance of a coupled climate forecast system initialized with observations, relative to the accompanying uninitialized system, to re-forecast annual-mean tropical and North Pacific Sea Surface Temperature (SST) departures at the decadal time scale is described. The study finds that the correlation skill of the leading Empirical Orthogonal Function (EOF) mode of North Pacific SSTs, i.e., the Pacific Decadal Oscillation (PDO), is limited to about one year, while the second mode, i.e., the North Pacific Gyre Oscillation (NPGO), is skillfully forecast throughout the 10-year forecast range. In the tropical Pacific, it is found that the correlation skill of the leading EOF mode of tropical Pacific SSTs, i.e., El Niño Southern Oscillation (ENSO), and the second mode, i.e., ENSO-Modoki, is limited to about two years and one year. A main contributor to forecast skill of the NPGO are the effects of the long-term trend on Pacific SSTs while little impact of the initialization was found.


1. Introduction

Climate in the North Pacific (NP) varies on seasonal to decadal and longer time scales and it influences terrestrial climate over the adjacent continents. For example, the climatic state of the NP is an important predictor for droughts and heavy rainfall events in North America [Cook et al., 2011]. The decadal prediction of NP climate departures is therefore of major relevance for societies as well as stakeholders in affected regions. The focus of this study is on the assessment of how a state-of-the-art coupled climate forecast system initialized with observations forecasts the evolution of the leading two Empirical Orthogonal Function (EOF) modes of NP Sea Surface Temperature (SST) departures (see their spatial manifestations in Figure 1 (top)) at the decadal time scale. These two modes combined explain more than 50% of the total NP annual-mean SST variance.

Many studies have focused on the most prominent expression of NP SST departures [e.g., Latif and Barnett, 1994], called the Pacific Decadal Oscillation (PDO) [Mantua et al., 1997]. Despite promising decadal spectral peaks of the PDO index, empirical studies imply that the PDO is rather the result of the oceanic response to a combination of forcings, i.e., in the tropical Pacific [Lienert et al., 2011] and local variability of SST anomalies in the Kuroshio and the climatological Aleutian Low, acting on distinct time scales [Schneider and Cornuelle, 2005] than an ocean driven mode. The predictability of the PDO interannually is therefore thought to be low [Newman, 2007]. At the grid-point level, the NP is the region with the least decadal SST predictive skill in a range of climate prediction systems [Kim et al., 2012; Doblas-Reyes et al., 2011; Guemas et al., 2012].

On the other hand, the second mode of NP SST departures, called the North Pacific Gyre Oscillation (NPGO) [Di Lorenzo et al., 2008] or “Victoria Mode” [Bond et al., 2003], have attracted less attention. Shifts of NP climate at the beginning of the 21st century that apparently took place independently of variations in the PDO index have been explained by variations in the NPGO index [Bond et al., 2003]. The NPGO is driven by fluctuations in the strength of the NP current that feeds the Alaskan and Californian current. Dramatic changes of coastal upwelling and ecosystems in the north-east Pacific have recently been linked to the NPGO [Chenillat et al., 2012; Di Lorenzo et al., 2008, 2009]. Part of the NPGO variance is apparently of central equatorial Pacific origin, a result that led [Di Lorenzo et al. [2010]] to the conclusion that the decadal predictability of the NPGO is limited because of the uncertain response of the tropical Pacific to the transient changes in external forcing [e.g., Merryfield, 2006; Collins et al., 2010].

This study aims to answer the question of whether interannual-to-decadal climate re-forecasts contain useful information on the future evolution of the leading two EOF modes of tropical and NP SST departures. After introducing the climate forecast system, observational datasets and the statistical techniques used in this paper, we start in the results section with the verification (taking into account the role of the trend) of the re-forecasts in the NP region (Subsection 3.1). We finish with a similar analysis but for the tropical Pacific — presented in Subsection 3.2.
2. Methods

In seasonal to decadal climate prediction studies, the quality of the decadal forecasts is usually inferred from historical forecasts (re-forecasts, also known as “hindcasts”) that are verified against the observational reference. Here, we use decadal re-forecasts from the UK Met Office Decadal Prediction System (DePreSys) [Smith et al., 2007] that performed the perturbed parameter part of the EU-FP7 ENSEMBLES project [Doblas-Reyes et al., 2009]. DePreSys comprises nine different model variants of the Third Hadley Center Coupled Ocean-Atmosphere GCM (HadCM3) by perturbing poorly constrained model parameters. The freely running model variants sample a wide range of climate properties, with climate sensitivities ranging from 2.6 to 7.1˚C and ENSO amplitudes (simulated standard deviation of monthly central equatorial Pacific SST anomalies) ranging from 0.5 to 1.2˚C compared to the observed value of 0.8˚C [Smith et al., 2010].

The initial states for DePreSys Assim are obtained separately for each model variant. The coupled model is nudged to monthly estimates of the ocean state and six-hourly estimates of the atmospheric state. The estimates used in the nudging are obtained by adding observed anomalies to an estimate of the model climate. The observed anomalies come from ECMWF’s 40-year Re-analysis [Uppala et al., 2005] up to 2002 and the ECMWF operational analysis afterwards in the atmosphere and from an objective ocean temperature and salinity analysis [Smith and Murphy, 2007] where SSTs are from the Hadley Center sea ice and SST dataset (HadISST) [Rayner et al., 2003]. The nudged simulations start in 1958 from the transient simulations that used prescribed forcings. The predictions are initialized using model restarts obtained from the nudged simulations for 1 November of each year over the period 1960–2005. This method usually leads to a smaller model drift and thus a smaller bias removal than the full field initialization method.

A parallel set of DePreSys forecasts (NoAssim) is produced without initial state information, i.e., using initial states from freely running historical simulations as opposed to model states initialized with observations, thus allowing...
an estimate of the benefit of the initialization. NoAssim simulations start from the transient simulations on 1 November each year from 1960 to 2005. The radiative forcing in all experiments—Assim and NoAssim—is inferred from the observed composition of atmospheric greenhouse gases and aerosols up to 2000 and the A1B scenario afterwards, while solar activity and volcanic aerosols are dealt with in forecast mode, i.e., repetition of the previous 11-year solar cycle and prescribed volcanic aerosol load damped with e-folding time of 1 year. A forecast comprises nine ensemble members (one by each model variant) for both Assim and NoAssim. The performance of decadal near-surface temperature forecasts of DePreSys in the Pacific is comparable to other systems [see p. 41 in Doblas-Reyes et al., 2010].

In this study, the Extended Reconstructed SST version 3b observational dataset [Smith et al., 2008] serves as the SST reference from November 1960 to October 2011, the verification period. The forecast started in 2005 is therefore verified in the first six years only, the one started in 2004 in the first seven years, etc. The effect of the different SST datasets used for the initialization and the verification on forecast error at the start of the forecast is expected to be small, because the differences in the initial conditions of the model variants of DePreSys Assim, that sample the observed uncertainty, are larger than the differences in the observational datasets.

The confidence intervals for the scores of time series in this study are based on the bootstrap where re-forecast/reference pairs have been resampled 500 times with replacement [Lanzante, 2005; Jolliffe, 2007]. We select blocks of 4 consecutive years for the bootstrap in order to take into account that the consecutive years are not independent of each other. In this manner, our method accounts for serial correlation [Zwiers, 1990]. We have chosen to select blocks of 4 years, because the observed autocorrelation of the NP modes drops to insignificant values after this period. The estimated confidence level has then been computed from the scores of each of the 500 samples. The average correlation between swapping forecast members treated as “observations” and the ensemble mean of the remaining members yields our potential skill estimate [similar to the one in Doblas-Reyes et al., 2009 and Kharin et al., 2009]. It provides an estimate of the perfect model skill, which may be reached after improvements to the forecast system, based on the spread of the model ensemble.

Following Newman [2007], the PDO and the NPGO are identified as the leading two EOFs of detrended annual-mean (Nov–Oct) observed NP SST anomalies (20°–65°N). To detrend, the linear regression of observed SST anomalies against the observed global-mean near-surface air temperature (GMT) from the Hadley Center/Climatic Research Unit [HadCRUT3, Brohan et al., 2006] dataset is removed at each grid point. The spatial patterns of the leading two EOF modes of NP SST anomalies are shown in Figure 1 (top). The first mode explains roughly 36% and the second mode 17% of the total SST variance in the observational dataset of annual means.

The estimated slope of the observed trend in NP SST is illustrated in Figure 2 (left). Over the past 51 years, both the Arctic and south-western NP have experienced a warming of 1.5–2 K. On the other hand, the eastern part of the basin has experienced a cooling of 0.5–1 K. For a comparison between the trend component vs. the interannual SST variability, Figure 2 (right) shows the estimated slope of the trend normalized by the residual SST variance. We note that the trend component dominates the year-to-year variability in the Arctic NP where local SST variability independent of the trend is relatively small. In the Kuroshio region east of Japan, however, both the trend and residual variability contribute the year-to-year SST variability.

It has been noted previously that besides the information on the initial state, the long-term global temperature trend due to increasing greenhouse gases contributes to the forecast performance as well [Murphy et al., 2010]—even at the seasonal time scale [Doblas-Reyes et al., 2006]. In decadal re-forecasts, the trend may turn out to be an even larger contributor to forecast quality. The predictive skill of DePreSys is thus assessed in two ways: (1) the full field case that includes the effects of the trend on NP SSTs and (2) the detrended case where the effects of the trend on NP SSTs are removed. For the full field case, the time series of the modes are computed by projecting the full forecast or observed anomalies onto the observed spatial EOF patterns. For the detrended case, the linear regression of forecast SST anomalies against the forecast GMT (dependent on the forecast year) is removed from the forecast at each grid point. The regression coefficients are estimated as a function of forecast year across all forecasts (i.e., started in 1960–2005). Regarding the observed detrended case, the...
3. Results

3.1. Decadal North Pacific SST Forecast

[15] The performance of DePreSys in the NP is considered based on the extensive set (i.e., 46) of 10-year long re-forecasts. Figure 1 (bottom) shows the anomaly correlation (AC) skill as a function of forecast time for the full time series (two columns on the left) and for the detrended time series (two columns on the right) in DePreSys Assim (first and third columns) and DePreSys NoAssim (second and fourth columns). For an illustrative example of the NPGO index forecast by this system (forecast years 1, 3, 5, 7, and 9), the reader is referred to Figure 4. This figure shows the NPGO forecasts of each member (color) vs. the observed evolution (black).

[16] The skill (solid line) of DePreSys Assim to forecast the evolution of the PDO index (Figure 1, top left of bottom part) measured by the AC is significantly positive with a 95% confidence level (shading) exclusively in the first forecast year (reaching 0.6) confirming empirical predictability studies [Newman, 2007]. The AC coefficient is statistically not distinguishable from zero in the second forecast year and for the remaining forecast times. The average skill over the 10-year forecast range (filled circle, confidence interval) is therefore very small, but positive. Without the information on the initial state, no PDO predictive skill is found in DePreSys NoAssim in any forecast year (top panel of second column). When the long-term trend in NP SSTs is removed at each grid point, a very similar result is obtained (top two panels on the right) than for the full field case. In other words, the PDO predictive skill in DePreSys is virtually equivalent with and without the trend, and the skill arises mainly from the information in the initial conditions. This agrees with the lack of a trend in Figure 3. In addition, the pattern correlation of the observed PDO vs. the observed trend pattern of 0.1 indicates that they are almost orthogonal.

[17] With regard to the predictability of the PDO in DePreSys, whether the trend is included or not, the PDO predictability in the perfect model framework where the potential skill estimate is the average correlation between swapping forecast members treated as “observations” and the ensemble mean of the remaining members is damped during the first four years of the forecast (dashed red line). The PDO persistence, computed as the projection of the observed NP anomalies from the previous year onto the PDO pattern, skill (dash-dotted black line) lies inside the uncertainty of the Assim skill, but below the potential skill for both the full and the detrended time series. This means that there is no more information in these PDO forecasts from DePreSys than from persisting the information of the previous year.

[18] In contrast to the PDO, the NPGO (Figure 1, bottom panel on the left) AC skill of DePreSys Assim is significantly positive (lower confidence interval above zero) over most of the course of a 10-year re-forecast for the full time series. The potential skill is outside the confidence interval of the actual skill indicating that there might be room for improvements in the forecast system. The Assim skill significantly beats persistence in the first two years where persistence skill is below the lower confidence interval of the actual skill. Even without any initial state information, we find significantly positive NPGO correlation skill in DePreSys NoAssim (middle panel of second column). The high predictability of the NPGO is mainly due to the long-term trend. When the effects of the trend are removed from the NP SSTs, the skill of the NPGO is reduced substantially (middle two panels on the right). For the detrended case, the AC skill averaged over the 10-year forecast range (filled circle, confidence interval) is smaller (0.11) than for the full field case (0.29). This shows that the trend is a major but not the only contributor to the skill of the NPGO. This is in line with the fact that the pattern correlation of the observed NPGO vs. the observed trend pattern is 0.6 which indicates that they are not independent.

[19] In DePreSys, the NPGO includes an artificial trend that is either inexistant or much smaller in the observations (Figure 3, bottom). This is illustrated in Figure 4, where Assim forecast (colored dots) vs. the observed values (black dots) are shown for all start dates for forecast years 1 (top), 3, 5, 7, and 9 (bottom). The DePreSys re-forecasts tend to drift toward negative NPGO values (warmer NP) in the full field case (Figure 4, left). The forecasts vary about zero in forecast year 1 (top) while they tend to be negative in forecast year 9 (bottom) after around 1990. In the detrended case (Figure 4, right), the forecasts oscillate about zero in both forecast year 1 (top) and forecast year 9 (bottom). The observed NPGO evolution (black dots) includes a small negative trend during the verification period of 51 years that is enhanced in the second half of the period. The trend in Assim (Figure 4, left), although overestimated, contributes to NPGO forecast AC skill. When the trend is removed (Figure 4, right), less NPGO forecast AC skill in Assim is found.
The potential skill of the NPGO is significantly above the actual skill in both DePreSys Assim and NoAssim. In the detrended case, however, the potential NPGO skill is statistically not distinguishable from the actual skill. The fact that in Assim, the forecasts tend to take negative NPGO values after 1990 in the full field case (Figure 4, left) in contrast to the detrended case (Figure 4, right) is the reason why we obtain positive potential skill for the full field case, because the covariance between the ensemble members of the same sign tends to be positive. This result should be interpreted as a warning in the use of potential skill estimates as upper limits of skill.

The model variants of DePreSys are not able to reasonably represent the impact of the GMT trend on NP SSTs. A comparison between the observed GMT trend in NP SST (top left) from 1960 (Nov)–2006 (Oct) and for lead year 1 in the nine model variants (other panels) is illustrated.
in Figure 5. The pattern correlations of the model variants range between 0.46 and 0.73. A similar comparison, but for forecast year 10 is provided in Figure 6. In forecast year 10, the pattern correlations of the model variants range between −0.19 and 0.69. Not only is the bias impact of GMT trend on NP SSTs present in the first forecast year, the bias is also growing throughout the forecast. This is consistent with the fact that the NPGO comprises an artificial trend in DePreSys, responsible for the high potential skill, because the biased trend in NP SSTs projects onto the NPGO. The NP GMT trend pattern in Assim tends to the NoAssim pattern as forecast time increases.

[22] Initializing the system with the observed climate state generally improves the forecast quality of both of the two NP modes in the first year of the full time series. With p-values (two-sided) of the differences of the ACs of 0.0007 (PDO) and 0.027 (NPGO), we note that the difference is significant at a 95% level or higher for the PDO and the NPGO. For the
remaining forecast years, no significant difference between Assim and NoAssim is obtained. [23] At the grid-point level, some forecast systems exhibit regions with positive or zero AC skill for temperature all across the NP in the first year [see Figure 3 in Kim et al., 2012]. Beyond the first year, however, very little or even negative AC skill is left in the central and eastern NP in most of the systems with positive skill remaining in the south western part of the basin [Kim et al., 2012; Doblas-Reyes et al., 2011; Guemas et al., 2012]. These results agree well with the skill obtained for the main NP SST modes.

3.2. Decadal Tropical Pacific SST Forecast

[24] It is well known that part of NP SST variability is of tropical origin [e.g., Trenberth et al., 1998; Alexander et al., 2002]. For monthly means, e.g., ENSO explains about 25% of the PDO variance [Lienert et al., 2011], while for annual means, ENSO explains about 31% of the PDO
The predictive skill of the PDO may therefore be in part forced by tropical Pacific SSTs or vice versa [Chiang and Vimont, 2004]. The study is thus continued by assessing the skill of the leading modes of the annual-mean tropical Pacific SST anomalies.

[25] The spatial manifestations of ENSO (first mode) and ENSO-Modoki (second mode) in the tropical Pacific are illustrated in Figure 7 (top). It is seen that the ENSO and ENSO-Modoki AC skill is virtually identical for the full field and detrended case (Figure 7, bottom). For example, the ENSO AC is distinguishable from zero until forecast year 3 for the full time series. When the trend is removed the ENSO skill is reduced slightly so that it remains significantly positive in the first year only. The skill of the ENSO-Modoki mode (in both the full field and detrended cases) in DePreSys Assim is limited to the first forecast year. The evolution of the full and detrended time series (Figure 8) confirms that the long-term trend projects onto ENSO, but not onto the second mode (ENSO-Modoki).

[26] In DePreSys Assim, the predictive skill of the main NP SST modes in forecasts years 1 and 2 may be in part due to the predictive skill of the main SST modes in the tropical Pacific. On the other hand, the positive AC skill of ENSO in year 3 in Assim independent of the trend may be in part due to the response of the tropical Pacific to, e.g., the NPGO in the previous year where Assim shows significantly positive AC skill beyond the trend [Vimont et al., 2001].
4. Conclusions and Discussion

[27] This study has analyzed the decadal performance of initialized vs. uninitialized coupled climate prediction systems to reforecast the leading two EOF modes of annual-mean NP and tropical Pacific SST departures for (1) the full time series and (2) for the detrended time series where the long-term trend is removed from the grid-point SST. The results show information in the initial state improves skill of decadal PDO forecasts only in the first year. The NPGO is skillful over the course of a 10-year forecast, but this skill results almost exclusively from predictions of the trend. In the detrended case, most of the skill of the NPGO is limited to the first two years. This result is not documented anywhere else for climate forecast systems with dynamical models. On the other hand, the PDO skill is not affected by the long-term trend.

[28] In the tropical Pacific, decadal ENSO (ENSO-Modoki) forecasts are skillful exclusively in the first three years (first year) due to information in the initial state. The trend plays a minor (no) role for the predictive skill of ENSO (ENSO-Modoki). The low PDO skill is line with the skill of ENSO—one of the sources of PDO predictability—that is limited to the initial few years.

[29] The initialized climate forecast system fails to capture the impact of GMT on annual-mean NP SSTs. This bias increases with forecast time as the model GMT trend in NP SSTs tends to the one in the uninitialized system. This introduces a bias in the forecast evolution of the NPGO. The result that DePreSys is not capable to capture the GMT trend on NP SSTs tends to the one in the uninitialized system. This introduces a bias in the forecast evolution of the NPGO. The result that DePreSys is not capable to capture the GMT trend on NP SSTs tends to the one in the uninitialized system. This introduces a bias in the forecast evolution of the NPGO. The result that DePreSys is not capable to capture the GMT trend on NP SSTs tends to the one in the uninitialized system. This introduces a bias in the forecast evolution of the NPGO. The result that DePreSys is not capable to capture the GMT trend on NP SSTs tends to the one in the uninitialized system. This introduces a bias in the forecast evolution of the NPGO. The result that DePreSys is not capable to capture the GMT trend on NP SSTs tends to the one in the uninitialized system. This introduces a bias in the forecast evolution of the NPGO. The result that DePreSys is not capable to capture the GMT trend on NP SSTs.


