



Assessing the skill of the Pacific Decadal Oscillation (PDO) in a decadal prediction experiment

George J. Boer¹ · Reinel Sospedra-Alfonso¹

Received: 30 July 2018 / Accepted: 11 July 2019 / Published online: 22 July 2019
© The Author(s) 2019

Abstract

A modified approach to the assessment of the prediction skill of “modes of variability” is proposed and applied to a decadal prediction experiment. In particular, the skill of predicting the Pacific Decadal Oscillation (PDO) is investigated. The approach depends on separately calculating the EOFs of the observations, the ensemble of forecasts, and an ensemble of simulations made with the same model and external forcing. The skill of predicting and simulating the spatial structure of the modes is captured by comparing forecast and simulated EOFs with the observation-based EOFs. This is in contrast to the case where forecasts and simulations are expanded in observation-based EOFs, or other structure functions, which gives no direct information about the model-based EOF structures themselves. The skill of predicting the temporal evolution of EOFs is separately captured by comparing the associated expansion functions. Finally, the contribution of the modes to the overall prediction skill is obtained by weighting the spatial and temporal skills with the variances involved. The behaviour of the first mode, identified as the PDO, is given particular attention. Perhaps not unexpectedly, the EOF structure of the forecasts more closely resembles that of the simulations than that of the observations, but both reproduce the structure of the observed PDO quite well with spatial correlations near 0.8. The temporal correlation of the expansion functions is near 0.7 for year 1 forecasts and declines toward zero subsequently. The overall correlation skill for the North Pacific is dominated by the PDO with a small contribution from the second mode and none from the third mode.

Keywords Decadal prediction · Pacific Decadal Oscillation · PDO · Modal skill · CanCM4 decadal hindcasts

1 Introduction

Many climate studies investigate “modes” of variability which are typically identified in terms of statistical constructs. Modes are often identified by indices (e.g. ENSO, AMV, NAM, SAM etc.) or, somewhat more generally, in terms of empirical orthogonal functions (EOFs), rotated EOFs, non-linear EOFs etc. (e.g., Luo and Yamagata 2002, and the references therein). The modes, so identified, are one way of organizing observed and model data in an “economical” way which may also lead to their interpretation and investigation in terms of physical processes and structures.

Prediction on all timescales is a feature of physically-based theories and climate prediction on decadal timescales is becoming an operational activity (Kushnir et al. 2018). For climate, both predictability (i.e. the ability to be predicted) as well as the forecast skill (i.e. the current ability to forecast the evolution) are of interest (Boer 2004; Smith et al. 2007; Murphy et al. 2010; Doblas-Reyes et al. 2013; Boer et al. 2013; Cassou et al. 2018).

A modified approach to the assessment of the skill of “modes of variability” is proposed and applied to the case of the Pacific Decadal Oscillation (PDO) based on the output of the Canadian Centre for Climate Modelling and Analysis (CCCma) multi-year decadal prediction experiment (Kharin et al. 2012; Boer et al. 2013, 2018a, b). In particular, the spatial structures and temporal evolution of the modes of variability of the observations, the ensemble of predictions, and an ensemble of simulations of sea surface temperature (SST) are separately calculated, compared, and assessed with the first mode identified as the PDO. This approach avoids the expansion of all three data sources in terms of

✉ Reinel Sospedra-Alfonso
reinel.sospedra-alfonso@canada.ca

George J. Boer
gjb.cccma@gmail.com

¹ Canadian Centre for Climate Modelling and Analysis,
Environment and Climate Change Canada, Victoria, BC,
Canada

a single set of spatial basis functions and thereby retains their own “natural” modes of variability so to say. In this approach, the skill of predicting North Pacific SSTs has contributions depending on the agreement between the spatial structures and temporal evolutions of the modes and their contribution to the overall SST variability.

2 The PDO

Mantua et al. (1997) review the then state of knowledge of the PDO and provide many references. They define the PDO as the leading EOF of monthly SST anomalies in the North Pacific Ocean poleward of 20°N. The calculation is applied to anomalies where monthly mean global average SSTs are subtracted out in order to remove (at least most of) the global warming signal in the data. We adopt this approach, although applied to annually averaged SSTs, in order to be reasonably consistent with other investigations and also because there is no an obvious better way to do this.

Mantua and Hare (2002) in a subsequent review discuss Pacific Decadal Variability (PDV) and characterize the PDO as the “dominant pattern of PDV”, and refer to other studies of PDV including those identified as the Interdecadal Pacific Oscillation (IPO, Power et al. 1999) and the North Pacific Oscillation (NPO, Gershunov and Barnett 1998). Mantua and Hare (2002) also review some of the studies of the processes thought to be involved in the PDO and in PDV. Newman et al. (2016) “revisit” the PDO, characterized as the dominant year-round pattern of monthly North Pacific SST variability, and argue that it is a combination of physical processes and time scales operating both locally and forced remotely from the tropics.

While the PDO is defined for the North Pacific, it is viewed more widely by means of the PDO “index” which is the principal component (PC) associated with the leading EOF pattern. The regression of this index against other variables such as SST (e.g., Figure 1, Mantua et al. 1997; Newman et al. 2016; Cassou et al. 2018), sea level pressure and surface wind stress suggests connections between the North Pacific, the broader Pacific, and more remote regions.

3 Predicting the PDO

There is an increasing number of studies of the skill of prediction on time scales of a year or more generally referred to as decadal predictions. The results of some of these decadal prediction studies are surveyed in Chapter 11 of the IPCC 2013 (Kirtman et al. 2013). One of the activities of the Decadal Climate Prediction Project (DCPP) is a multi-model hindcast study of decadal prediction as a contribution to the Coupled Model Intercomparison Project phase 6 (CMIP6)

(Boer et al. 2016). An on-going collaborative effort produces temperature predictions for the next several years (Smith et al. 2013) while the WCRP Grand Challenge on near-term climate prediction (e.g., Kushnir et al. 2018) is providing the ground work for the operationalization of such forecasts.

Prediction studies need, as a basis, a sequence of retrospective forecasts or “hindcasts”, which make use of the data available at the time the hindcast is initialized, and which are verified against subsequent observations. The hindcasts permit the calculation of skill measures that are a necessary part of the utility and success of any forecasting system.

The PDO is an important component of long-term temperature variability characterized by its spatial structure (the leading EOF of North Pacific sea surface temperature) and temporal evolution (the associated principal component or PDO index). Many studies assess the ability of freely running climate models to reproduce the behaviour of the PDO although not its predictive skill. As discussed by Newman et al. (2016) for instance, the structure of the simulated PDO is considered and compared with that based on observations. Results from CMIP5 model simulation demonstrate considerable scatter with the majority of spatial correlation values ranging between about 0.6 and 0.9. Newman et al. also compare some aspects of the temporal behaviour of the PDO indices noting that, because of the chaotic nature of the climate system, simulations differ from one another and from observations. In the absence of a dominant externally forced signal there is no expectation that the simulated index should agree in time with that of the observations. A prediction of the PDO, as opposed to its simulation, should reproduce both the spatial structure of the mode as well as the temporal evolution of its index.

Studies assessing the skill of predictions of the PDO come in several flavors. One aspect of PDO forecasts considers the prediction of sign changes in the index (transitions between warm and cold events or the reverse) as in Meehl et al. (2016) and references therein. Another approach considers the overall skill of predicting the evolution of the PDO from year to year (e.g., Kim et al. 2014; Mehta et al. 2019, and the references therein). The latter approach is adopted here when assessing overall PDO prediction skill. Some studies also look at skill measures of SSTs evaluated at grid points or as averages over the North Pacific region (Mochizuki et al. 2010; Guemas et al. 2012; Kim et al. 2014) and this is also part of the analysis here.

The evaluation of the overall skill of the PDO may be approached in several ways. One approach is to expand both the forecasts and the verification data in terms of the EOFs of the latter and to assess the skill of the associated PDO indices. Another approach expands the verification data and the forecasts in terms of their own EOFs and compares the PDO indices. The structure of the EOFs and the temporal behaviour of the PDO indices are confounded in these

approaches and the consequences of differences between the structure of the forecast PDO and its temporal evolution are not distinct.

In the approach developed here, the ability of the forecasting system to skilfully predict both the structure of the PDO (and other North Pacific modes of SST variability) and their temporal evolution is explicitly calculated. Moreover, while the PDO is a dominant mode of North Pacific temperature variability, the skill of the PDO itself represents only a portion of the overall predictive skill for North Pacific SST and the contribution of the PDO (and other modes) to regional skill is also assessed.

4 Data and statistics

The results analyzed here are from a decadal forecast experiment consisting of ensembles of ten forecasts initialized near the end of each year from 1960 to 2015 together with an ensemble of ten historical climate change simulations begun from the preindustrial period. The same external forcing parameters are specified for both the forecasts and the simulations. The initial conditions for the ten ensemble members are a combination of information from independent coupled model “assimilation” runs and ocean reanalyses as described in Merryfield et al. (2013). The experiment and some of the prediction results are described in Boer et al. (2013). Annual mean SST is the variable of interest in this study. Observation-based SST verification data is from version 5 of the US National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed Sea Surface Temperature (ERSSTv5, Huang et al. 2017). The observations, forecast and simulations for the period 1971 to 2015 are used in this analysis. This is consistent with the recommendation of the DCP (Boer et al. 2016), to keep the period of analysis fixed and as long as possible for all forecast ranges.

The statistical significance of the overall correlation skill of forecasts and simulations of North Pacific SST, and of the contributions of the PDO and other modes of variability, is assessed using a non-parametric bootstrap procedure. Following Wilks (1997), a moving block bootstrap approach is employed in order to account for temporal and spatial correlations in the data. With this procedure, for every forecast range and separately for the simulations, 1000 samples are generated by randomly drawing b block pairs (with replacement) of l consecutive years of observed and modelled fields from the 45-year pool of all year-overlapping l -block pairs (each draw is done once for the observations and for each ensemble member separately). The EOF expansions described in the next section are applied to the resulting datasets and skill measures are computed for the samples.

For every skill measure, the 5%- and 95%-quantile estimates of all samples are used as the 90% confidence interval. The computations shown here correspond to b equal the nearest integer to $45 / l$ with blocks of length $l = 7 \approx \sqrt{45}$ years (Wilks 1997), although the results are not very sensitive to the block size.

5 Approach

Forecasts of the climate system attempt to trace out the actual evolution of the climate system from some observation-based initial state. A climate forecast will exhibit forced and internally generated variability where the forced component of the forecasts should be at least similar to that of the simulations when the same external forcing is specified for both as is the case here. For the purpose of this analysis, anomalies of annual mean SST from the long-term average are considered at each forecast range and assessed against those from observations. Removing the long-term means from the data serves as a first order bias correction as recommended by Boer et al. (2016). Also, following Mantua et al. (1997) (see also Mantua and Hare 2002), the forced component is removed, at least partially, by further subtracting out the global mean SST anomaly for each year at each grid cell.

To the extent that the forced component is removed, the forecast and simulated anomalies represent internally generated components of variability. A modal representation of this variability is developed by means of an EOF expansion and the forecast quality of the modes is investigated in the context of multi-annual forecasts. Observation-, forecast- and simulation-based anomalies, $X(\lambda, \phi, t)$, $Y_k(\lambda, \phi, t)$, $U_k(\lambda, \phi, t)$ respectively, are functions of location (λ, ϕ) , initial time (t_j) , and forecast range τ , for $t = t_j + \tau$ where τ labels yearly averaged forecasts for years 1 to 10. The index k identifies the ensemble member in the case of ensembles of forecasts and simulations. For notational convenience, the dependencies on location, start time and forecast range are simplified as

$$X(\lambda, \phi, t) \Leftrightarrow X(\lambda, \phi, t_j, \tau) \Leftrightarrow X_{j\tau}(\lambda, \phi) \Leftrightarrow X_{j\tau} \Leftrightarrow X$$

and similarly for the ensembles of forecasts and simulations. Using this notation, the observations, forecasts, and simulations are expressed as

$$X = \chi + x \quad (1)$$

$$Y_k = \psi + y_k \quad (2)$$

$$U_k = u_k \quad (3)$$

where ψ , χ are internally generated predictable components and x , y , u unpredictable “noise” components (Boer et al. 2013).

In the forecasts, the internally generated components are not independent, as they are in the simulations, but depend on initial conditions and the ability of the model to predict the behaviour of the climate system. The predictable part of this variability, represented as ψ , is common across an ensemble of forecasts while the remaining unpredictable or noise components y_k will evolve differently and independently and will average out across a large enough ensemble. The actual system has analogous predictable χ and noise x variability. By contrast, the simulations U_k are not initialized from observation-based information and, provided the forced component is removed, represent unforced and unpredicted internal variability u_k .

An ensemble mean forecast or simulation is represented by a subscript A or by curly brackets with

$$Y_A = \{Y_k\} = \frac{1}{m} \sum_{k=1}^m Y_k = \psi + \frac{1}{m} \sum_{k=1}^m y_k = \psi + \{y_k\} \rightarrow \psi \tag{4}$$

$$U_A = \{U_k\} = \frac{1}{m} \sum_{k=1}^m U_k = \frac{1}{m} \sum_{k=1}^m u_k = \{u_k\} \rightarrow 0 \tag{5}$$

where m is the ensemble size. Ensemble averaging does not affect the predictable component ψ , which is the same for each ensemble member, while the unpredictable components y_k, u_k average to zero for a large enough ensemble as indicated by the arrows.

Under these circumstances, ensemble averaging is an operational way of identifying the predictable component of the forecast. The ensemble mean forecast is typically more skillful than an individual forecast, at least in terms of correlation, mean square error (MSE) and mean square skill score (MSSS), since the noise in the forecast is suppressed by the averaging in (4) and degradation of the scores is thereby, at least partially, avoided.

5.1 Observation-based EOFs

For the sequence of observations X , the EOF analysis proceeds by expressing the grid of SSTs over the Pacific area of interest from 20°N to 60°N as a column vector \mathbf{X} . The sample covariance matrix

$$C_X = \overline{\mathbf{X}\mathbf{X}^T}$$

is estimated from the data and is the basis of the EOF analysis, where the overbar is the average over start times t_j . The eigenvectors \mathbf{e}_α of the covariance matrix are related to the eigenvalues η_α as

$$C_X \mathbf{e}_\alpha = \eta_\alpha \mathbf{e}_\alpha$$

The eigenvectors are represented geographically as $e_\alpha(\lambda, \varphi)$ and are orthonormal $\langle e_\alpha e_\beta \rangle = \delta_\alpha^\beta$ with respect to the area

average, represented by angular brackets, over the PDO area of analysis. The observations may then be expanded in terms of these EOFs with

$$\begin{aligned} X(\lambda, \varphi, t) &= \sum_{\alpha} a_{\alpha}(t) e_{\alpha}(\lambda, \varphi) \\ \sigma_X^2 &= \langle X^2 \rangle = \sum_{\alpha} \sum_{\beta} \overline{a_{\alpha} a_{\beta}} \langle e_{\alpha} e_{\beta} \rangle \\ &= \sum_{\alpha} \overline{a_{\alpha}^2} = \sum_{\alpha} \sigma_{a_{\alpha}}^2 \end{aligned} \tag{6}$$

where σ_X^2 is the overall sample variance and $\sigma_{a_{\alpha}}^2$ is the variance of the individual modes. The expansion functions a_{α} are the principal components (PCs) and are obtained from the original fields X as

$$a_{\alpha} = \langle e_{\alpha} X \rangle.$$

The expansion functions contain the temporal information and the EOFs the spatial information of the modes. The variance accounted for by the first 3 modes of the observations is shown in the upper panel of Fig. 1 along with that of the forecasts and simulations as discussed subsequently. The first two observation-based EOFs e_{α} of North Pacific SST, the first of which is the PDO, are displayed in Fig. 2. In what follows we concentrate mainly on the PDO although some information on the second and third modes is also included. The second mode of North Pacific variability has been

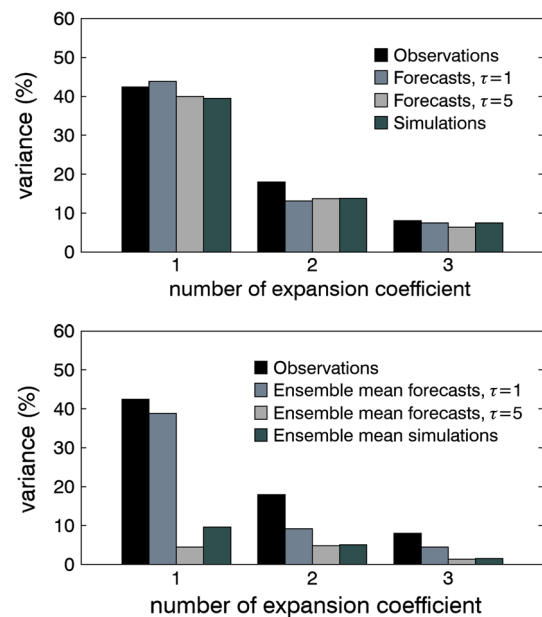
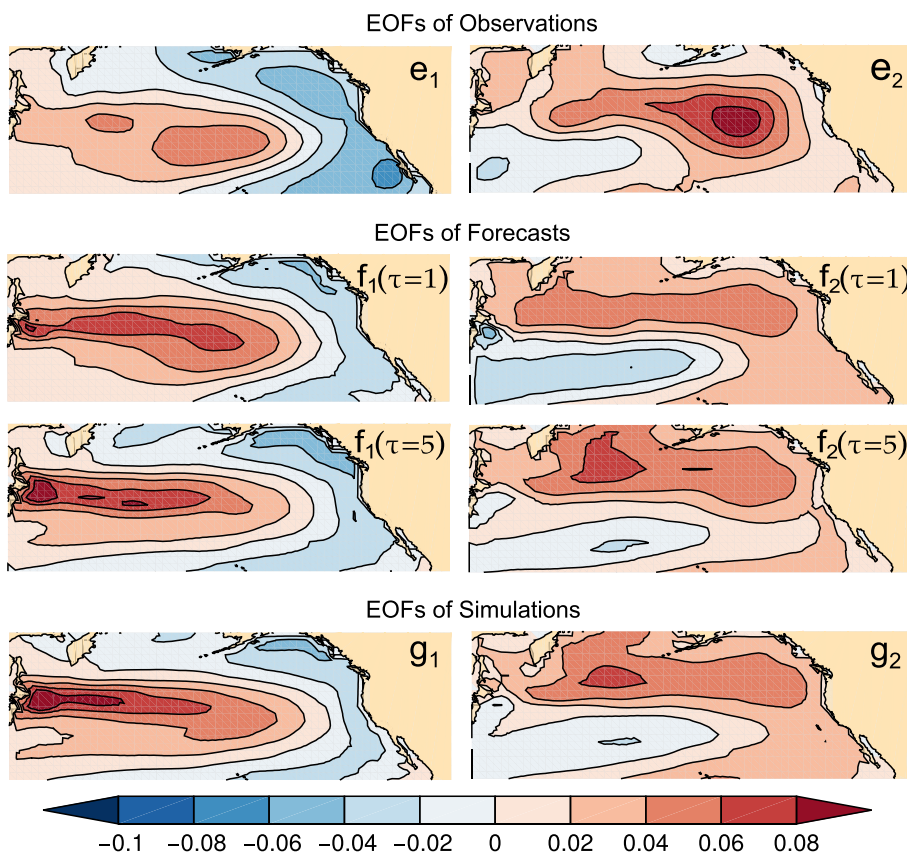


Fig. 1 Upper panel: Fraction of the total variance accounted for by the first three modes of the observations (black), forecasts for $\tau =$ year 1 (dark gray) and $\tau =$ year 5 (light gray), and for the simulations (olive). Bottom panel: As upper panel but for the ensemble mean of the forecasts and simulations

Fig. 2 The spatial structure of the first two EOFs of the observations (e_1, e_2), of the forecasts at range $\tau =$ year 1 ($f_1(\tau = 1), f_2(\tau = 1)$) and $\tau =$ year 5 ($f_1(\tau = 5), f_2(\tau = 15)$), and of the simulations (g_1, g_2)



studied in some detail elsewhere (e.g., Bond et al. 2003). The third and subsequent components are not well separated, have little variance, and as will be seen, are unskillful.

5.2 Forecast-based EOFs

For an ensemble of forecasts at range τ the associated elements of the covariance matrix are the ensemble average, indicated by $\{.\}$, of the covariances of the individual forecasts. Expressing the grid of forecasts as column vectors, \mathbf{Y}_k , the covariance matrix is

$$C_Y(\tau) = \left\{ \overline{\mathbf{Y}_k \mathbf{Y}_k^T} \right\}$$

with associated eigenvectors \mathbf{f}_α and eigenvalues μ_α with

$$C_Y \mathbf{f}_\alpha = \mu_\alpha \mathbf{f}_\alpha.$$

Here the EOFs $f_\alpha = f_\alpha(\lambda, \varphi, \tau)$ are functions also of forecast range τ and

$$\begin{aligned} Y_k &= \sum_\alpha b_{k\alpha}(t_j, \tau) f_\alpha(\lambda, \varphi, \tau) \\ b_{k\alpha} &= \langle f_\alpha Y_k \rangle \\ \sigma_Y^2 &= \left\langle \left\{ \overline{Y_k^2} \right\} \right\rangle = \sum_\alpha \sum_\beta \left\{ \overline{b_{k\alpha} b_{k\beta}} \right\} \langle f_\alpha f_\beta \rangle \\ &= \sum_\alpha \left\{ \overline{b_{k\alpha}^2} \right\} = \sum_\alpha \left\{ \sigma_{b_{k\alpha}}^2 \right\} \end{aligned} \tag{7}$$

Of course the typical forecast is an ensemble mean forecast which averages out some of the noise thereby improving the scores. For the ensemble mean forecast,

$$Y_A = \{Y_k\} = \sum_\alpha \{b_{k\alpha}\} f_\alpha = \sum_\alpha b_\alpha f_\alpha \tag{8}$$

$$b_\alpha = \langle f_\alpha Y_A \rangle = \{b_{k\alpha}\} \tag{9}$$

$$\sigma_{Y_A}^2 = \left\langle \overline{Y_A^2} \right\rangle = \sum_\alpha \overline{b_\alpha^2} = \sum_\alpha \sigma_{b_\alpha}^2. \tag{10}$$

The ensemble mean forecast Y_A is represented in terms of the EOFs of the ensemble of forecasts f_α , not of the ensemble mean forecast itself, where b_α is the expansion function. The b_α 's are not the PCs of the ensemble mean forecast.

There are three points here. Firstly, the forecast-based EOFs f_α need not be the same as the observation-based EOFs e_α and that differences between them should be considered when considering the skill of the modes. Secondly, the EOFs f_α are functions of forecast range and so are expected to evolve as τ increases. Lastly, the forecast EOFs are not the same as the EOFs of the ensemble mean forecast. The percentage of the variance accounted for by the leading modes of the ensemble of forecasts are shown in the upper panel of Fig. 1 for $\tau = 1$ and 5 years. The first two forecast EOFs are plotted in the middle panels of Fig. 2.

5.3 Simulation-based EOFs

Paralleling the forecasts-based relations in Sect. 5.2, the corresponding relations for the simulations are

$$\begin{aligned} C_U &= \left\{ \overline{U_k U_k^T} \right\} \\ C_U g_\alpha &= v_\alpha g_\alpha \\ U_k &= \sum_\alpha c_{k\alpha}(t_j) g_\alpha(\lambda, \varphi) \end{aligned} \quad (11)$$

$$\begin{aligned} c_{k\alpha} &= \langle g_\alpha U_k \rangle \\ \sigma_U^2 &= \left\langle \left\{ \overline{U_k^2} \right\} \right\rangle = \sum_\alpha \sum_\beta \left\{ \overline{c_{k\alpha} c_{k\beta}} \right\} \langle g_\alpha g_\beta \rangle \\ &= \sum_\alpha \left\{ \overline{c_{k\alpha}^2} \right\} = \sum_\alpha \left\{ \sigma_{c_{k\alpha}}^2 \right\} \end{aligned} \quad (12)$$

and, for the ensemble average simulation,

$$U_A = \sum_\alpha c_\alpha g_\alpha \quad (13)$$

$$c_\alpha = \langle g_\alpha U_A \rangle = \{c_{k\alpha}\} \quad (14)$$

$$\sigma_{U_A}^2 = \left\langle \overline{U_A^2} \right\rangle = \sum_\alpha \overline{c_\alpha^2} = \sum_\alpha \sigma_{c_\alpha}^2. \quad (15)$$

Figures 1 and 2 also display the variance fractions and the first 2 simulation-based EOFs g_α .

6 EOFs and expansion functions

Observations, forecasts, and simulations are denoted by X, Y_k, U_k with associated EOFs $e_\alpha, f_\alpha, g_\alpha$ carrying the spatial patterns and expansion functions $a_\alpha, b_{k\alpha}, c_{k\alpha}$ carrying the time information. The fraction of the overall variances associated with the first 3 modes of the expansions of the annual mean temperatures of the observation-based data, the forecasts, and the simulations are shown in the upper panel of Fig. 1. The first mode (i.e. the PDO) captures over 40% of the variance, the second mode captures about half that, and the third mode half as much again. In the upper panel, the forecast and simulated modes have variance fractions which are similar to those based on the observations so are successful in that respect.

The lower panel of Fig. 1 again plots variance fractions but here the variance fractions are based on the ensemble means of the forecasts and simulations. Following (7) and (10), the variance fractions for the forecast PDO mode in the upper and lower panels have the form $\left\{ \sigma_{b_{k1}}^2 \right\} / \sigma_Y^2$ and $\sigma_{b_1}^2 / \sigma_Y^2$, respectively, and similarly for the simulated PDO

mode following (12) and (15). The ensemble mean is a rough measure of the predictable fraction of the variance and is expected to decline from $\tau = \text{year 1}$ to $\tau = \text{year 5}$ as prediction skill wanes, as is seen for the PDO in the lower panel of Fig. 1.

If the forced component is successfully removed from the simulation results, the PDO is that of the freely running model. As simulations are uninitialized they lack a predictable component and the ensemble mean should tend to zero for a large ensemble average. The variance fraction of the ensemble mean in the lower panel is much reduced from the overall variance fraction in the upper panel for the PDO as expected. As forecast range increases, the effect of initial conditions declines and, at long enough range, a forecast is expected to be indistinguishable from a simulation. This is consistent with the PDO variance fraction for the simulations (olive bar) being similar to that of the forecasts at $\tau = \text{year 5}$ (light gray bar) in the lower panel. The variance fraction of the PDO in the simulations is modestly larger than that of the year 5 forecasts which suggests that there may remain some synchronization in the simulated PDOs across the ensemble. Removing the global mean SST from the Pacific SSTs may not be entirely effective at removing the forced signal arising from greenhouse gas forcing or other sources such as volcanoes (Malik et al. 2018) or anthropogenic aerosols (Boo et al. 2015; Smith et al. 2016; Frankignoul and Gastineau 2017), and this is suggested also in subsequent results in Fig. 5.

The first two EOFs of the observations, forecasts and simulations ($e_\alpha, f_\alpha, g_\alpha$) are plotted in Fig. 2. The EOFs of the forecasts f_α are functions of forecast range τ while those for the observations and simulations are not. The expansion functions for the observations and the ensemble mean forecasts and simulations ($a_\alpha, b_\alpha, c_\alpha$) are plotted in Fig. 3 since it is the ensemble mean forecast that is verified against observations when using the standard skill measures of correlation, MSE and MSSS. Differences between observed and modelled EOFs and between expansion functions contribute to differences in the skill of the modes as discussed subsequently.

7 Modal statistics and skill measures

The second order statistics of interest are obtained by “lining up” the forecasts and observations with respect to forecast range τ and calculating the statistics by averaging over all start dates t_j . Where skill is assessed by deterministic measures, the ensemble mean forecast is used since ensemble averaging acts to suppress the noise in the forecasts as in (4) above.

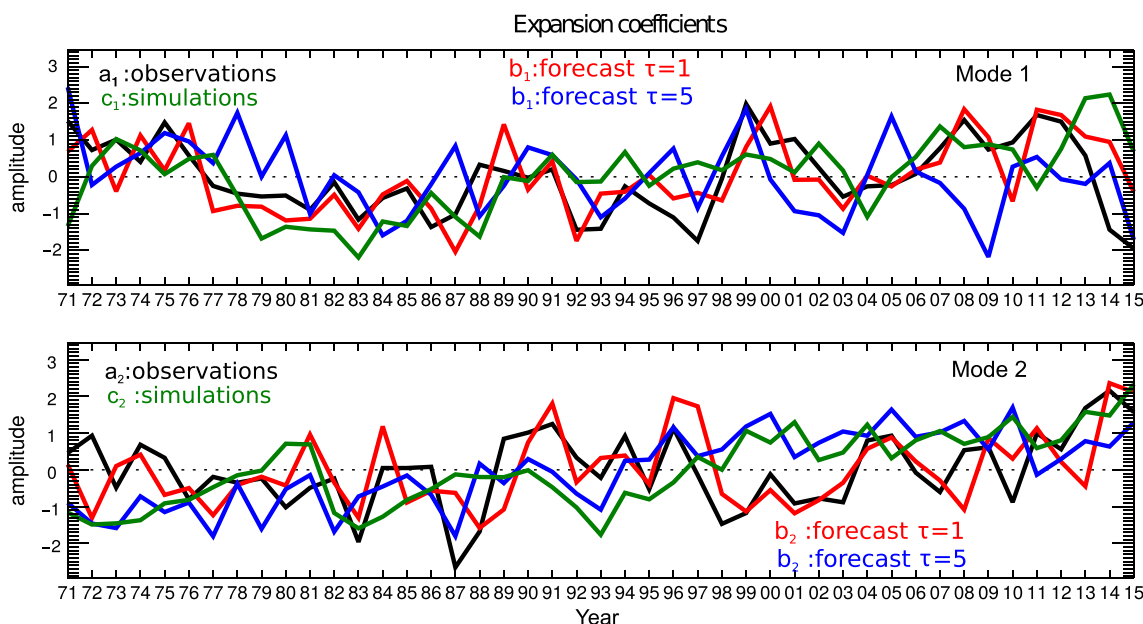


Fig. 3 The temporal evolution of the normalized expansion coefficients for the first two EOFs of the observations (black), of the ensemble mean forecasts at forecast range $\tau = \text{year } 1$ (red) and $\tau = \text{year } 5$ (blue), and of the ensemble mean of the simulations (green)

7.1 Variances and covariances

Averages over the North Pacific area of interest are indicated by angular brackets. Averaged variance components for the observations and ensemble mean forecasts and simulations are rewritten from (6, 10, 15) as

$$\sigma_X^2 = \langle \overline{X^2} \rangle = \sum_{\alpha} \overline{a_{\alpha}^2} + \langle \overline{X_*^2} \rangle = \sum_{\alpha} \sigma_{a_{\alpha}}^2 + \sigma_{X_*}^2 \tag{16}$$

$$\sigma_{Y_A}^2 = \langle \overline{\{Y_k\}^2} \rangle = \sum_{\alpha} \overline{b_{\alpha}^2} + \langle \overline{Y_*^2} \rangle = \sum_{\alpha} \sigma_{b_{\alpha}}^2 + \sigma_{Y_*}^2 \tag{17}$$

$$\sigma_{U_A}^2 = \langle \overline{\{U_k\}^2} \rangle = \sum_{\alpha} \overline{c_{\alpha}^2} + \langle \overline{U_*^2} \rangle = \sum_{\alpha} \sigma_{c_{\alpha}}^2 + \sigma_{U_*}^2 \tag{18}$$

where it is now understood that the summations retain only the first few modes of interest and higher order terms, indicated by a star, are carried as a remainder. The standard deviations of the expansion functions in (16)–(18) are listed in Table 1 for the first three modes and for the total (where the total standard deviation is the square root of the sum of the variances over all modes). As expected σ_{Y_A} is reasonably close to σ_X at $\tau = \text{year } 1$ but approaches σ_{U_A} , the value for the free running model simulations, at later τ .

The covariance between observations X and ensemble mean forecast Y_A is

$$\begin{aligned} C_{XY_A} &= \langle \overline{XY_A} \rangle \\ &= \sum_{\alpha\beta} \overline{a_{\alpha}b_{\beta}} \langle e_{\alpha}f_{\beta} \rangle + C_* \\ &= \sum_{\alpha\beta} \sigma_{a_{\alpha}}\sigma_{b_{\beta}} r(a_{\alpha}, b_{\beta})s(e_{\alpha}, f_{\beta}) + C_* \\ &= \sum_{\alpha} C_{\alpha} + \hat{C} \end{aligned} \tag{19}$$

where

$$\begin{aligned} C_{\alpha} &= \sigma_{a_{\alpha}}\sigma_{b_{\alpha}} r(a_{\alpha}, b_{\alpha})s(e_{\alpha}, f_{\alpha}) \\ \hat{C} &= \sum_{\alpha \neq \beta} \sigma_{a_{\alpha}}\sigma_{b_{\beta}} r(a_{\alpha}, b_{\beta})s(e_{\alpha}, f_{\beta}) + C_* \end{aligned} \tag{20}$$

and

$$r(a_{\alpha}, b_{\beta}) = \frac{\overline{a_{\alpha}b_{\beta}}}{\sigma_{a_{\alpha}}\sigma_{b_{\beta}}} \tag{21}$$

$$s(e_{\alpha}, f_{\beta}) = \langle e_{\alpha}f_{\beta} \rangle. \tag{22}$$

Here $r(a_{\alpha}, b_{\alpha})$ is the temporal correlation between the expansion functions of a given mode and $s(e_{\alpha}, f_{\alpha})$ is the spatial correlation between the EOFs. The remainder term \hat{C} has contributions from the “off diagonal” ($\alpha \neq \beta$) terms in the summation as well as from

$$C_* = \langle \overline{X_* Y_*} \rangle + \sum_{\alpha} \langle \overline{a_{\alpha} Y_* e_{\alpha}} \rangle + \sum_{\beta} \langle \overline{b_{\beta} X_* f_{\beta}} \rangle$$

the “left over” variability terms in (19)–(20) which need not be zero. Ideally \hat{C} is small. Similar relationships hold for the covariance C_{XU_A} between the observations and simulations.

7.2 Skill measures

The overall correlation skill for the North Pacific region is written as

$$R = \frac{C_{XY_A}}{\sigma_X \sigma_{Y_A}} = \sum_{\alpha} \frac{C_{\alpha}}{\sigma_X \sigma_{Y_A}} + \frac{\hat{C}}{\sigma_X \sigma_{Y_A}} = \sum_{\alpha} R_{\alpha} + R_* \tag{23}$$

where the skill of mode α is characterized as

$$R_{\alpha} = \left(\frac{\sigma_{a_{\alpha}}}{\sigma_X} \right) \left(\frac{\sigma_{b_{\alpha}}}{\sigma_{Y_A}} \right) r(a_{\alpha}, b_{\alpha}) s(e_{\alpha}, f_{\alpha}) = w(a_{\alpha}, b_{\alpha}) r(a_{\alpha}, b_{\alpha}) s(e_{\alpha}, f_{\alpha}) \tag{24}$$

and R_* is the remainder term. The factor

$$w(a_{\alpha}, b_{\alpha}) = \left(\frac{\sigma_{a_{\alpha}}}{\sigma_X} \right) \left(\frac{\sigma_{b_{\alpha}}}{\sigma_{Y_A}} \right), \tag{25}$$

weights the product of the spatial and temporal correlations to give R_{α} the contribution of the mode to the overall correlation R for the PDO area.

These expressions encapsulate several aspects of modal skill as developed here. R is the correlation for the PDO area with the correlation R_{α} of the retained modes and with contribution R_* representing any left over correlation beyond the retained modes as well as from the “miss-projection” of modes β on modes α . Each of the terms in (23)–(24) gives some information about the skill of the PDO, and of any other modes considered.

Although we concentrate on correlation skill (R), a similar approach applies to the MSE (E) and MSSS (M) scores for an ensemble mean forecast expressed as the sum of the contributions from the α th observation-based and forecast modes plus what is “left over” with

$$E = \sigma_X^2 + \sigma_{Y_A}^2 - 2C_{XY_A} = \sum_{\alpha} E_{\alpha} + E_*$$

$$M = 1 - \frac{E}{\sigma_X^2} = \sum_{\alpha} M_{\alpha} + M_*$$

The various statistics pertinent to the skill are summarized in Table 2.

The values of R_{α} and M_{α} will be small and that of E_{α} large if the observation-based and forecast-based EOFs are spatially disjoint (small $s(e_{\alpha}, f_{\alpha})$) and/or if their expansion functions are temporally disjoint (small $r(a_{\alpha}, b_{\alpha})$) and/or if the weight function $w(a_{\alpha}, b_{\alpha})$ is small. Considering $r(a_{\alpha}, b_{\alpha})$, $s(e_{\alpha}, f_{\alpha})$ and $w(a_{\alpha}, b_{\alpha})$ separately conveys information about the relationships between the observation-based and forecast-based modes. The statistics in Tables 3 and 4 aid in analyzing and understanding the forecast quality of the modes. The relative size of the remainder compared to the diagonal terms gives another indication of the suitability of an EOF-based forecast analysis.

8 Correlation skill

The terms in the correlation skill decomposition in (23-24) are evaluated from observations and the results of a decadal prediction experiment and of climate change simulations.

8.1 EOF pattern skill

The first EOF, representing the observation-, forecast-, and simulation-based PDO in Fig. 2 displays a family resemblance with f_1 visually resembling g_1 somewhat more than e_1 . In other words, even at year 1 of the forecasts the PDO spatial structure partakes of the PDO structure of the freely running model simulations. The agreement between the observation-based EOF patterns and the forecast-based EOFs is measured by $s(e_{\alpha}, f_{\alpha})$ in (22) and plotted in Fig. 4 as a function of forecast range. The agreement between the observation-based EOF patterns and the simulations-based EOFs is measured by $s(e_{\alpha}, g_{\alpha})$ and is also plotted in Fig. 4.

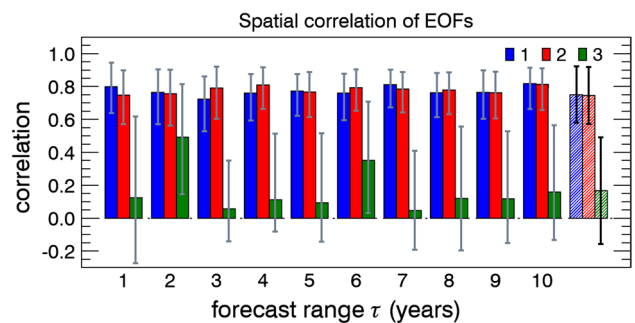


Fig. 4 The spatial correlations $s(e_{\alpha}, f_{\alpha})$ between the observation-based EOF structures e_{α} and the forecast-based EOF structures f_{α} are plotted for the first three modes as blue, red and green solid bars respectively. The corresponding correlations $s(e_{\alpha}, g_{\alpha})$ of observation-based and simulated EOF structures are plotted as hatched bars. The 90% confidence intervals are indicated

The point here is that model and observation-based EOF patterns need not be the same and that differences between them imply a loss of skill when $s(e_\alpha, f_\alpha)$ is less than one. This aspect of the skill at predicting the PDO is not apparent when both observations and forecasts are expanded in terms of observation-based EOFs for instance. The spatial correlations between the EOF structures is rather constant for the first two EOFs, but not for the third EOF which lacks statistical significance, as seen in Fig. 4. The expectation that f_1 would evolve from a structure similar to e_1 at $\tau = \text{year } 1$ to a structure more resembling g_1 at $\tau = \text{year } 5$ is not as dramatic as might be expected.

More generally, if $s(e_\alpha, f_\alpha)$ is small, it is apparent that the α^{th} forecast- and observation-based modes are disjoint and the predictive skill of mode α is low. The $s(e_\alpha, f_\beta)$ matrix also gives an indication of the usefulness of analyzing forecast skill in terms of EOFs. Only if “miss-projection” is reasonably small, i.e. if $s(e_\alpha, f_\beta)$ is reasonably diagonal, does it make sense to go on to consider the temporal forecast skill $r(a_\alpha, b_\alpha)$ of the expansion functions in (21). The bold entries in Table 3 show that the forecast and simulation modes agree very well and that the off diagonal terms are small. This is less the case for the agreement between forecast and simulation EOFs and observation-based EOFs. Although the bold diagonal values are largest, there is some “miss-projection” involving modes 1 and 2. Differences in model-based and observation-based EOFs as indicated in Table 3 contribute to the error of the forecast of the PDO and of other modes.

8.2 Temporal skill

The correlations $r(a_\alpha, b_\alpha)$, measuring the temporal agreement between the expansion coefficients of Fig. 3 for the first three modes ($\alpha = 1, 2, 3$ in blue, red, and green respectively) of the observations and ensemble mean forecasts are plotted against forecast range τ in Fig. 5. The corresponding correlations $r(a_\alpha, c_\alpha)$, between the expansion coefficients of the observations and ensemble mean of the simulations, which are not functions of forecast range, are also plotted.

The skill of the PDO if measured by the temporal correlation of the expansion functions $r(a_1, b_1)$ ranges from near 0.7 at year 1 and is not significantly different from zero beyond year 5. The skill of the second mode if measured in this way, drops to zero by year 3 and approaches a somewhat higher value thereafter, although it is not statistically significant beyond the first year of the forecast. Mode 3 is not skillful by this measure. The skill of the simulations does not depend on τ and, as expected, is not statistically significant. Since the EOFs are spatially well correlated in Fig. 4 the mismatch is apparently directly due to the decorrelation of the expansion functions with forecast range. The PDO displays the relatively monotonic decline of skill toward zero with

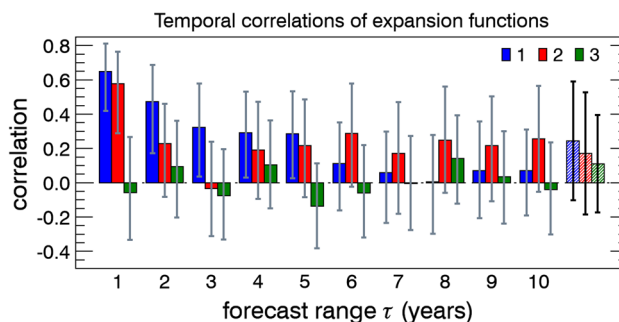


Fig. 5 Temporal correlations $r(a_\alpha, b_\alpha)$ between the expansion functions of the observations and the ensemble mean forecasts are plotted for the first three modes as blue, red and green solid bars respectively. The corresponding correlation $r(a_\alpha, c_\alpha)$ of observation-based and simulated expansion functions are plotted as hatched bars. The 90% confidence intervals are indicated

Table 1 In the decomposition (16)–(18), the temporal standard deviations of the expansion functions of the observations, σ_{a_α} ; of the ensemble mean forecasts for forecast ranges $\tau = \text{year } 1$ and year 5, σ_{b_α} ; and the ensemble mean of the simulations, σ_{c_α} for the first three modes

Standard deviations	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	Total	
$\sigma_{a_\alpha} = \sqrt{a_\alpha^2}$	8.6	5.6	3.8	σ_X	13.3
$\sigma_{b_\alpha} = \sqrt{b_\alpha^2} (\tau = 1)$	9.4	4.6	3.2	σ_{Y_A}	12.5
$\sigma_{b_\alpha} = \sqrt{b_\alpha^2} (\tau = 5)$	3.3	3.4	1.8	σ_{Y_A}	6.6
$\sigma_{c_\alpha} = \sqrt{c_\alpha^2}$	4.7	3.4	1.9	σ_{U_A}	6.9

The values in the last column are for the standard deviation calculated as the square root of the total variance. Units: °C

increasing forecast range that is expected for an internally generated component of variability in isolation.

Temporal correlations between these expansion functions are given in Table 4. The diagonal terms are generally the largest, at least for modes 1 and 2, but off-diagonal terms are not all small. The differences between EOFs will result in differences in expansion functions and this will affect their temporal correlation. If the EOFs all agreed (or if the fields were expanded in a single set of basis functions) the temporal correlation between expansion functions might be more diagonal but would mask the information that the forecast and observation-based modes differ.

8.3 Weights

Finally, the “weight” term $w(a_\alpha, b_\alpha)$ in (25) measures the contribution of the mode to the modal correlation R_α in (24). It depends on the fraction of the observed and predicted variance that the α^{th} components of the observations and ensemble

Table 2 Modal skill statistics

	Expansion	Variances
Observation	$X = \sum_{\alpha} a_{\alpha} e_{\alpha} + X_*$	$\sigma_X^2 = \sum_{\alpha} \sigma_{a_{\alpha}}^2 + \sigma_{X_*}^2$
Ensemble mean forecast ranges $\tau = 1$ and 5 years	$Y_A = \sum_{\alpha} b_{\alpha} f_{\alpha} + Y_*$	$\sigma_{Y_A}^2 = \sum_{\alpha} \sigma_{b_{\alpha}}^2 + \sigma_{Y_*}^2$
Simulation	$U_A = \sum_{\alpha} c_{\alpha} g_{\alpha} + U_*$	$\sigma_{U_A}^2 = \sum_{\alpha} \sigma_{c_{\alpha}}^2 + \sigma_{U_*}^2$
Contribution of the α^{th} mode		
Covariance	$C_{\alpha} = \sigma_{a_{\alpha}} \sigma_{b_{\alpha}} r(a_{\alpha}, b_{\alpha}) s(e_{\alpha}, f_{\alpha})$	
Correlation	$R_{\alpha} = \left(\frac{\sigma_{a_{\alpha}} \sigma_{b_{\alpha}}}{\sigma_X \sigma_{Y_A}} \right) r(a_{\alpha}, b_{\alpha}) s(e_{\alpha}, f_{\alpha}) = w(a_{\alpha}, b_{\alpha}) r(a_{\alpha}, b_{\alpha}) s(e_{\alpha}, f_{\alpha})$	
MSE	$E_{\alpha} = \sigma_{a_{\alpha}}^2 + \sigma_{b_{\alpha}}^2 - 2\sigma_{a_{\alpha}} \sigma_{b_{\alpha}} r(a_{\alpha}, b_{\alpha}) s(e_{\alpha}, f_{\alpha})$	
MSSS	$M_{\alpha} = 2 \left(\frac{\sigma_{Y_A}}{\sigma_X} \right) R_{\alpha} - \frac{\sigma_{b_{\alpha}}^2}{\sigma_X^2}$	

Table 3 Spatial correlations $s(a_{\alpha}, b_{\beta})$, $s(a_{\alpha}, c_{\beta})$ and $s(b_{\alpha}, c_{\beta})$ between the EOFs of the observations and the ensemble of forecasts at ranges $\tau = 1$ and 5 years, and the ensemble of simulations, for the first three modes

	FCST $\tau = 1$			FCST $\tau = 5$			SIM		
	f_1	f_2	f_3	f_1	f_2	f_3	g_1	g_2	g_3
OBS									
e_1	0.80	0.50	-0.15	0.77	0.29	-0.34	0.75	0.41	-0.26
e_2	-0.51	0.75	0.02	-0.35	0.77	-0.05	-0.47	0.75	0.03
e_3	-0.10	0.04	0.12	-0.12	0.00	0.09	-0.06	-0.13	0.17
SIM									
g_1	0.96	0.05	0.05	0.97	-0.17	0.10			
g_2	-0.01	0.94	0.16	0.10	0.97	0.11			
g_3	0.00	-0.20	0.89	-0.09	-0.12	0.87			

Table 4 Temporal correlations $r(a_{\alpha}, b_{\beta})$, $r(a_{\alpha}, c_{\beta})$ and $r(b_{\alpha}, c_{\beta})$ between the expansion functions of the observations and the ensemble mean forecasts and simulations for the first three modes

	FCST $\tau = 1$			FCST $\tau = 5$			SIM		
	b_1	b_2	b_3	b_1	b_2	b_3	c_1	c_2	c_3
OBS									
a_1	0.65	0.49	0.07	0.29	-0.01	-0.09	0.24	0.45	0.08
a_2	-0.35	0.58	0.17	0.11	0.22	-0.04	0.10	0.17	-0.12
a_3	0.02	-0.02	-0.06	0.33	0.17	-0.14	0.14	0.16	0.11
SIM									
c_1	0.50	0.20	-0.31	-0.03	-0.13	0.19			
c_2	0.24	0.25	-0.11	0.43	0.71	-0.64			
c_3	-0.28	-0.62	0.51	-0.35	-0.41	0.37			

mean forecasts account for. The weight terms are less than 1 and the variance fraction $\sigma_{a_{\alpha}}^2 / \sigma_X^2$ is just that accounted for by the observations while $\sigma_{b_{\alpha}}^2 / \sigma_{Y_A}^2$ is the corresponding term but for the ensemble mean forecasts. The weight functions $w(a_{\alpha}, b_{\alpha})$ depending on forecast range, and $w(a_{\alpha}, c_{\alpha})$ for the ensemble mean simulations are plotted in Fig. 6. As expected, the weight function associated with $\alpha = 1$, the PDO, is the

largest with the weight function for $\alpha = 2$ intermediate and with the third mode having very little weight.

8.4 Correlation skill of the PDO

There are different versions of ‘‘correlation skill’’ depending on what is being considered in (24). The temporal

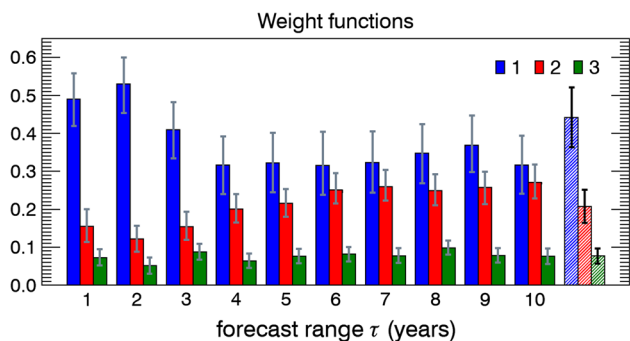


Fig. 6 The weight function from (25) for the first three modes, giving the contribution of modal correlations to the overall correlation skill, are plotted as blue, red and green solid bars respectively. The corresponding weights for the first three modes of the simulations are plotted as hatched bars. The 90% confidence intervals are indicated

correlation between the expansion functions, $\hat{R}_\alpha = r(a_\alpha, b_\alpha)$ will be different from the correlation skill \tilde{R}_α that would be obtained if the α^{th} mode were considered in isolation, which would be

$$\tilde{R}_\alpha = \frac{\overline{a_\alpha b_\alpha \langle e_\alpha f_\alpha \rangle}}{\sigma_{a_\alpha} \sigma_{b_\alpha}} = r(a_\alpha, b_\alpha) s(e_\alpha, f_\alpha)$$

Since $s(e_\alpha, f_\alpha) < 1$, unless the observation- and forecast-based EOFs are the same $\tilde{R}_\alpha < \hat{R}_\alpha$ which, in turn, will be smaller than R_α , the contribution to the overall correlation skill so that

$$\hat{R}_\alpha > \tilde{R}_\alpha > w(a_\alpha, b_\alpha) \tilde{R}_\alpha = R_\alpha.$$

A mode may be well predicted according to $\hat{R}_\alpha = r(a_\alpha, b_\alpha)$ but less so according to \tilde{R}_α and less so again in terms of its contribution R_α to the overall correlation skill R .

Figure 7 plots the terms $\hat{R}_1 > \tilde{R}_1 > R_1$ for

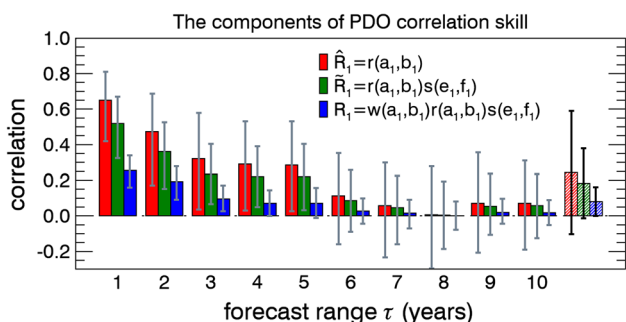


Fig. 7 The components of the PDO correlation R_1 from (24) for the forecasts (solid bars) and simulations (hatched bars). The temporal correlation of the expansion coefficients (red) is discounted by the imperfect correlation between the observation- and forecast-based EOFs (green) and subsequently weighted by the variances involved to produce the contribution R_1 (blue) of the PDO to the overall area correlation R . The 90% confidence intervals are indicated

$$\begin{aligned} \hat{R}_1 &= r(a_1, b_1) \\ \tilde{R}_1 &= r(a_1, b_1) s(e_1, f_1) \\ R_1 &= w(a_1, b_1) r(a_1, b_1) s(e_1, f_1) \end{aligned}$$

illustrating the contributions to the PDO correlation skill R_1 associated with the spatial and temporal correlations and the weight function. The temporal skill of the PDO expansion functions $r(a_1, b_1)$ is comparatively high. It is discounted modestly by $s(e_1, f_1)$, representing the mismatch between the forecast and observation-based EOFs patterns associated with the PDO and further by the the weight function $w(a_1, b_1)$ involving the variances. The PDO accounts for a portion of the overall variance and only part of its variance is skillfully predicted.

8.5 Modal contributions to North Pacific correlation skill

Figure 8 plots the cumulative contributions $R = R_1 + R_2 + R_3 + R_*$ to the overall correlation from the first 3 modes and the residual. The overall correlation skill R in (23) for annual mean SST for the North Pacific PDO area (given by the overall height of the bar) is about 0.6 for the first year and declines comparatively rapidly for years 2–4 and then is more or less stable or slightly declining at later times. The 90% bootstrap confidence intervals shown in the figure indicate that the overall correlation skill is statistically significant for the first 4 years of the forecasts and for the simulations, and is marginally or not significant for later years. The forecast skill at range 1 is significantly larger than that of the simulations, as implied by the corresponding confidence intervals in the figure, but it is not so for longer ranges according to the 90% confidence level

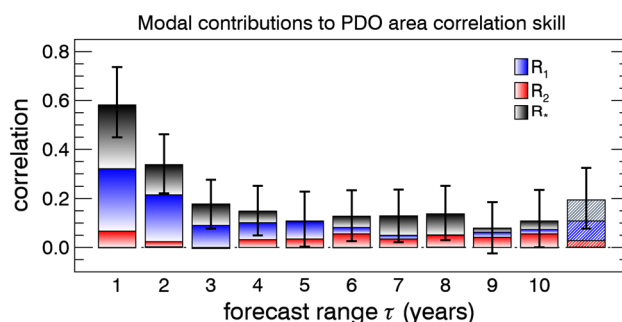


Fig. 8 The cumulative contributions $R = R_1 + R_2 + (R_3 = 0) + R_*$ to the overall SST correlation from the first three modes plus the residual. The contribution from the third mode is effectively zero. For each forecast range (solid bars) and for the simulations (hatched bars), the contribution of modes R_1 and R_2 are indicated respectively by the blue and red sectors and that of R_* by the shades of black. The 90% confidence intervals of the overall correlation skill are indicated

from the bootstrap procedure applied to the differences in the correlations (not shown). The overall correlation for the simulations presumably reflects the residual effects of the forced signal that has not been completely removed i.e. by subtracting out the global mean SST at each point. The forecast correlation value is slightly smaller at longer forecast ranges than that of the simulations. This suggests that either the “forced signal” removal is more successful in the forecasts than in the simulations or, perhaps due to errors associated with initialization, the forecast skill of the PDO “overdecays” and would recover to the simulation value at longer τ .

The PDO correlation R_1 is the main contributor to the overall correlation skill R at least for the first 5 years while mode 2 correlation R_2 contributes very little and mode three correlation R_3 essentially contributes not at all. The residual correlation R_* contributes to the overall skill during the first year of the forecast and modestly afterwards.

9 Summary and conclusions

The PDO is, as identified here, the leading empirical orthogonal function (EOF) of annual mean SST anomalies in the Pacific Ocean between 20°N and 60°N after removing the global mean SST, together with its associated principal component (PC) time series. It is the dominant contribution to annual mean SST variance, accounting for some 40% of the total based on the ERSSTv5 data set. The CCCma climate model is used both for simulations and for decadal predictions and the leading EOFs of the simulations and the predictions are identified as the respective simulated and forecast EOFs of the PDO. There are corresponding higher order modes which account for less of the variability. The EOF analysis applies to the 1971–2015 period for all of the observations, forecasts and simulations.

The observation-based EOFs differ from those of the forecasts and simulations which also differ from one another. These represent differences in the structures of the respective representations of the PDO (and higher order modes) in the observations, forecasts, and simulations. The forecast EOFs are functions of forecast range τ while the EOFs of the simulations and observations are not. The CCCma model reproduces the spatial pattern of the observation-based PDO, i.e. the leading EOF, quite well (spatial correlation skill near 0.8) for both simulations and forecasts with rather modest differences with forecast range. This is the case also for the second, but not the third mode.

The temporal behaviour of the observation based, forecast, and simulated PDOs is carried by the respective expansion functions. For the forecasts the correlation skill compared to the observations behaves more or less as expected with values near 0.7 for year 1 forecasts and with values

declining toward zero subsequently. The value of the temporal correlation skill of the simulated PDO with the observations is comparatively small but non-zero, at least as calculated, suggesting some residual of the forced component of variability may be involved.

The correlation skill of the PDO, considered in isolation, is the product of the spatial correlation of the EOFs and the temporal correlation of the expansion functions. While the PDO is the dominant mode it does not account for all of the annual mean SST variability and its contribution to the overall correlation for the region is determined by a weight function involving the fraction of the variability accounted for by the observed and forecast modes.

The approach adopted here discriminates between observation-, forecast-, and simulation-based versions of the PDO's spatial structure in order to compare the respective PDOs as such rather than as their projections on a particular set of basis function. Results obtained when expanding the forecasts and simulations in terms of the observation-based leading EOFs, for instance, do not consider this mismatch which could be considerable. In the case of the CCCma model's simulations and forecasts, the model generated PDO patterns are comparatively close to the observation based version, as measured by spatial correlation.

The overall correlations skill for the North Pacific is dominated by the PDO with a small contribution from the second mode and none from the third mode. The non-PDO components associated with the rest of the variability also contribute modestly.

This general approach can be applied to other skill measures and, potentially, to estimates of the predictability of the PDO and other modes. The novelty and virtues of the method are that the leading EOF, i.e. the PDO, of the observations is compared to the leading EOF of the forecasts and simulations, that the temporal skill of the PDO is calculated separately, and that the several contributions to the overall skill for the North Pacific PDO region is obtained. The approach is also suitable for comparing PDO skill across models in terms of EOF structures, expansion coefficients, skill components, and overall area skill.

Acknowledgements We are grateful to our colleagues Bill Merryfield and Ali Asaadi for their helpful comments on an earlier version of the paper. We are also pleased to acknowledge the important contributions of many members of the CCCma team in the development of the model and the forecasting system and to Woo-Sung Lee for her contribution in producing the forecasts.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Boer GJ (2004) Long timescale potential predictability in an ensemble of coupled climate models. *Clim Dyn* 23:29–44
- Boer GJ, Kharin VV, Merryfield WJ (2013) Decadal predictability and forecast skill. *Clim Dyn* 41:1817–1833
- Boer GJ, Smith DM, Cassou C, Doblas-Reyes F, Danabasoglu G, Kirtman B, Kushnir Y, Kimoto M, Meehl GA, Msadek R, Mueller WA, Taylor KE, Zwiers F, Rixen M, Ruprich-Robert Y, Eade R (2016) The decadal climate prediction project (DCPP) contribution to CMIP6. *Geosci Mod Dev* 9:3751–3777
- Boer GJ, Kharin VV, Merryfield WJ (2019a) Differences in potential and actual skill in a decadal prediction experiment. *Clim Dyn* 52(11):6619–6631
- Boer GJ, Merryfield WJ, Kharin VV (2019b) Relationships between potential, attainable, and actual skill in a decadal prediction experiment. *Clim Dyn* 52(7–8):4813–4831
- Bond NA, Overland JE, Spillane M, Stabeno P (2003) Recent shifts in the state of the North Pacific. *Geophys Res Lett* 30(23):2183
- Boo KO, Booth BBB, Byun YH, Lee J, Cho C, Shim S, Kim KT (2015) Influence of aerosols in multidecadal SST variability simulations over the North Pacific. *J Geophys Res Atmos* 120:517–531
- Cassou C, Kushnir Y, Hawkins E, Pirani A, Kucharski F, Kang IS, Caltabiano N (2018) Decadal climate variability and predictability: challenges and opportunities. *Bull Am Meteor Soc* 99:479–486
- Doblas-Reyes FJ, Andreu-Burillo I, Chikamoto Y, García-Serrano J, Guemas V, Kimoto M, Mochizuki T, Rodrigues LRL, van Oldenborgh GJ (2013) Initialized near-term regional climate change prediction. *Nat Commun* 4:1715
- Frankignoul C, Gastineau G (2017) Estimation of the SST response to anthropogenic and external forcing and its impact on the Atlantic Multidecadal Oscillation and the Pacific Decadal Oscillation. *J Clim* 30:9871–9895
- Gershunov A, Barnett TP (1998) Interdecadal modulation of ENSO teleconnections. *Bull Am Meteor Soc* 79:2715–2725
- Guemas V, Doblas-Reyes FJ, Lienert F, Soufflet Y, Du H (2012) Identifying the causes of the poor decadal climate prediction skill over the North Pacific. *J Geophys Res* 117:D20111
- Huang B, Thorne PW, Banzon VF, Boyer T, Chepurin G, Lawrimore JH, Menne MJ, Smith TM, Vose RM, Zhang HM (2017) Extended reconstructed sea surface temperature version 5 (ERSSTv5), upgrades, validations, and intercomparisons. *J Clim* 30:8179–8205
- Kharin VV, Boer GJ, Merryfield WJ, Scinocca JF, Lee W-S (2012) Statistical adjustment of decadal predictions in a changing climate. *Geophys Res Lett* 39:L19705
- Kim HM, Ham YG, Scaife AA (2014) Improvement of initialized decadal predictions over the North Pacific ocean by systematic anomaly pattern correction. *J Clim* 27:5148–5162
- Kirtman B, Power SB, Adedoyin JA, Boer GJ, Bojariu R, Camilloni I, Doblas-Reyes FJ, Fiore AM, Kimoto M, Meehl GA, Prather M, Sarr A, Schär C, Sutton R, van Oldenborgh G, Vecchi G, Wang HJ (2013) Near-term climate change: projections and predictability. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: the physical science basis. contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Kushnir Y, Scaife AA, Arritt R, Balsamo G, Boer G, Doblas-Reyes F, Hawkins E, Kimoto M, Kolli R, Kumar A, Matei D, Matthes K, Muller WA, O’Kane T, Perlwitz J, Power S, Raphael M, Shimp A, Smith D, Tuma M, Wu B (2019) Towards operational predictions of the near-term climate. *Nat Clim Change* 9:94–101
- Luo JJ, Yamagata T (2002) Four decadal ocean-atmosphere modes in the North Pacific revealed by various analysis methods. *J Ocean* 58:861–876
- Malik A, Brönnimann S, Perona P (2018) Statistical link between external climate forcings and modes of ocean variability. *Clim Dyn* 50:3649–3670
- Mantua NJ, Hare SR (2002) The Pacific decadal oscillation. *J Ocean* 58:35–44
- Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC (1997) A Pacific interdecadal climate oscillation with impacts on Salmon production. *Bull Am Meteor Soc* 78:1069–1079
- Meehl GA, Hu A, Teng H (2016) Initialized decadal prediction for transition to positive phase of the interdecadal Pacific oscillation. *Nat Commun* 7(11718):1–7
- Mehta VM, Mendoza K, Wang H (2019) Predictability of phases and magnitudes of natural decadal climate variability phenomena in CMIP5 experiments with the UKMO HadCM3, GFDL-CM2.1, NCAR-CCSM4, and MIROC5 global earth system models. *Clim Dyn* 52:3255–3275
- Merryfield WJ, Lee WS, Boer G (2013) The Canadian seasonal to interannual prediction system. Part I: models and initialization. *Mon Weather Rev* 141:2910–2945
- Mochizuki T, Ishii M, Kimoto M, Chikamoto Y, Watanabe M, Nozawa T, Sakamoto TT, Shiogama H, Awaji T, Sugiura N, Toyoda T, Yasunaka S, Tatebe H, Mori M (2010) Pacific decadal oscillation hindcasts relevant to near-term climate prediction. *PNAS* 107:1833–1837
- Murphy J, Kattsov V, Keenlyside N, Kimoto M, Meehl G, Mehta V, Pohlmann H, Scaife A, Smith D (2010) Towards prediction of decadal climate variability and change. *Proc Environ Sci* 1:287–304
- Newman M, MA A, TR A (2016) The Pacific decadal oscillation, revisited. *J Clim* 29:4399–4427
- Power S, Casey T, Folland C, Colman A, Mehta V (1999) Interdecadal modulation of the impact of ENSO on Australia. *Clim Dyn* 15:319–324
- Smith DM, Cusack S, Colman AW, Folland CK, Harris G, Murphy JM (2007) Improved surface temperature prediction for the coming decade from a global climate model. *Science* 317:796–799
- Smith DM, Scaife AA, Boer GJ, Caian M, Doblas-Reyes FJ, Guemas V, Hawkins E, Hazeleger W, Hermanson L, Ho CK, Ishii M, Kharin V, Kimoto M, Kirtman B, Lean J, Matei D, Merryfield WJ, Muller WA, Pohlmann H, Rosati A, Wouters B, Wyser K (2013) Real-time multi-model decadal climate predictions. *Clim Dyn* 41:2875–2888
- Smith DM, Booth BBB, Dunstone NJ, Eade R, Hermanson L, Jones GS, Scaife AA, Sheen KL, Thompson V (2016) Role of volcanic and anthropogenic aerosol in the recent global surface warming slowdown. *Nat Clim Change* 6:936–941
- Wilks DS (1997) Resampling hypothesis tests for autocorrelated fields. *J Clim* 10:65–82

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