

*Decadal predictability and prediction skill
of sea surface temperatures in the South
Pacific region*

**Ramiro I. Saurral, Javier García-Serrano,
Francisco J. Doblas-Reyes, Leandro
B. Díaz & Carolina S. Vera**

Climate Dynamics

Observational, Theoretical and
Computational Research on the Climate
System

ISSN 0930-7575

Clim Dyn

DOI 10.1007/s00382-020-05208-3



Your article is protected by copyright and all rights are held exclusively by Springer-Verlag GmbH Germany, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



Decadal predictability and prediction skill of sea surface temperatures in the South Pacific region

Ramiro I. Saurral^{1,2,3} · Javier García-Serrano⁴ · Francisco J. Doblas-Reyes^{5,6} · Leandro B. Díaz^{1,2,3} · Carolina S. Vera^{1,2,3}

Received: 23 June 2019 / Accepted: 12 March 2020

© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

The South Pacific Ocean is a key driver of climate variability within the Southern Hemisphere at different time scales. Previous studies have characterized the main mode of interannual sea surface temperature (SST) variability in that region as a dipolar pattern of SST anomalies that cover subtropical and extratropical latitudes (the South Pacific Ocean Dipole, or SPOD), which is related to precipitation and temperature anomalies over several regions throughout the Southern Hemisphere. Using that relationship and the reported low predictive skill of precipitation anomalies over the Southern Hemisphere, this work explores the predictability and prediction skill of the SPOD in near-term climate hindcasts using a set of state-of-the-art forecast systems. Results show that predictability greatly benefits from initializing the hindcasts beyond the prescribed radiative forcing, and is modulated by known modes of climate variability, namely El Niño-Southern Oscillation and the Interdecadal Pacific Oscillation. Furthermore, the models are capable of simulating the spatial pattern of the observed SPOD even without initialization, which suggests that the key dynamical processes are properly represented. However, the hindcast of the actual phase of the mode is only achieved when the forecast systems are initialized, pointing at SPOD variability to not be radiatively forced but probably internally generated. The comparison with the performance of an empirical prediction based on persistence suggests that initialization may provide skillful information for SST anomalies, outperforming damping processes, up to 2–3 years into the future.

Keywords Near-term climate predictions · IPO · ENSO · SST · South Pacific

1 Introduction

The climate of the Southern Hemisphere is modulated by variability in the South Pacific (SP) region at several time scales, including seasonal (Guan et al. 2014), interannual (Linsley et al. 2000; Barros and Silvestri 2002; Shakun and Shaman 2009) and multi-decadal (Power et al. 1999). Yet, limited literature on the characterization of SP climate variability and dynamics exists so far, which in part has been attributed to the sparse data available over the basin (Basher and Zheng 1998). Nonetheless, recent active research on the field has provided further insight into our knowledge of the region (e.g. Power et al. 2017; Saurral et al. 2018; Lou et al. 2019), particularly focused on the main modes of sea surface temperature (SST) variability, their drivers and impacts.

In a recent study, Saurral et al. (2018) characterized the main modes of SST variability on interannual time scales in the basin. They found that the leading mode exhibits a dipolar structure of SST anomalies covering subtropical and extratropical latitudes (see Fig. 1) that is highly and

✉ Ramiro I. Saurral
saurral@cima.fcen.uba.ar

¹ Universidad de Buenos Aires, Consejo Nacional de Investigaciones Científicas y Técnicas, Centro de Investigaciones del Mar y la Atmósfera (CIMA/UBA-CONICET), Buenos Aires, Argentina

² Departamento de Ciencias de la Atmósfera y los Océanos (DCAO/FCEN/UBA), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

³ Institut Franco-Argentin d'Estudes sur le Climat et ses Impacts, Unité Mixte Internationale (UMI-IFAECI/CNRS-CONICET-UBA), Buenos Aires, Argentina

⁴ Group of Meteorology, Universitat de Barcelona (UB), Barcelona, Spain

⁵ Barcelona Supercomputing Center (BSC), Barcelona, Spain

⁶ Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

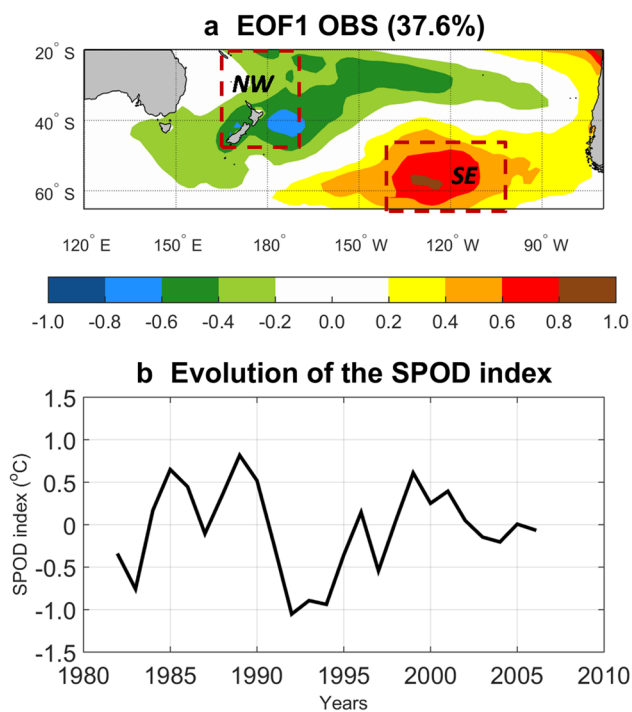


Fig. 1 **a** Spatial pattern of the leading mode of observed SST anomalies in the SP region, and **b** time evolution (period 1982–2006) of the mode. This mode accounts for 37.6% of the variance. The northwestern and southeastern centers of the SPOD are highlighted as “NW” and “SE”, respectively, in **a**

significantly correlated with El Niño-Southern Oscillation (ENSO) and the Interdecadal Pacific Oscillation (IPO; Power et al. 1999). The second mode is associated with SST anomalies propagating eastward in the extratropics linked to the South Pacific Gyre (their Figs. 2, 8). Both patterns depict significant oscillations at interannual time scales: the first mode has periodicities of about 4 and 8 years and the second mode oscillates at approximately 7 years. More recently, Lou et al. (2019) used a global ocean model to identify the sources of SST variability over the SP region at decadal and longer time scales and found that the topography of the SP basin helps in low-pass filtering the SST and, therefore, increasing the fraction of variance explained by lower-frequency modes compared to the North Pacific region. At the same time, Chung et al. (2019) showed that part of the Tropical Pacific variability is explained and driven by variations in salinity and SST over the SP region, providing further evidence of the importance of the SP region in sustaining and modulating the global climate.

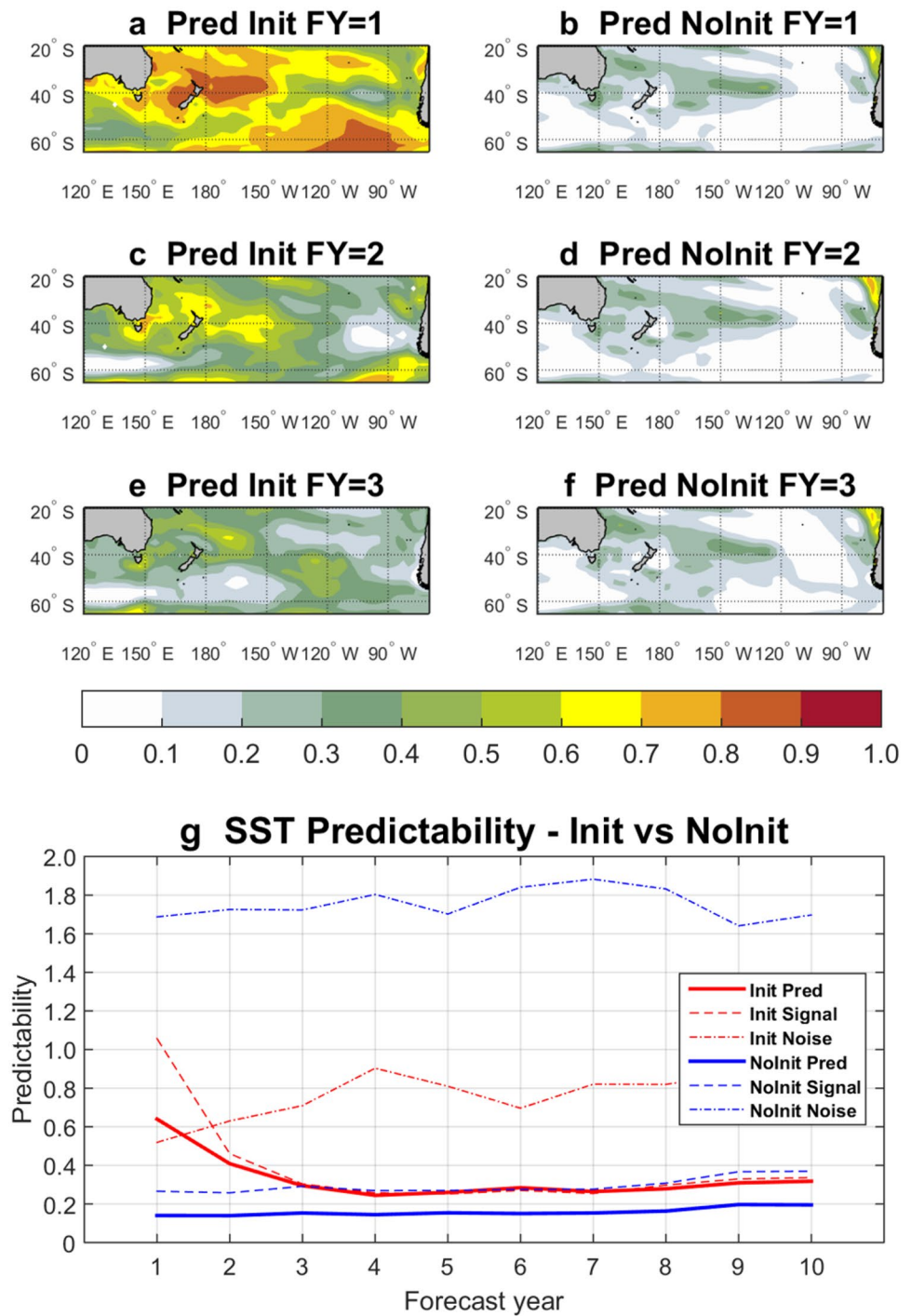
Previous studies have identified potential sources of climate predictability in the SP region. Among them, Guan et al. (2014) concluded that seasonal SST variability in the area can be predicted with acceptable skill up to nine months

in advance. At longer time scales, Ding et al. (2013) showed that the climate shifts observed in the Pacific around 1976 and 1998, as well as the anomalous atmospheric conditions following such changes, could have been successfully predicted 4 years in advance using initialized coupled simulations in retrospective-forecast mode. Guémas et al. (2013) found that some skill may be present over selected areas of the SP basin up to 2–5 years into the future. However, no quantification of the predictability of SST variability in the SP region has yet been carried out.

The presence of significant seasonal-to-interdecadal variability in the SP region leads to the question of whether these variations may be skillfully predicted. In this regard, the set of decadal climate predictions produced under the Fifth Phase of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) appears as a suitable tool to effectively assess multi-year forecast skill of SST variability in the SP region. The CMIP5 data comprise simulations from different coupled global climate models (GCMs) especially developed for analyzing near-term climate predictability on interannual-to-decadal timescales. Two sets of simulations are used: initialized and non-initialized hindcasts (Doblas-Reyes et al. 2013; Kirtman et al. 2013). The initialized (*Init*) or *decadal* hindcasts are retrospective forecasts that employ the best estimate of the observed climate state at the start date and are run for 10 years into the future. Although standard hindcasts with a 5-year interval between start dates may be enough to evaluate the level of skill, yearly hindcasts allow assessing the forecast quality more accurately and with higher statistical robustness (e.g. García-Serrano and Doblas-Reyes 2012; García-Serrano et al. 2015). Therefore, the latter set includes simulations initialized every year from 1960 to 2005. The set of non-initialized hindcasts (*NoInit*) consists of historical runs started in the nineteenth or twentieth century from a long control simulation where no information of the actual state of the climate system is prescribed, and only radiative forcing is applied. In this study, the historical simulations are continued beyond 2005 with the projections using the scenario RCP4.5. The assessment of the actual impact of initialization on the forecast quality is carried out by contrasting predictions skill from the *Init* and *NoInit* hindcasts.

Recent studies have shown that the actual skill of decadal predictions for some atmospheric variables such as precipitation over South America is limited, even with the added value of initialization (e.g. Smith et al. 2019). As such, this study follows an approach usually employed in seasonal prediction studies, in which prediction skill of a given variable (e.g. precipitation) is quantified through the analysis of the prediction skill of known modes of climate variability that are related to this variable (e.g. Osman and Vera 2020). In the present study, the focus is put on the

Fig. 2 a–f Predictability of SST in the SP region in (left column) Init and (right column) Nolnit hindcasts, from the first to the third forecast years (first to third row). **g** Basin-mean potential predictability evolution along the forecast years considering Init and Nolnit (blue and red thick curves, respectively). Also shown are terms related to the variance of the signal (dashed lines) and variance of the noise (dashed-dotted lines)



documented relationship between SST variability in the SP region and rainfall and temperature anomalies over the Southern Hemisphere.

Multi-year prediction skill in the Pacific Ocean has been the subject of several recent studies. For instance, Lienert and Doblas-Reyes (2013) analyzed the decadal prediction problem in the North Pacific region and found that the

second mode of variability, called the North Pacific Gyre Oscillation, can be predicted with some skill almost 10 years in advance. Guémas et al. (2012) found very little skill in decadal hindcasts over the North Pacific region, in agreement with other multi-model assessments (e.g. Kim et al. 2012; Doblas-Reyes et al. 2013), and linked it to the poor representation of anomalous SST events occurring

in the second half of the twentieth century. More recently, González and Goddard (2016) analyzed how a set of CMIP5 decadal hindcasts predicts ENSO variability in seasonal to interannual scale. They concluded that ENSO may be predicted up to 2- to 3-years in advance, in agreement with previous findings using other forecast systems (Jin et al. 2008; Volpi et al. 2013).

The main objective of this study is to quantify predictability as well as forecast skill in a set of hindcasts to simulate and predict interannual-to-decadal variability of SST anomalies in the SP region. A key aspect of the study is to systematically compare the performance of Init and NoInit as a way to assess the impact of initialization on the forecast skill, as well as to explore whether the SST variability is internally generated or radiatively forced.

Given the reported relationships of SST anomalies in this basin with temperature and precipitation variability over several land areas of the Southern Hemisphere (Saurral et al. 2018), this study could provide useful information for implementing regional climate services oriented to different socio-economic sectors like water management and agricultural planning. The paper is organized as follows: Sect. 2 describes the data and methodology used in the study. Section 3 contains the results and related discussion, including the assessment of predictability, a comparison between objective metrics of prediction skill in the Init and NoInit hindcasts and also an analysis of its link to temperature and precipitation predictions in the Southern Hemisphere. Finally, the conclusions are drawn in Sect. 4.

2 Data and methods

2.1 Observations

Observed monthly mean SST data was taken from version 1.1 of the HadISST dataset (Rayner et al. 2003) which has a horizontal resolution of $1^\circ \times 1^\circ$ and includes in-situ measurements from the Met Office Marine Data Bank as well as satellite-derived SST information starting in 1982. As discussed in Saurral et al. (2018), the lack of measurements over the SP basin before 1982 could affect the results and lead to biased conclusions. In fact, a recent paper by Volpi et al. (2017) using different initialization approaches applied to the EC-Earth climate model found that the highest skill scores (i.e. best model performance) over the SP region are achieved in NoInit experiments, a result that the authors attribute to the sparseness of in-situ observations acting to preclude robust initialization there. For this reason, and despite the availability of hindcasts from

the 1960s onward, it was decided to narrow the period of analysis from January 1982 to December 2015 (34 years) in order to consider only years with sufficient, reliable observational data (see further discussion in Saurral et al. 2018). This study employs SST annual-mean values, so the original monthly values were averaged to derive 34 values (one per year) for every grid point. Most of the analysis is performed over the SP region, considered as the area bounded by Australia to the west and South America to the east, and from 20° S to 65° S.

As mentioned in the Introduction, the leading mode of interannual SST variability in the basin given by a Principal Component Analysis (PCA/EOF) shows a dipolar structure of anomalies, which is centered around 40° S/ 170° W and 55° S/ 130° W respectively (Saurral et al. 2018; see Fig. 1a). The ability of the decadal prediction systems to simulate and forecast this pattern, called the South Pacific Ocean Dipole (SPOD) Huang and Shukla (2006), is addressed. In this paper, the SPOD index is computed as the difference of area-averaged SST anomalies in the box delimited by 20° S– 48° S/ 165° E– 170° W (the NW pole) minus those over 44° S– 65° S/ 140° W– 100° W (the SE pole, Fig. 1b; see locations of the two poles in Fig. 1a). The analysis of the corresponding principal components (PCs) associated with the EOFs of SST provides identical results, but the approach of area-averaging is adopted for the sake of reproducibility and future applicability.

2.2 Hindcasts

Several forecast systems are used for the analysis of near-term climate prediction. These include nine physically-perturbed variants of the Met Office Decadal Climate Prediction System (Smith et al. 2010) and four models from the CMIP5 set of yearly hindcasts, namely HadCM3 (Gordon et al. 2000), MIROC5 (Hasumi and Emori 2004), GFDL-CM2.1 (Yang et al. 2013; Zhang et al. 2017; from now on, simply GFDL) and EC-Earth2.3 (Du et al. 2012; hereafter EC-Earth). As introduced in Sect. 1, each forecast system contributes with Init and NoInit decadal hindcasts. Initialization in DePreSys and EC-Earth is on November 1st, while in the other models is on January 1st of the following year (i.e. two months later). Anomaly initialization (a.i.) was employed in DePreSys, HadCM3 and MIROC5; while EC-Earth and GFDL make use of full-field initialization (FFI).

All simulations include the effects of solar variability, greenhouse gases and anthropogenic aerosols, although CMIP5 models also consider the role of volcanic eruptions and the related injection of aerosols into the stratosphere while DePreSys projects the volcanic aerosol load available at the start date. As the DePreSys dataset consists of nine

different (perturbed physics) versions of a single model, it is considered here as a multi-model with one member per model version. The CMIP5 GCMs have varying ensemble sizes: 10 ensemble members in GFDL and HadCM3, 5 (3) in EC-Earth Init (NoInit) and 6 (3) in MIROC5 Init (NoInit). Most of the analysis is done upon the multi-model ensemble-mean, without discussing particular merits of the individual forecast systems.

Variables analyzed include monthly SST, 2-m air temperature and total precipitation from both Init and NoInit hindcasts over the Southern Hemisphere. As for the observations, all monthly bias-corrected anomalies were averaged to obtain annual means, with one value per year at each grid point.

2.3 Statistical tools

Predictability is quantified following the definition proposed by Schubert et al. (2002), in which predictability of a quantity x is given by the ratio between the variance of the signal and the total variance (which results from the sum of contributions from the signal and the noise). Variance of the signal (S^2) and noise (N^2) are defined as:

$$S^2 = \frac{n}{n-1} \overline{(\langle x \rangle - \langle x \rangle)^2} - \frac{1}{m} N^2 \quad (1)$$

$$N^2 = \frac{m}{m-1} \overline{(\langle x \rangle - \langle x \rangle)^2} \quad (2)$$

Brackets (overbars) indicate averaging over the m ensemble members (the n years). Predictability is therefore defined as the ratio between signal and total variance as follows:

$$P = \frac{S^2}{S^2 + N^2} \quad (3)$$

The quantification of the actual prediction skill is based on the comparison between SST variability in the observations and the different forecast systems. As a first step, for each forecast year the observed and simulated SST climatology were computed by averaging SST across the start dates, and then subtracting the obtained climatology from the raw values to remove the drift in each model (García-Serrano and Doblas-Reyes 2012). To avoid masking the skill with the signal resulting from long-term trends in SST, the trend at each grid point was removed in the observations (hindcasts) through a simple linear regression against the observed (simulated) global-mean temperature across all start dates as a function of the forecast year.

The skill of the hindcasts is determined by computing the anomaly correlation coefficient (ACC) as a function of the forecast year for each forecast system. The same

computation was done in parallel using persistence of the observed anomalies as an empirical prediction system relying on damped processes (e.g. García-Serrano et al. 2012). The leading mode of SST variability was also computed for each forecast system considering Init and NoInit separately so that to assess if the SPOD pattern is simulated by the GCMs and whether initialization is required to properly capture the dipole-like structure or not.

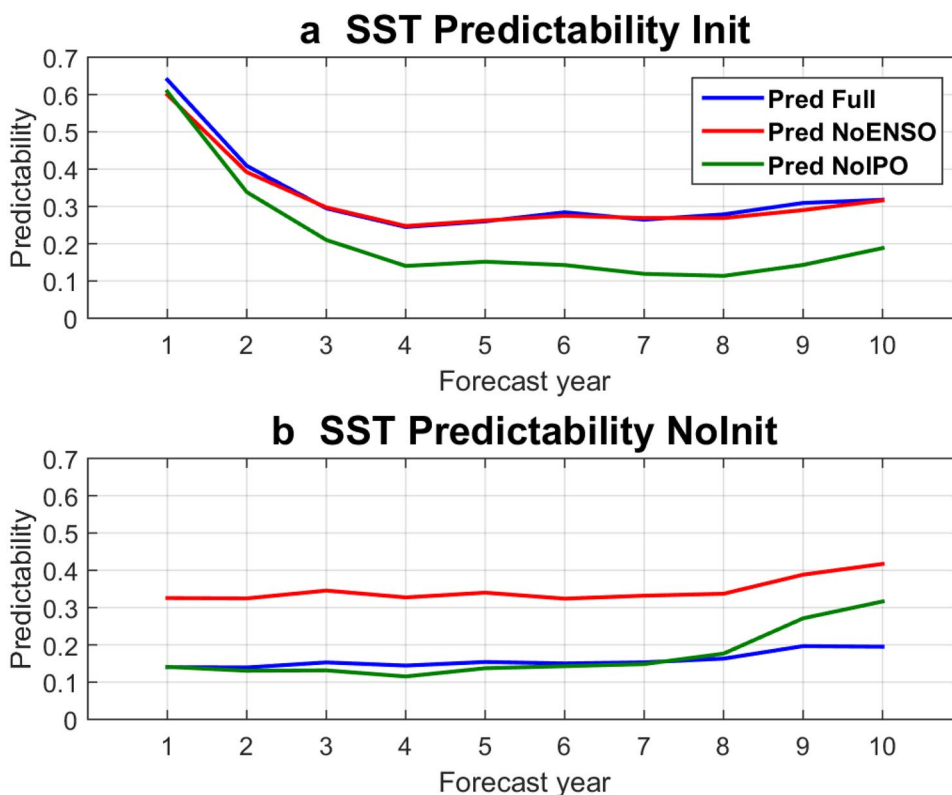
Correlation maps were computed using the observed SPOD index against hindcast precipitation and temperature anomalies in order to explore climate impacts over the Southern Hemisphere that are associated with the SPOD skill. The diagnostic aims to assess the benefits, if any, of the initialization in hindcasting the target teleconnections (García-Serrano et al. 2015).

3 Results

3.1 Predictability of SST anomalies in the SP region

Figure 2 shows predictability of SST anomalies as a function of the forecast year, considering Init and NoInit hindcasts (Fig. 2a–f). Initialization clearly leads to an increase in predictability over the basin, most noticeably in the first 2 forecast years (Fig. 2a, c) with values surpassing 0.6 in two distinct regions: east of New Zealand, and southwest of the southern coast of South America near 60° S. Although the magnitude of predictability clearly diminishes at the third forecast year, it remains above the values derived from NoInit (Fig. 2e, f), suggesting an added value from initialization. The basin-averaged predictability along the forecast time is displayed in Fig. 2g, as well as the terms associated with the signal and noise variances, for Init (red) and NoInit (blue) separately. As expected from the grid-point analysis, there are clear differences during the first and second forecast years. These differences can be mostly attributed to changes in the variance of the signal, which is markedly larger in Init as compared to NoInit in forecast years 1 and 2. At the same time, the NoInit hindcasts contain much more noise than their Init counterparts, which naturally acts to shape the behavior of predictability. It is interesting to note that from the third forecast year onwards, the variance of the signal in Init and NoInit becomes virtually identical, suggesting that initialization does not provide any improvement in its predictability after the second forecast year. Still, the variance of the noise remains substantially larger in NoInit than in Init (dashed-dotted lines) which leads to larger predictability values in Init even until the tenth forecast year and shows how initialization can effectively act to narrow

Fig. 3 Lead time evolution of SP basin-averaged predictability of SST in (top row) Init and (bottom row) NoInit hindcasts, considering the full SST variability (blue curves) and SST without the effects of ENSO (red curves) and IPO (green curves)



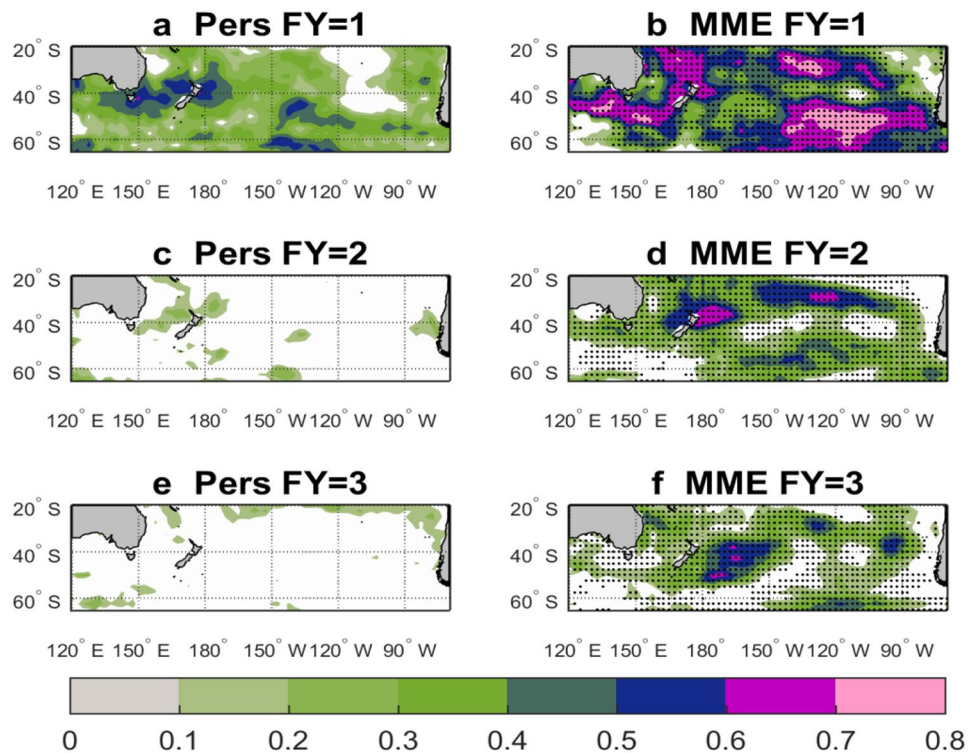
the dispersion among ensemble members along the entire hindcast period.

The relationship between SST variability in the SP region with the activity of large-scale modes of variability of the climate system such as ENSO and the IPO brings the question of whether these modes may be contributing to the SP predictability. To address this point, Fig. 3 shows the evolution of predictability in Init and NoInit considering the actual evolution of SST (“full”; same as in Fig. 2g) as well as after removing the effects of ENSO and IPO, separately. This is achieved by means of regressing out the least-square fit of SST anomalies at each grid point to the Niño 3.4 index and the IPO time series (estimated as the leading principal component of detrended SST over 50° S–50° N/100° E–70° W; Doblas-Reyes et al. 2013), in Init and NoInit respectively.

In the case of Init (Fig. 3a), removing ENSO or IPO leads to a small decrease in predictability during the first and second forecast years. At longer lead times, the contribution of the lower-frequency variability linked to the IPO is substantial up to the end of the hindcast period which results in a sharp decrease in predictability when the effect of IPO is removed from the total variability,

while excluding the interannual variability linked to ENSO at these long lead times results in no change of predictability due to the limited predictability horizon of ENSO itself (e.g. Hou et al. 2018). Regarding NoInit (Fig. 3b), it is worth noting how the variability associated with ENSO represents an important source of unpredictable noise acting in detriment of predictability in the system at all lead times, as the largest values are obtained when considering SST variability with no ENSO effect. Variability associated with the IPO has almost no effect on predictability in NoInit at shorter lead times, but then becomes beneficial at lead times larger than 8 years. It is clear from this analysis that predictability in NoInit hindcasts is heavily affected by the incorrect phasing of both ENSO and IPO at the start of their simulations, which is carried all along the forecast period and is more clearly driven by ENSO. At the same time, it is interesting to note that previous authors have found that this predictability behavior is independent of the nature of the forecast systems considered (i.e., with or without ocean dynamics included) and the same results would be obtained even using simple, slab-ocean coupled models (Srivastava and DelSole 2018).

Fig. 4 ACC computed between (left column) persisted SST anomalies and (right column) multi-model ensemble mean Init hindcast anomalies in (top to bottom) forecast years 1, 2 and 3. Black dots indicate where predictions from the hindcasts have larger ACC values than those from persistence at the corresponding forecast time



3.2 SPOD prediction skill

The objective of this section is to assess the actual skill of the different forecast systems to predict the interannual variability of SST in the SP basin. Figure 4 shows the ACC, with respect to observations, of the empirical predictions based on persistence (first column) and the Init hindcasts (second column) for forecast years 1, 2 and 3 (rows). The climate system exhibits large memory at annual time scale, as can be seen from the persistence-based prediction for forecast year 1 (Fig. 4a). The exception is the central-eastern part of the SP basin, where persistence skill is very low. Simple damping processes are not skillful one year later, yielding a marked drop in correlation for forecast years 2 and 3, where only a few spots across the SP basin retain ACC values exceeding 0.1 (Fig. 4c, e). In the case of the initialized forecast systems, the multi-model ensemble-mean is very skillful for forecast year 1 in most of the SP basin (Fig. 4b). But, more surprisingly, they also perform well and outperform persistence for forecast year 2 (Fig. 4d), particularly over the central part of the basin. Among them, DePreSys, EC-Earth and MIROC5 are those depicting the largest ACC values, with MIROC5 showing the best performance (not shown). At forecast year 3 (Fig. 4f), there are still areas with ACC values above 0.6 in the multi-model

ensemble-mean, mostly over the central SP region. These results illustrate the added value of dynamical climate forecasting as compared to simple empirical prediction for regional multi-annual SST variability. As the skill maps in Fig. 4 (right column) project on the leading mode of variability, i.e. the SPOD (Fig. 1a), in the following the separate role of initialization and radiative forcing on the SPOD skill is assessed.

Figure 5 shows the observed (black line) and simulated evolution of the SPOD index in Init (first column) and NoInit (second column) for the first three forecast years (rows), showing the multi-model ensemble-mean (colored lines) as well as the spread of the individual ensemble-means (colored shading). The ACC score is also computed in each case. Initialization clearly improves the prediction of the SPOD variability in terms of both, amplitude and phase. The multi-model ensemble-mean ACC is much larger for Init (0.51) than for NoInit (0.14) in the first forecast year. The amplitude of the SPOD index, which in the observations and the ensemble-mean of Init is approximately 0.5 °C, drops dramatically below 0.2 °C in the case of the ensemble-mean of NoInit. It is worth noting that this feature cannot be explained by differences in the representation of the SPOD mode of SST variability in the forecast systems, as the mode is equally simulated in the models

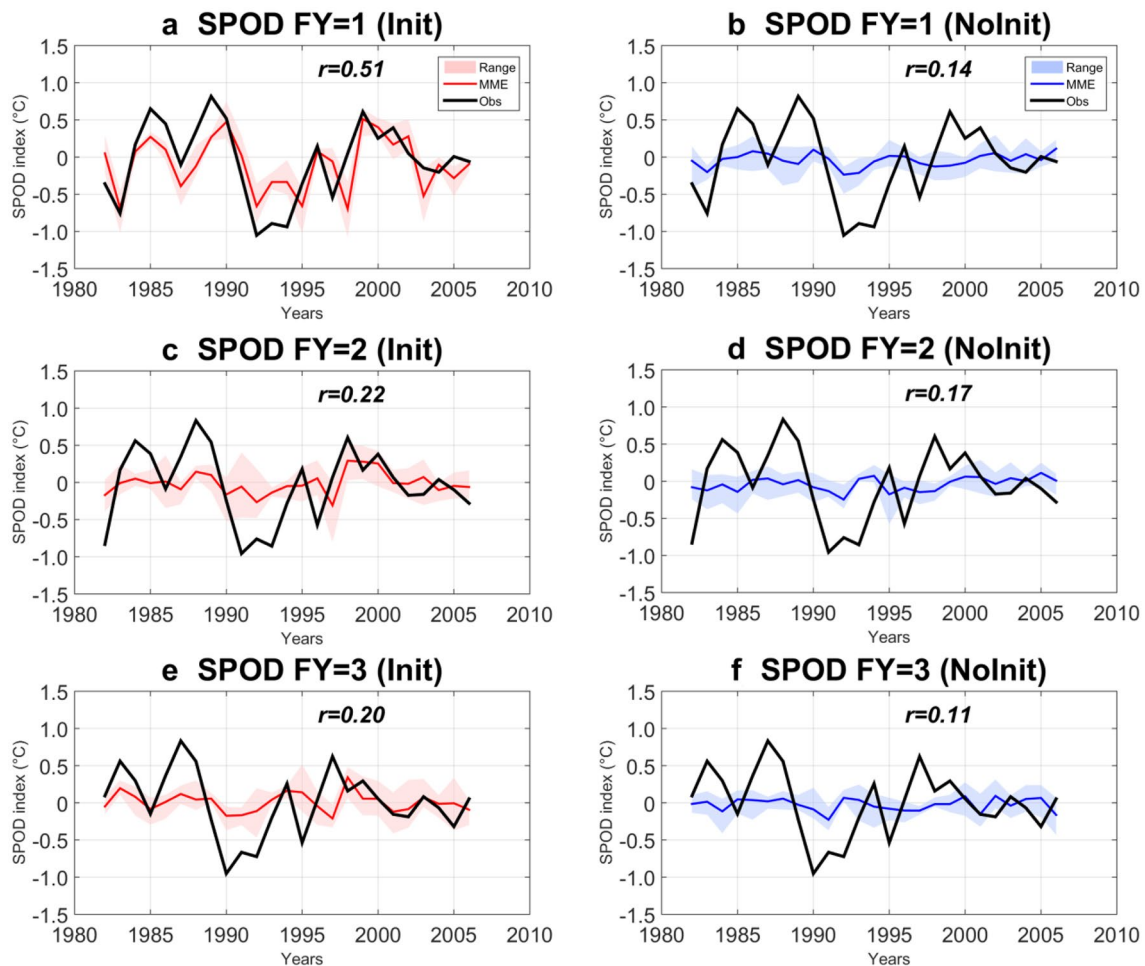


Fig. 5 Observed (black thick curve) and multi-model ensemble mean predicted evolution of the SPOD index (see text for its definition) in (left column) Init and (right column) NoInit hindcasts, for forecast years (top to bottom) 1, 2 and 3. The red and blue bands show the

ensemble spread in each case. Numbers in the upper part of each figure indicate the correlation coefficient between the observed and predicted time series. Values between 0.39 and -0.39 are not significantly different from zero ($\alpha=5\%$)

regardless initialization (Fig. 9 in the “Appendix”). Instead, this result suggests that the SPOD is mostly part of internal variability and is not radiatively forced; yet, while in NoInit there is no consistency among the members/forecast systems (i.e. the ensemble-mean/multi-model ensemble-mean cancel each other out), initialization leads the Init hindcasts to be in phase among them and with observations. Note that ACC decreases with forecast time in Init, although it remains above that of NoInit and the predicted SPOD index systematically keeps depicting coherence in the phase, but not statistically significant.

Figure 6 shows the ACC evolution of the SPOD index as a function of forecast time using persistence (black) as well as Init (red) and NoInit (blue) from the different forecast systems (thin lines), respectively; also displayed is the

skill of the Init and NoInit multi-model ensemble-means (thick lines). In the case of persistence (and in agreement with the results discussed in Fig. 3) high ACC values are obtained in forecast year 1, comparable to Init scores, followed by a rapid decline. The negative correlations in forecast years 7–8 are consistent with the oscillation of this mode at periods around 8 years (Saurral et al. 2018). Regarding the skill of the forecast systems, it is clear that initialization improves the prediction of the SPOD variability, and not only in the first forecast year, for which is statistically significant, but also up to the second or third forecast year, although by then it is no longer statistically significant. Note that the evolution of the SPOD index in NoInit (Fig. 5, second column) translates into no skill at all in these three forecast years (Fig. 6), thereby confirming

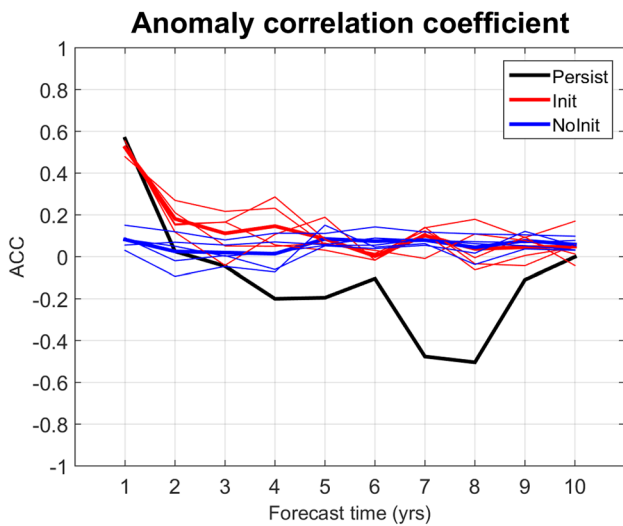


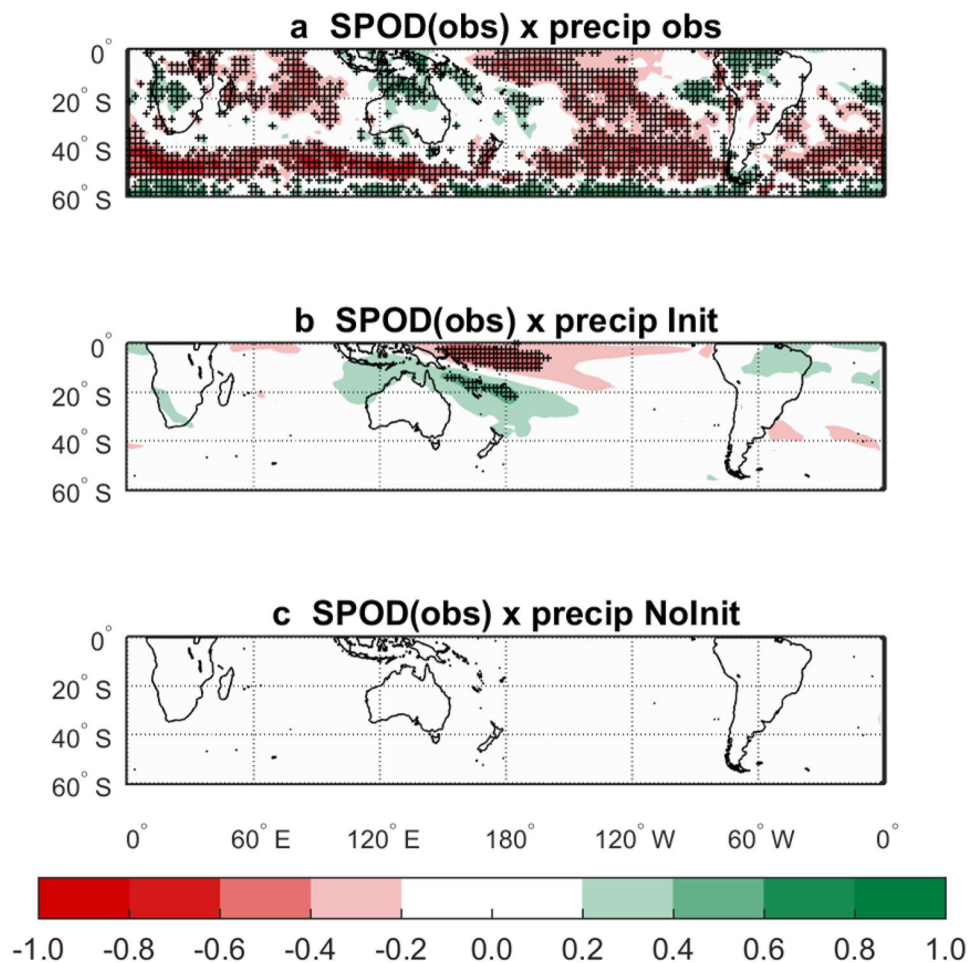
Fig. 6 Ensemble-mean ACC for each forecast system against the observed SPOD index as a function of the forecast time, considering (black curve) persistence of SST anomalies, (red curves) Init hindcasts and (blue curves) NoInit hindcasts, as a function of the forecast year

that the role of the radiatively-forced variability in the SPOD is minor. Likewise, the Init hindcasts outperform persistence in forecast years 2–3. However, the grid-point SST skill displayed in Fig. 3 (cf. Init vs. persistence) is not associated with statistically significant SPOD skill at these forecast times. Hence, together with the results in Fig. 2, it is envisaged that there is room for improvement in near-term climate forecasting up to 2–3 years ahead in the SP basin since the predictability does not yet lead to prediction skill.

3.3 Skill of the relationship between SPOD and surface climate anomalies

The objective of this section is to identify the benefits, if any, of initialization in hindcasting the observed SPOD impact on temperature and precipitation anomalies over the Southern Hemisphere. It aims to assess the potential translation of SPOD skill in Fig. 6 into skill at capturing the SPOD influence on surface climate; the approach unambiguously identifies grid points in which the observed

Fig. 7 a Correlation coefficient between observed SPOD index and observed annual mean precipitation anomalies. Multi-model ensemble mean correlation maps of hindcast annual mean precipitation anomalies in the b Init and c NoInit hindcasts against the observed SPOD index in the first forecast year. The black crosses highlight those correlation coefficients that are significant at the 10% confidence level according to a two-tailed t-test

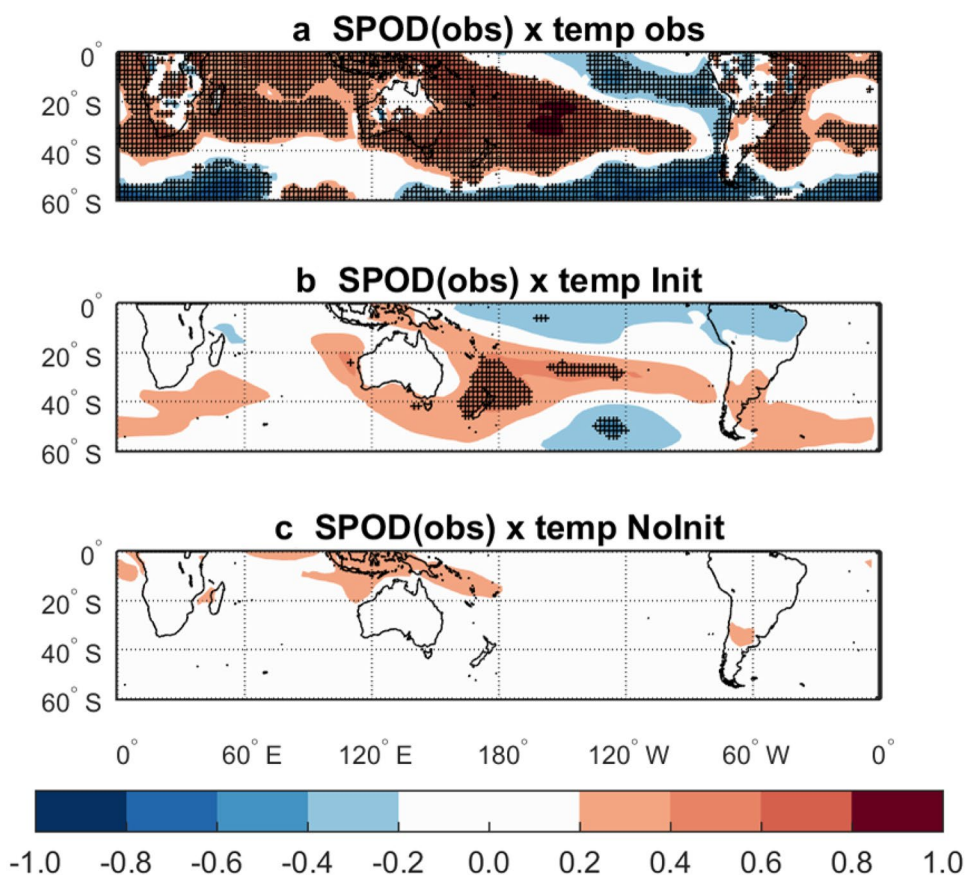


SPOD fluctuations are more representative of the hindcast variability (see Sect. 2). The analysis is restricted to only forecast year 1 and encompasses both local and remote (teleconnection-driven) impacts on the atmospheric variability. Figure 7 shows the correlation map between the observed SPOD index and observed (Fig. 7a) as well as multi-model ensemble mean hindcast precipitation anomalies for Init (Fig. 7b) and NoInit (Fig. 7c). In the observations, the SPOD is highly and significantly correlated with precipitation anomalies over much of the Southern Hemisphere, most markedly over the ocean areas and specific parts of South America, South Africa and northern Australia (Fig. 7a). In the hindcasts, a remarkable feature of Init is that the SPOD skill is associated with large correlation coefficients around Australia and over the western tropical Pacific, depicting a dipolar-type pattern of rainfall anomalies. Interestingly, there are also positive correlations over northern South America in all the forecast systems (not shown), although the signal is not statistically significant. Hints of an opposite relationship are present over the southwestern South Atlantic region, but overall these correlations are not significant either. Teleconnections forced

by ENSO variability explain part of these relationships, and they are correctly captured by the initialized hindcasts (not shown). Indeed, ENSO variability has important effects on rainfall anomalies over Australia (e.g. Ropelewski and Halpert 1987; Power et al. 2006), interpreted as a latitudinal migration of the South Pacific Convergence Zone (SPCZ), away from the continent during El Niño conditions and into northern Australia during La Niña years. Also, the region with high correlations covering northern South America is affected by the remote influence of ENSO (e.g. Vera et al. 2004; Kayano et al. 2009; Krishnamurthy and Misra 2010; Tedeschi and Collins 2016; García-Serrano et al. 2017), as well as southeastern South America (e.g. Ropelewski and Halpert 1987; Grimm et al. 2000). Regarding NoInit, the correlations are noticeably smaller, not even reaching 0.2, and not significant; this is consistent with the internal nature of the SPOD-related variability and thus the lack of skill discussed above.

The correlation between the observed SPOD index and 2-m air temperature anomalies, both in observations and the forecast systems, is analyzed in Fig. 8, following a similar approach as for precipitation. Again, a clear impact

Fig. 8 As in Fig. 7 but for annual mean temperature anomalies



arises from initialization, as no significant correlation is achieved in NoInit (Fig. 8c), in agreement with its lack of SPOD skill coming from the inconsistency among members/forecast systems. In the Init hindcasts (Fig. 8b), as in observations (Fig. 8a), the correlations show a wave-like structure arching over the SP region. This feature is particularly robust in DePreSys, EC-Earth and GFDL, while in HadCM3 it is a bit distorted (not shown). These results could be explained by the ability of the forecast systems to simulate the SPOD spatial pattern (see Fig. 9), in which HadCM3 is less spatially coherent than the other models. Together with the analysis of precipitation, the results illustrate that initialization provides skill in forecast year 1 for surface climate over the Southern Hemisphere, in association with the SPOD variability and likely mediated by ENSO.

4 Conclusions

This study assessed predictability and prediction skill in a set of decadal prediction hindcasts from state-of-the-art forecast systems in the SP region. Focus was made on quantifying their skill to predict the first mode of SST variability in the region. The assessment was performed for both, the skill in SST predictions and the related skill in representing the teleconnections with surface temperature and precipitation over the Southern Hemisphere.

It is shown that the forecast systems realistically simulate the leading mode of SST variability associated with a dipole of anomalies between subtropical and extratropical latitudes. This ability is present regardless initialization, indicating that the climate dynamics of the SPOD are properly represented in the models. Results also show that initialization provides added skill in the prediction of the SPOD and its variability beyond the prescribed radiative forcing. In fact, a comparison between the Init and NoInit hindcasts shows that Init has higher skill in forecasting the SPOD up to the third forecast year, also largely outperforming the skill of empirical predictions based on persistence beyond the first forecast year. However, the SPOD skill in forecast years 2–3 is not statistically significant. Analysis of predictability based on signal-to-noise variances suggests that larger prediction skill levels might be achievable at least for the second forecast year, which encourages further work in near-term prediction beyond ENSO. A first test-suite for improvement may be the upcoming Decadal Climate Prediction Project (DCPP; Boer et al. 2016) contribution to CMIP6.

There is a clear translation of the SPOD skill into capturing the relationship with surface climate thanks to initialization. The Init hindcasts proved to be skillful

at representing the observed relationship between the SPOD and temperature and precipitation anomalies, mostly thanks to the contribution from ENSO-related variability. This could serve as a potential tool for near-term predictions of climate anomalies over the Southern Hemisphere, such as mega-drought periods in South America (Boisier et al. 2016). In the light of the results found here, the skill of such predictions could be favored under active ENSO/IPO conditions. Still, it should be stressed that, due to constraints in the observational coverage, the reliable period with available hindcasts only consists of 34 years. A longer set of initialized experiments might need to be explored in order to gain some insight into the stationarity of the forecast quality (Müller et al. 2014).

Acknowledgements The authors would like to thank the two anonymous reviewers, whose comments and suggestions led to a significant improvement of the manuscript. This study also benefited from discussion with Fabian Lienert. It was supported by the University of Buenos Aires through Grant UBACYT-20020100100803, UBACYT-2002170100428BA, CONICET through PIP2009-00444, ANPCyT through Grants PICT07-00400 and PICT2016-1898, the EUCP (GA 776613) and PRIMAVERA (GA 641727) EU-funded projects, and the CLIMAX (Belmont Forum/ANR-15-JCL/-0002-01) Project. J.G-S was funded by the Spanish “Ramón y Cajal” programme (RYC-2016-21181). The authors acknowledge the Red Española de Supercomputación (RES) and PRACE for awarding access to Marenostrum 3 at the Barcelona Supercomputing Center through the HiResClim project, as well as the World Climate Research Programme’s Working Group on Coupled Modelling and the diverse climate modeling groups for producing and making available the model outputs used in this paper. The support given by Oriol Mula-Valls and Virginie Guémas at early stages of this research is particularly appreciated.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Leading mode of SST variability in the SP region in the forecast systems

Figure 9 shows the spatial pattern associated with the leading mode of interannual SST variability in the SP region in forecast year 1 of the (left) initialized and (right) uninitialized hindcasts. The resemblance of the Init vs. NoInit patterns in each forecast system indicates that the SPOD mode is internally generated by the models. Additional discussion on this issue can be found in the main text.

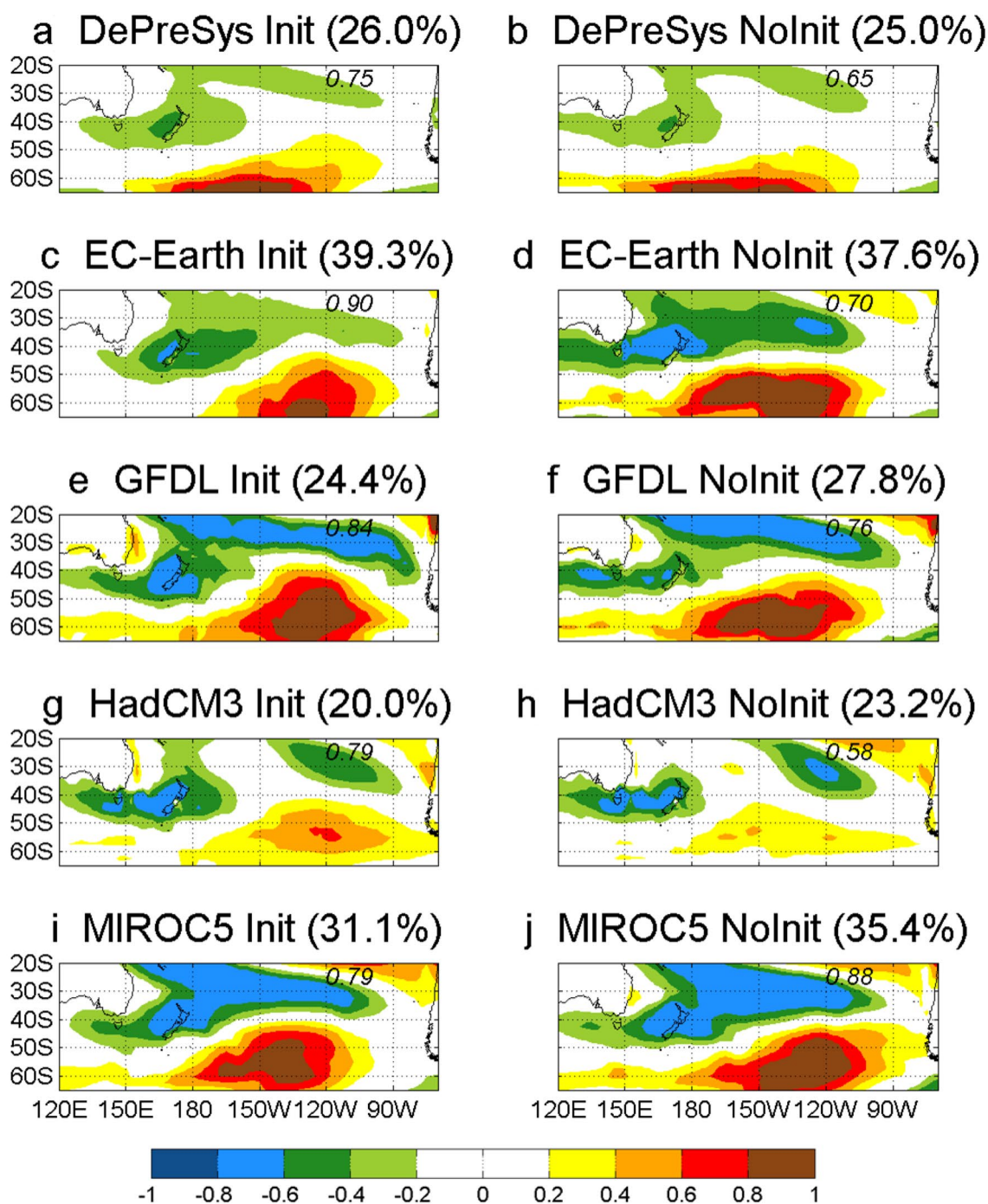


Fig. 9 Leading mode of SST interannual variability in the SP region in (from top to bottom) DePreSys, EC-Earth, GFDL, HadCM3 and MIROC5, considering Init (left column) and NoInit (right column) experiments. The numbers in the sub-titles denote the fraction of variance explained by the mode in each case. The smaller numbers

within each sub-figure indicate the spatial correlation between the simulated mode in the hindcast and the observed field. The computation of the leading mode was done taking, for each forecast system, all the ensemble members together

References

Barros VR, Silvestri GE (2002) The relation between sea surface temperature at the subtropical south-central Pacific and

precipitation in southeastern South America. *J Clim* 15:251–267
 Basher RE, Zheng X (1998) Mapping rainfall fields and their ENSO variation in data-sparse tropical south-west Pacific Ocean regions. *Int J Climatol* 18:237–251

- Boer GJ et al (2016) The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. *Geosci Model Dev* 9:3751–3777
- Boisier JP, Rondanelli R, Garreaud RD, Muñoz F (2016) Anthropogenic and natural contributions to the Southeast Pacific precipitation decline and recent megadrought in central Chile. *Geophys Res Lett* 43:413–421
- Chung CTY, Power SB, Sullivan A, Delage F (2019) The role of the South Pacific in modulating Tropical Pacific variability. *Sci Rep* 9:18311
- Ding H, Greatbach RJ, Latif M, Park W, Gerdes R (2013) Hindcast of the 1976/77 and 1998/99 climate shifts in the Pacific. *J Clim* 26:7650–7661
- 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. <https://doi.org/10.1038/ncomms2704>
- Du H, Doblas-Reyes FJ, García-Serrano J, Guemas V, Soufflet Y, Wouters B (2012) Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations. *Clim Dyn* 39:2013–2023
- García-Serrano J, Doblas-Reyes FJ (2012) On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Clim Dyn* 39:2025–2040
- García-Serrano J, Doblas-Reyes FJ, Coelho CAS (2012) Understanding Atlantic multi-decadal variability prediction skill. *Geophys Res Lett* 39:L18708
- García-Serrano J, Guémas V, Doblas-Reyes FJ (2015) Added-value from initialization in predictions of Atlantic multi-decadal variability. *Clim Dyn* 44:2539–2555
- García-Serrano J, Cassou C, Douville H, Giannini A, Doblas-Reyes FJ (2017) Revisiting the ENSO teleconnection to the tropical North Atlantic. *J Clim* 30:6945–6957
- González PLM, Goddard L (2016) Long-lead ENSO predictability from CMIP5 decadal hindcasts. *Clim Dyn* 46:3127–3147
- Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Clim Dyn* 16:147–168
- Grimm AM, Barros VR, Doyle ME (2000) Climate variability in southern South America associated with El Niño and La Niña events. *J Clim* 13:35–58
- Guan Y, Zhu J, Huang B, Hu ZZ, Kinter JL III (2014) South Pacific Ocean Dipole: a predictable mode on multiseasonal time scales. *J Clim* 27:1648–1658
- Guémas 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 Atmos* 117:D20111
- Guémas V, Corti S, García-Serrano J, Doblas-Reyes FJ, Balmaseda MA, Magnusson L (2013) The Indian Ocean: the region of highest skill worldwide in decadal climate prediction. *J Clim* 26:726–739
- Hasumi H, Emori S (2004) K-1 coupled GCM (MIROC) description. K-1 technical report 1. University of Tokyo, Tokyo, p 34
- Hou Z, Li J, Ding R, Karamperidou C, Duan W, Liu T, Feng J (2018) Asymmetry of the predictability limit of the warm ENSO phase. *Geophys Res Lett* 45:7646–7653
- Huang B, Shukla J (2006) Interannual SST variability in the southern subtropical and extra-tropical ocean. Technical report 223, Center for Ocean-Land-Atmosphere Studies, Calverton
- Jin EK, Kinter JL III, Wang B et al (2008) Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Clim Dyn* 31:647–664
- Kayano MT, de Oliveira CP, Andreoli RV (2009) Interannual relations between South American rainfall and tropical sea surface temperature anomalies before and after 1976. *Int J Climatol* 29:1439–1448
- Kim H-M, Webster PJ, Curry JA (2012) Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophys Res Lett* 39:L10701
- Kirtman B, Power S, 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 GJ, Vecchi G, Wang HJ (2013) Near-term climate change: projections and predictability. *Climate change 2013: the physical science basis*. In: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, pp 953–1028. <https://doi.org/10.1017/CBO9781107415324.023>
- Krishnamurthy V, Misra V (2010) Observed ENSO teleconnections with the South American monsoon system. *Atmos Sci Lett* 11:7–12
- Lienert F, Doblas-Reyes FJ (2013) Decadal prediction of interannual tropical and North Pacific sea surface temperature. *J Geophys Res Atmos* 118:5913–5922
- Linsley BK, Wellington GM, Schrag DP, Ren L, Salinger MJ, Tudhope AW (2000) Geochemical evidence from corals for changes in the amplitude and spatial pattern of South Pacific interdecadal climate variability over the last 300 years. *Clim Dyn* 22:1–11
- Lou J, Holbrook NJ, O’Kane TJ (2019) South Pacific decadal climate variability and potential predictability. *J Clim* 32:6051–6069
- Müller WA, Pohlmann H, Sienz F, Smith D (2014) Decadal climate prediction for the period 1901–2010 with a coupled climate model. *Geophys Res Lett* 41:2100–2107
- Osman M, Vera CS (2020) Predictability of extratropical upper-tropospheric circulation in the Southern Hemisphere by its main modes of variability. *J Clim* 33:1405–1421
- Power S, Casey T, Folland C, Colman A, Mehta V (1999) Inter-decadal modulation of the impact of ENSO on Australia. *Clim Dyn* 15:319–324
- Power S, Haylock M, Colman R, Wang X (2006) The predictability of interdecadal changes in ENSO activity and ENSO teleconnections. *J Clim* 19:4755–4771
- Power S, Saurral RI, Chung C, Colman R, Kharin V, Boer G, Gergis J, Henley B, McGregor S, Arblaster J, Holbrook N, Liguori G (2017) Towards the prediction of multi-year to decadal climate variability in the Southern Hemisphere. *Past Glob Changes Mag* 25:32–40
- Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res D Atmos*. <https://doi.org/10.1029/2002JD002670>
- Ropelewski CF, Halpert MS (1987) Global and regional scale precipitation patterns associated with El Niño/Southern Oscillation. *Mon Wea Rev* 115:1606–1626
- Saurral RI, Doblas-Reyes FJ, García-Serrano J (2018) Observed modes of sea surface temperature variability in the South Pacific region. *Clim Dyn* 50:1129–1143
- Schubert SD, Suarez MJ, Pegion PJ, Kistler MA, Kumar A (2002) Predictability of zonal means during boreal summer. *J Clim* 15:420–434
- Shakun JD, Shaman J (2009) Tropical origins of North and South Pacific decadal variability. *Geophys Res Lett* 36:L19711
- Smith DM, Eade R, Dunstone NJ, Fereday D, Murphy J, Pohlmann H, Scaife AA (2010) Skilful multi-year predictions of Atlantic hurricane frequency. *Nat Geosci* 3:846–849

- Smith DM, Eade R, Scaife AA, Caron LP, Danabasoglu G, DelSole TM, Delworth T, Doblas-Reyes FJ, Dunstone NJ, Hermanson L, Kharin V, Kimoto M, Merryfield WJ, Mochizuki T, Müller WA, Pohlmann H, Seager S, Yang X (2019) Robust skill of decadal climate predictions. *Npj Clim Atmos Sci* 2:13
- Srivastava A, DelSole T (2018) Decadal predictability without ocean dynamics. *PNAS* 114:2177–2182
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Am Meteorol Soc* 93:485–498
- Tedeschi RG, Collins M (2016) The influence of ENSO on South American precipitation during austral summer and autumn in observations and models. *Int J Climatol* 36:618–635
- Vera CS, Silvestri GE, Barros VR, Carril AF (2004) Differences in El Niño response over the Southern Hemisphere. *J Clim* 17:1741–1753
- Volpi D, Doblas-Reyes FJ, García-Serrano J, Guémas V (2013) Dependence of the climate prediction skill on spatio-temporal scales: internal versus radiatively-forced contribution. *Geophys Res Lett*. <https://doi.org/10.1002/grl.50577>
- Volpi D, Guémas V, Doblas-Reyes FJ, Hawkins E, Nichols NK (2017) Decadal climate prediction with a refined anomaly initialization approach. *Clim Dyn* 48:1841–1853
- Yang X, Rosati A, Zhang S, Delworth TL, Gudgel RG, Zhang R, Vecchi GA, Anderson WG, Chang YS, DelSole T, Dixon KW, Msadek R, Stern WF, Wittenberg AT, Zeng F (2013) A predictable AMO-like pattern in GFDL's fully-coupled ensemble initialization and decadal forecasting system. *J Clim* 26:650–661
- Zhang L, Delworth TL, Yang X, Gudgel RG, Jia L, Vecchi GA, Zeng F (2017) Estimating decadal predictability for the Southern Ocean using the GFDL CM2.1 model. *J Clim* 30:5187–5203

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