

Predictability and prediction of persistent cool states of the Tropical Pacific Ocean

Nandini Ramesh^{1,2}  · Mark A. Cane² · Richard Seager² · Dong Eun Lee²

Received: 9 July 2016 / Accepted: 6 November 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The Tropical Pacific Ocean displays persistently cool sea surface temperature (SST) anomalies that last several years to a decade, with either no El Niño events or a few weak El Niño events. These cause large-scale droughts in the extratropics, including major North American droughts such as the 1930s Dust Bowl, and also modulate the global mean surface temperature. Here we show that two models with different levels of complexity—the Zebiak–Cane intermediate model and the Geophysical Fluid Dynamics Laboratory Coupled Model version 2.1—are able to produce such periods in a realistic manner. We then test the predictability of these periods in the Zebiak–Cane model using an ensemble of experiments with perturbed initial states. Our results show that in most cases the cool mean state is predictable. We then apply this method to make retrospective forecasts of shifts in the decadal mean state and to forecast the mean state of the Tropical Pacific Ocean for the upcoming decade. Our results suggest that the Pacific will undergo a shift to a warmer mean state after the 2015–2016 El Niño. This could imply the cessation of the drier than normal conditions that have generally afflicted southwest North America since the 1997–1998 El Niño, as well as the twenty-first-century pause in global warming. Implications for our understanding of the origins of such persistent cool states and the possibility of improving predictions of large-scale droughts are discussed.

1 Introduction

The Tropical Pacific Ocean is a dominant force in global climate variability. The coupling between atmosphere and ocean dynamics in this region gives rise to a richly complex system that is able to produce a wide range of behaviors on a variety of timescales. The most famous of these is its interannual variability: El Niño–Southern Oscillation (ENSO). ENSO is an oscillation of the SSTs in the eastern and central parts of the equatorial Pacific between warm and cool phases, known as El Niño and La Niña events respectively. This oscillation has a period of 2–7 years, and there is a well-known asymmetry between the duration of La Niña events, which tend to last 1–2 years, and El Niño events, which usually last only for 2–4 seasons (Okumura and Deser 2010). A second observed mode of variability is the less well-understood Pacific Decadal Variability (PDV), occurring on decadal to multidecadal timescales, which is associated with El Niño-like conditions in its positive phase and La Niña-like conditions in its negative phase. In recent decades the Tropical Pacific shifted from a warm PDV phase between the 1976–1977 and 1997–1998 El Niño events to a protracted cool state that lasted at least until the 2015–2016 El Niño event. Both of these modes of variability have consequences for large parts of the globe via atmospheric teleconnections. On an intermediate timescale, there are also persistent La Niña-like states during which the SSTs in the eastern Tropical Pacific maintain a cool mean state over several years—longer than the usual duration of a La Niña event—with either no El Niño events, or a few very weak El Niño events. In a study of persistent droughts in North America, Seager et al. (2005) made note of this type of behavior in the instrumental record spanning the following periods: 1856–1865, 1870–1877, 1890–1896,

✉ Nandini Ramesh
nramesh@ldeo.columbia.edu

¹ Department of Earth and Environmental Sciences,
Columbia University, New York, NY, USA

² Lamont-Doherty Earth Observatory, Columbia University,
105F Oceanography, 61 Route 9W, Palisades, NY 10964,
USA

1932–1939 and 1948–1957; and then 1998–2002 (Seager 2007).

While the origins of the interannual variability are relatively well-understood, there is no clear consensus on the mechanisms that give rise to Tropical Pacific variability on longer timescales. Several proposed hypotheses rely on variability external to the region to produce decadal to multidecadal variability as a forced response through either oceanic or atmospheric teleconnections. The midlatitude Pacific, where the multidecadal signal was first observed (Graham 1994; Mantua et al. 1997), has been considered a candidate for the origins of this variability under a few hypotheses. It was proposed that temperature anomalies in the midlatitudes could propagate to the Tropics along the thermocline via advection by the subtropical cell (Deser et al. 1996; Gu and Philander 1997), where they may be able to alter the background state of the equatorial ocean, but subsequent work has shown that these anomalies are too small compared to local wind forcing to exert a significant influence on the tropical ocean on decadal timescales (Schneider et al. 1999). Another way for variability in the midlatitudes to influence the Tropics is by producing changes in the trade winds (Barnett et al. 1999) that generate SST anomalies via air–sea fluxes (Vimont et al. 2003).

Another set of hypotheses places the origins of longer-than-interannual variability of the Tropical Pacific entirely outside of the Pacific basin, identifying a role for atmospheric teleconnections from the Atlantic basin (Dong et al. 2006; Kang et al. 2014) or even an influence from volcanic aerosols (Adams et al. 2003).

However, it is possible that the coupled ocean–atmosphere system of the Tropical Pacific itself produces variability on these timescales. A study using a wind-forced shallow water model of the Pacific basin demonstrated that the variability of the equatorial thermocline on decadal-to-multidecadal timescales is dominated by wind stress forcing within 20° north or south of the equator (McGregor et al. 2007). This is consistent with a scaling analysis based on linear wave theory (Emile-Geay and Cane 2009) which revealed that midlatitude wind forcing is unable to exert any significant influence on the equatorial thermocline on timescales less than 50 years. These studies point to wind stress anomalies and oceanic planetary waves within the Tropical Pacific as being responsible for decadal variability. An alternative perspective on this phenomenon is provided by dynamical systems theory: the chaotic nature of the coupled ocean–atmosphere system of the Tropical Pacific can give rise to an oscillation between different regimes of behavior on decadal timescales (Timmermann and Jin 2002; Tziperman et al. 1994). Another possibility is that stochastic atmospheric variability within the Tropical Pacific region produces shifts in ENSO properties on decadal timescales (Flugel and Chang 1999).

Although the decadal signal is strongly expressed in the extratropical North Pacific, on which the Pacific Decadal Oscillation (PDO) index is based, this extratropical signal has been explained as a response to the decadal modulation of ENSO in the tropics (Zhang et al. 1997; Newman and Alexander 2003; Chen and Wallace 2015; Newman et al. 2016). Thus, while a variety of factors may play a role, it is not necessary to invoke external influence to explain the long-term behavior of the Tropical Pacific.

The mechanisms for persistent cool states have not been explored in as much depth, and it is possible that they may originate from phenomena similar to decadal variability. Persistent cool anomalies in the Tropical Pacific have become a subject of current interest because of the recent behavior of the Tropical Pacific, and because of their effects on hydroclimate around the globe. It has been argued that the global precipitation footprint of these cool states is similar to that of a La Niña event (Seager 2015). They have been shown to be responsible for prolonged droughts in southwestern North America, such as the Dust Bowl drought of the 1930s (Seager et al. 2005; Herweijer et al. 2006). Concurrent with these droughts were droughts in South America that affected Uruguay, southern Brazil, and northern Argentina (Herweijer and Seager 2008). Parts of Eastern Europe extending to Central Asia (Hoerling and Kumar 2003) have also experienced simultaneous droughts, along with Western Australia, East Africa, southern India and Sri Lanka (Lyon and DeWitt 2012; Herweijer and Seager 2008; Lyon 2014; Yang et al. 2014). Figure 1 shows the SST and precipitation anomalies across the globe (see Sect. 2 for sources of data) for four such events: while global patterns of precipitation are not identical across all of the events, the areas mentioned above consistently experience drought during these La Niña-like periods, suggestive of global atmosphere–ocean regimes orchestrated by the Tropical Pacific. For comparison, the lower two panels of Fig. 1 show the same for a positive and negative phase of the PDV. While the effects on North American and East African hydroclimate are similar, the conditions in many of these regions (e.g., Australia) are quite different between the negative phase of the PDV and the persistent cool states, which may be because the response is different or because of masking by other variability.

There are some subtle differences between the events themselves, as well as their impacts. They range in length from 7 to 10 years. A few of the events, such as that of the 1890s, 1950s and 1999–present, were not uniformly cool but had a cool mean state interrupted by El Niño events of small magnitude; while others, such as those of the 1870s and the 1930s were uninterrupted in their cool state. These periods also display an association with large El Niño events, with four of the six listed periods (1870–1877, 1890–1896, 1932–1939, and 1999–2014) beginning

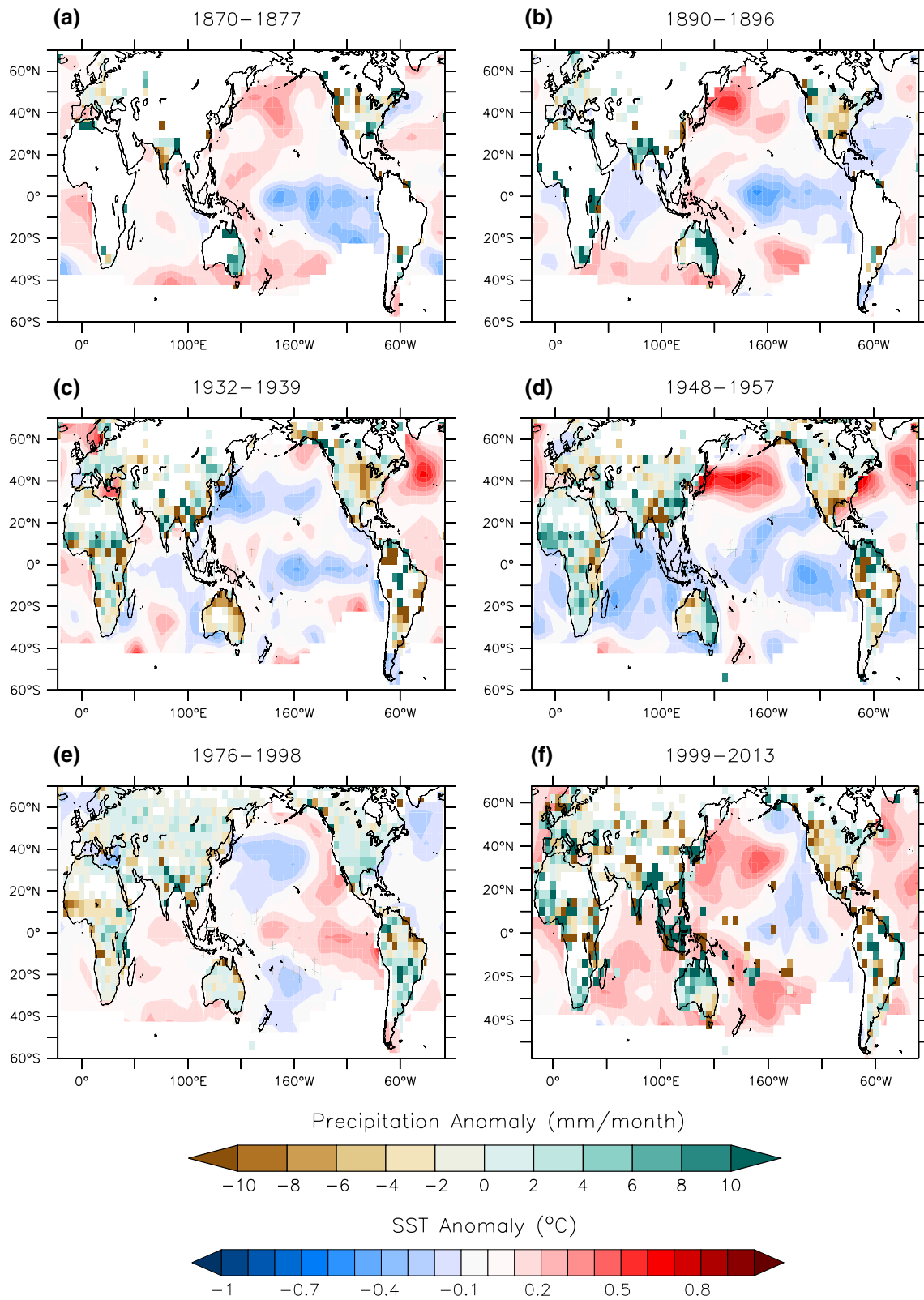


Fig. 1 a–d show de-trended SST and precipitation anomalies averaged over four persistent cool periods of the Tropical Pacific in the observational record. e and f Depict the same for periods of positive (1976–1998) and negative (1999–2013) PDV respectively

immediately after a high-magnitude El Niño event, and ending in a similar event (if the recent period was indeed ended by the 2015–2016 El Niño event).

Given the global impacts of these cool, quiescent periods in the Tropical Pacific, it is important to gain a better understanding of the dynamics that produce them, especially in the context of predicting them. Recently, cool states in the Tropical Pacific have been shown to play a role in modulating the global mean surface temperature (Meehl et al. 2011; Kosaka and Xie 2013; Delworth et al. 2015) and are therefore implicated in the “global warming hiatus” of the early 2000s: a period during which the warming of the global mean surface temperature has decelerated from the rapid warming of the 1970–1990s (Easterling and Wehner 2009). This provides further incentive to investigate this particular phenomenon. While climate models are able to simulate the interannual variability of the Tropical Pacific with varying degrees of success, it has not been shown that they are also able to capture extended cool periods in a realistic manner. In this study, we begin by identifying periods similar to the prolonged cool states in the observations in long unforced runs from two models of differing complexity (Sect. 3): the Zebiak–Cane model (ZC Model), and the Geophysical Fluid Dynamics Laboratory’s Coupled Model version 2.1 (GFDL Model). The ZC model is an intermediate-complexity model which simulates anomalies in the Tropical Pacific with respect to a prescribed mean state and has been shown to display decadal variability (Cane et al. 1995; Karspeck et al. 2004). The GFDL model is a full coupled general circulation model (GCM) which has been previously used to study the predictability of PDV (Wittenberg et al. 2014). We then assess the predictability of these periods in the ZC Model (Sect. 4). With the knowledge that the system has some long-term predictability both from previous work examining decadal variability (Karspeck et al. 2004) and our own results, we then apply our method of perturbed ensemble predictions to make retrospective forecasts of shifts in the mean state of the Tropical Pacific, and to forecast its state for the upcoming decade (Sect. 5).

2 Models and data

The two models used in this study are described below:

2.1 The Zebiak–Cane model

The Zebiak–Cane Model (Zebiak and Cane 1987) is an intermediate coupled model of the Tropical Pacific Ocean with a simplified global atmosphere and was the first model to produce a successful forecast of an El Niño event (Cane et al. 1986). Despite its simplicity compared to GCMs, it remains a useful tool in modelling ENSO both for scientific

analysis and operational forecasts (Chen et al. 2004). It has also been shown to exhibit the variability in this region over longer timescales (Cane et al. 1995). The model simulates the ocean above the thermocline in the region 124°E–80°W, 29°S–29°N with a deeper, motionless lower layer and, within the dynamic upper layer, a frictional Ekman layer of 50 m depth. It uses an adaptation of the Gill model of the atmosphere (Zebiak 1982) that responds to the SSTs produced by the ocean model, but does not have internal atmospheric noise. The Zebiak–Cane model is an anomaly model, meaning that it computes the anomalies of all variables with respect to a prescribed climatology. This model was run for a period of 100,000 years with no forcing. The use of this particular model offers a number of advantages. The first of these is its computational efficiency: it is fast to run and can therefore be used to produce large ensembles for analysis. By simulating anomalies with respect to a prescribed seasonal cycle, this model also avoids the drawbacks of having to simulate the seasonal cycle of the equatorial region accurately, which remains a challenge for most GCMs. The fact that many hypotheses point to mechanisms within the Tropical Pacific being responsible for its longer-than-interannual variability means that the decadal variability observed in the model (Cane et al. 1995) is likely to bear resemblance to the real-world variability despite its lack of an extratropical ocean; if the model instead is unable to simulate persistent cool states or decadal variability comparable to that in the observations, it may indicate that the extratropics or factors external to the Pacific basin are indeed a key driver of this phenomenon.

2.2 The GFDL coupled model (version 2.1)

This is a coupled GCM that is designed to capture the full complexity of climate variability with atmosphere, ocean, land, and ice components; and is among the models used in the Coupled Model Intercomparison Projects 3 and 5 (Delworth et al. 2006). A 4000-year unforced simulation from this model with no variability in solar or aerosol forcing (Wittenberg et al. 2014) was analyzed in this study. It is useful to examine persistent cool states in this particular model as it has been used to study the Tropical Pacific system both on decadal (Wittenberg et al. 2014) and interannual (Kug et al. 2010; Takahashi et al. 2011) timescales. A comparison of this model with the observations and the ZC model can also be instructive: if this model simulates persistent cool states with substantially more accuracy than the ZC model can, it would imply that the extratropics and the Atlantic Ocean in particular may in fact be crucial to generating this variability. If neither the ZC model nor this model are able to produce this variability, there may be something fundamental missing from the simulations, or it may be the case that radiative forcing in some form is required to give

rise to persistent cool states. The observational datasets used were the Kaplan Extended Sea Surface Temperature Version 2 (KSST) (Kaplan et al. 1998), a gridded SST dataset that spans the period from 1856 to the present, and station precipitation data from the Global Historical Climatology Network (GHCN) (Vose et al. 1992) developed and maintained by the National Oceanic and Atmospheric Administration's National Climate Data Center (NOAA NCDC) and the U.S. Department of Energy's Carbon Dioxide Information Analysis Center (DoE CDIAC).

3 Persistent cool states in the ZC model and GFDL model

Our first step was to identify whether analogs to the observed persistent cool states exist in the model simulations. We use the Niño 3 index as an indicator of the state of the Tropical Pacific in this study. This index is calculated as the time series of the spatial average of SSTs in the region spanning 90°W–150°W, 5°S–5°N in the eastern Tropical Pacific with the seasonal cycle removed, and is a common measure of ENSO variability. Before comparing the output from the model simulations with the observations, the Niño 3 time series from each of these were filtered using a low-pass Butterworth filter with a cutoff period of 4 months. Since this work focuses on longer-term variability, the high-frequency component of the variability in the observations can be considered noise, and was filtered out in order to make the two time series comparable.

Four long La Niña events were used to select analogs from the 100,000-year time series: 1870–1877, 1890–1896, 1932–1939, and 1948–1957. These were all periods of a mean cool state in the Tropical Pacific that coincided with major droughts in western North America as identified in previous studies (Seager et al. 2005; Herweijer and Seager 2008). For each of these events, the segment of the filtered Niño 3 time series beginning 3 years prior to the start of the event until the end of the event (for example, for the first event, the segment spanning January 1867 to December 1877) from the observations was correlated with each possible continuous segment of the same length from the filtered Niño 3 index from the models.

Figure 2 shows the best analog—i.e., the segment having the highest correlation with the observed time series—for each event in each of the models along with the Niño 3 time series from the observations. These high-correlation analogs not only have similar trajectories to the observed Niño 3 time series, but also capture the mean cool state for the duration of the event. Interestingly, in all cases the ZC model shows higher correlation coefficients than the GFDL model despite being the simpler of the two models, and is able to capture the quiescence of the observed Niño 3 time

series more accurately. This is possibly a consequence of the much longer time series from the ZC model.

The distributions of correlation coefficients calculated in this way for every continuous segment beginning in January of the model simulations for each of the events is shown in Figs. 3 and 4 for the ZC and GFDL models respectively. The same procedure was then repeated using the segment from the model simulation that had the highest correlation with the observations (the “best analog”) for each of the four events. The best analog, being from the simulation itself, attempts to capture the mode of variability produced by the model that most resembles the cool periods in the real world. Comparing the model run with this segment sheds light on whether the model reproduces itself much better than it does reality.

Figures 3 and 4 provide us with information on three aspects of the models' ability to reproduce the cool periods observed in the real world. Firstly, the fact that there do exist several instances of correlation coefficients greater than 0.5 indicates that both models are able to produce periods similar to the observed cool states that are repeatable and not just stray events. Secondly, the fact that the bulk of the segments in all of the cases correlate poorly or not at all with the observed cool periods is consistent with the models capturing these as a small subset of events, rather than the simulations being dominated by states resembling the persistent La Niña-like periods. A third piece of information comes from the blue lines: the distributions of correlation coefficients between segments from the model and its own best analog. In the GFDL model, these distributions show a distinct shift to the right, i.e., an increase in the correlation coefficients. Although the ZC model does not have this shift to the right, the blue curves display distributions with fatter tails than the black curves, meaning that the number of highly-correlated segments is higher when the best analog from the model itself is used. Since the models are imperfect in their representation of reality, they can be expected to favor a resemblance to their own behavior more than to the observations, giving rise to this shift. The less the model analogs look like reality, the more pronounced this shift to the right would be. The fact that using the best analog instead of the observed events still produces distributions with only a small fraction of the correlation coefficients being higher than 0.5 shows that the models capture the rarity and distinctiveness of the extended La Niña-like states in the observations.

In order to facilitate a comparison between the number of analogs found in the two model runs of different lengths, we split the 100,000 years of the ZC model output into 25 continuous segments of 4000 years (the same length as the GFDL model output), and calculated similar distributions of correlation coefficients using each of the 25 segments. Figure 5 shows the spread of

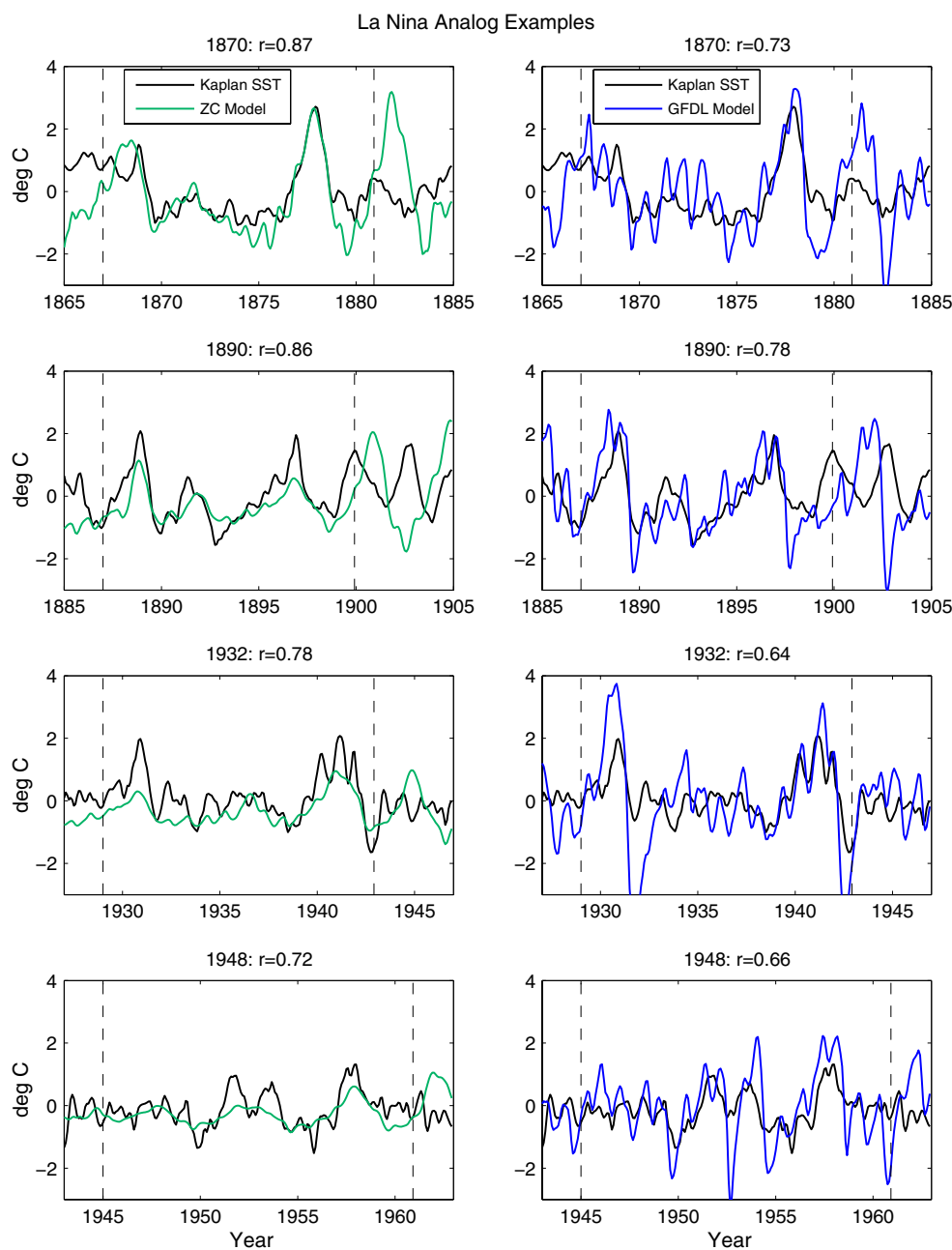


Fig. 2 The best analog as chosen based on the correlation coefficient (r) from the ZC Model (*left column*) and the GFDL Model (*right column*) for each of the four events, with the constraint that the Niño 3

value remains below $2\text{ }^{\circ}\text{C}$. The *dashed lines* indicate the beginning (3 years prior to the start of the event) and end (the end of the event) of the segments for which the correlation is shown

these distributions along with the corresponding distribution calculated from the GFDL model output for correlation coefficients greater than 0.5. The distribution of the high correlations in the GFDL output stays, for the most part, within the range of that of the 25 ZC model segments. This suggests that the higher correlation coefficients obtained from the ZC model may simply be a consequence of the ZC model run being far longer than the GFDL model run, and that we cannot conclude that

the ZC model has more skill in producing analogs to the observed cool states.

4 Predictability of persistent cool states in the ZC model

Since the models do capture the persistent cool states, it is useful to determine whether these states are predictable

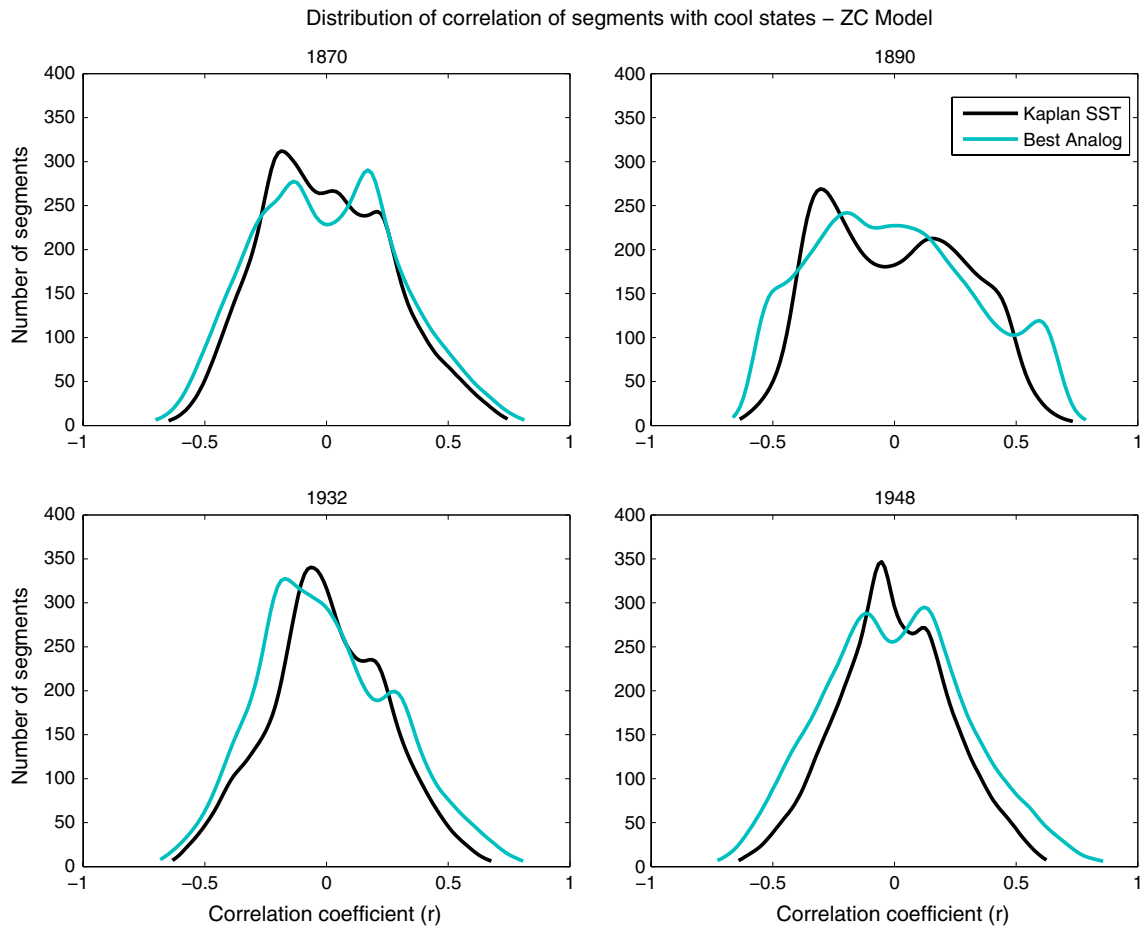


Fig. 3 The distribution of correlation coefficients obtained when the Niño 3 time series from each of the cool periods (1870–1877, 1890–1896, 1932–1939, 1948–1957) was correlated against each segment of the Niño 3 time series from the 100,000-year ZC model output

(black line). The blue line shows the distribution of correlation coefficients when the same calculation is performed using the best analog for each of the events from the model

in the models, or too sensitive to initial conditions to be predicted. We use ensembles of simulations from the ZC model in order to answer this question. This was done using a method similar to that of Karspeck et al. (2004). For each of the four events, twenty analogs were selected. There were two criteria applied in the selection of these analogs:

1. The analogs must not have an SST anomaly in their Niño 3 time series greater than 2 °C from the beginning to the end of the cool period. This criterion was applied in order to ensure that there were no large El Niño events within the cool period of the analog, as this would disqualify it from being similar to the cool periods in the observations.
2. Among the analogs that remain below 2 °C, the twenty with the highest correlation coefficients when correlated with the observations for each of the events were selected.

For each of these 80 selected analogs, we ran a set of 100 perturbed experiments. The initial state in these experiments was set as the state of the model 1 year prior to the start of the event in the analog with an added perturbation that was different for each of the 100 simulations. The perturbation was specified as a change to the initial SST field: at each spatial point, a number was randomly selected from a uniform distribution with a mean of zero and a standard deviation equal to one-fifth of the standard deviation of the full 100,000-year time series of SST anomalies at that spatial point, and added to the initial SST field at that point. Thus for each of the four events an ensemble of 2000 (20 analogs × 100 different perturbed initial states) simulations was run, giving a total of 8000 simulations. Each of these simulations was run for a period of 20 years.

Figure 6 shows the root mean squared error (RMSE) of the ensembles of predictions with respect to the time series that they were trying to predict. As expected, the predictions are unable to reproduce similar trajectories to the

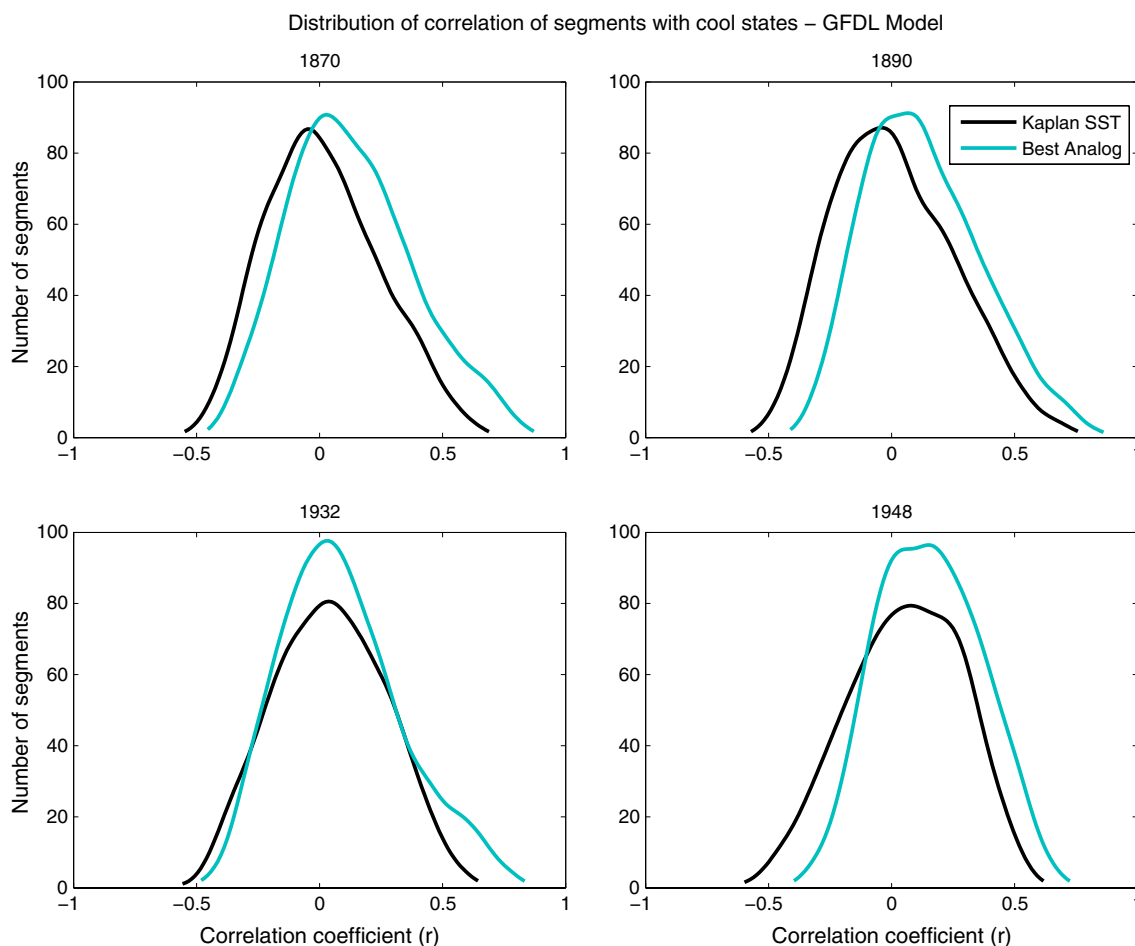


Fig. 4 As in Fig. 3, using the 4000-year simulation from the GFDL model

observed events beyond a few seasons—the RMSE values in all cases reach values near $1\text{ }^{\circ}\text{C}$ or greater at timescales of a few years or more. Of the four events, that of 1932 stands out as having the lowest RMSE values across analogs, with it remaining below $1\text{ }^{\circ}\text{C}$ for many years. The other extreme is the 1948 event, with the predictions showing a consistently high RMSE that climbs above $1\text{ }^{\circ}\text{C}$ in about a year.

Figure 7 allows us to examine the RMSEs in more detail by showing the distribution of the RMSE over the 100 predictions for each of the 20 analogs for each event over the full duration of the event. The analogs are ranked along the x-axis from highest to lowest correlation with the observed time series. The standard deviation of the observed time series that each prediction was attempting to reproduce is also indicated. Note that the success of the predictions does not display any consistent relationship to the magnitude of the correlation coefficient between the analogs and the observed time series. While the RMSE values for the 1932 event are low, as in Fig. 6, the large spread shows that there are also several ensemble members that make unsuccessful predictions.

Figures 6 and 7 show that it is unrealistic to expect the perturbed ensembles to predict the exact trajectories of the cool periods over their full duration. However, the ensembles may predict the occurrence of a persistent cool state without following the exact same trajectory; i.e., the long-term statistics of the persistent cool states may be predictable despite the sensitivity of the predictions to initial conditions. In order to explore this possibility, we consider two criteria that characterize these persistent cool states:

1. Did the predicted Niño 3 time series have a negative mean anomaly for a similar duration to the length of the cool period in the observations (and analogs)?
2. Did the predicted Niño 3 time series remain below $2\text{ }^{\circ}\text{C}$ for a similar duration to the length of the cool period in the observations (and analogs)?

Figure 8 depicts the spread of the percentage of correct predictions over the twenty analogs corresponding to each of the events according to the first criterion. This is shown as

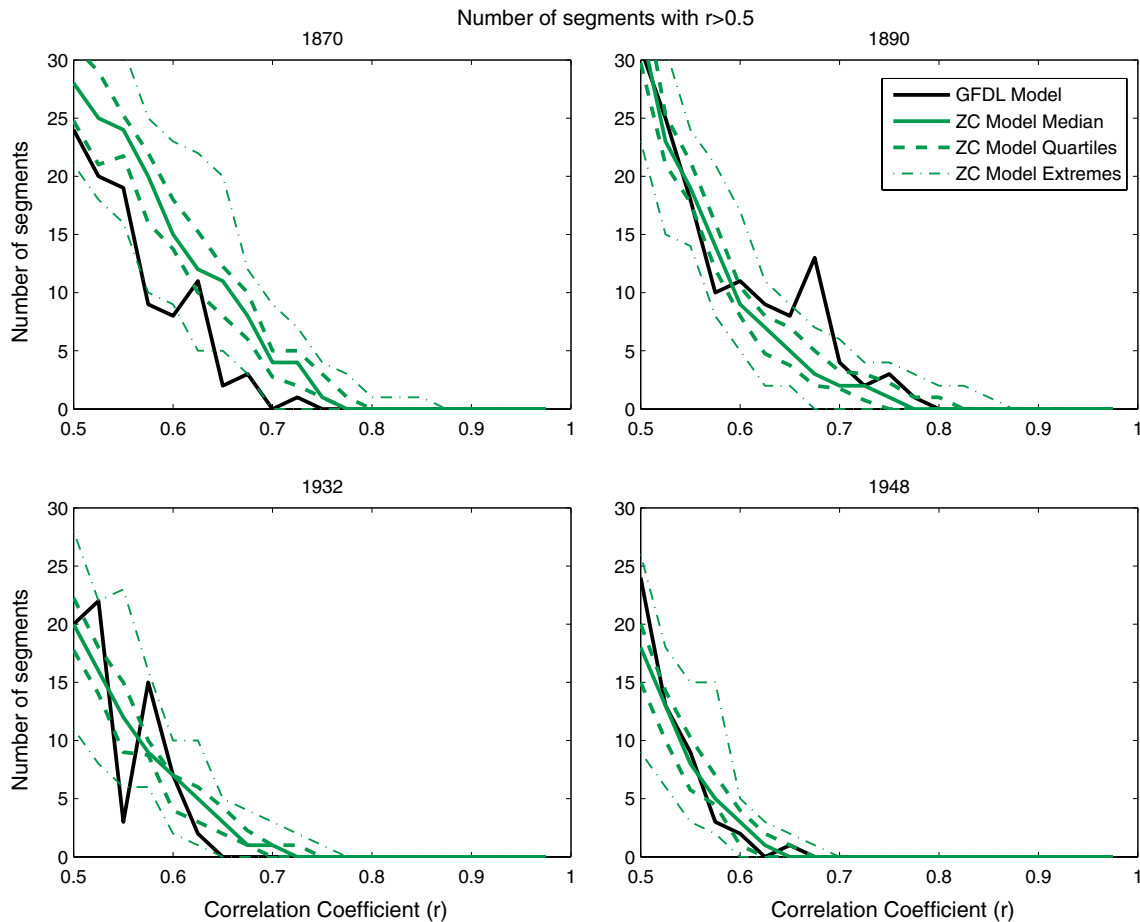


Fig. 5 The distribution of the segments having a higher correlation than 0.5 with the observed Niño 3 time series. The *black line* represents the GFDL model, while the *green lines* represent the median

(*solid line*), *upper and lower quartiles (heavy dashed lines)*, and *extreme values (light dash-dot lines)* over the 25 4000-year segments of the ZC model

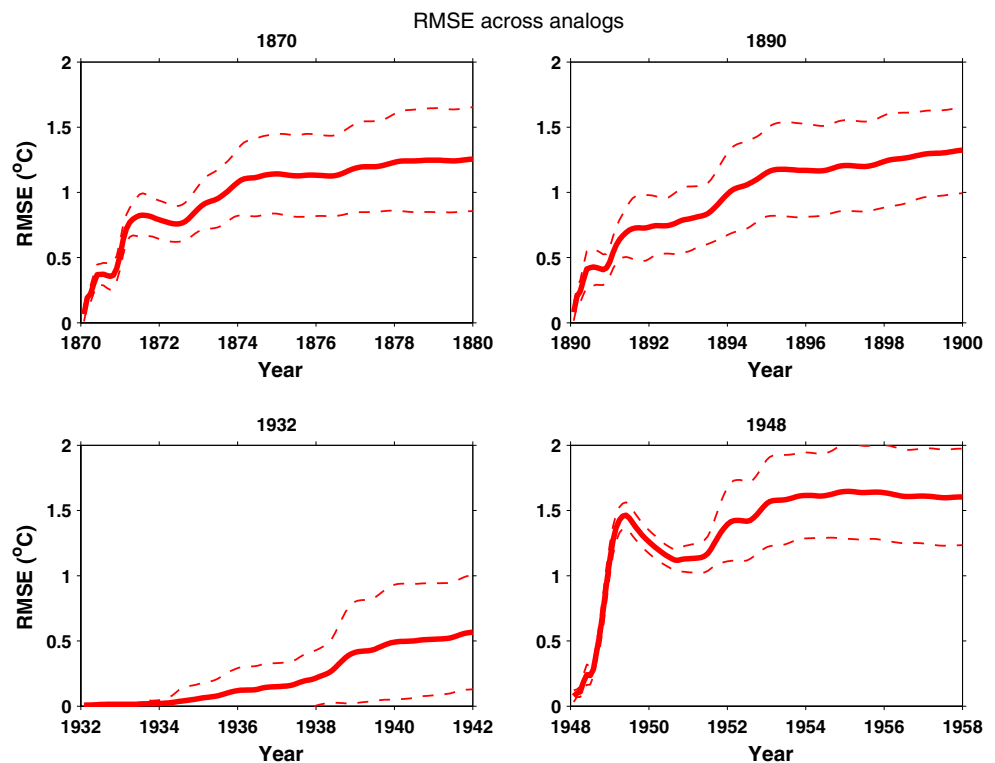
the evolution of the prediction of the long-term mean from 4 to 10 years after the initialization of the model. Evaluated in this way, the predictions fare better. For both the 1870 and 1890 events, the bulk of the analogs produce a correct prediction of the negative mean state, with only one or two outliers predicting a warm mean. For the 1932 event, the spread is so large as to fill the entire range of possibilities from 0 to 100%. The median prediction is correct near the timescale of the length of the event, however, indicating that the model does have some modest success even with this prediction. The incorrect prediction in earlier years shows that the model produces, in many instances, an El Niño event followed by a cooler period which allows the prediction to recover a cool mean state in later years. This means that even though the model is able to predict the analogs for this event quite well (as seen in the RMSE values from Figs. 6 and 7), the low similarity of the model analogs to this event in the observations (indicated by the correlation coefficients in Fig. 5) may be causing the cool mean state to be relatively poorly predicted. The model

largely fails to predict the 1948 event, with only a few outliers making a correct prediction. This is consistent with both the RMSE results in Fig. 7 which indicated that the 1948 event would be the least accurately predicted, and the correlation coefficient values from Fig. 5 which showed that the model has relatively few good analogs to this event.

The percentage over the entire ensemble (not separated according to analogs) that predict a negative mean is shown in Table 1. The values in this table confirm the successful predictions of 1870 and 1890, the modest success in later years in the case of 1932, and the poor performance for 1948. It is worth noting that the 4-year mean for the 1948 event is overwhelmingly predicted to be cool, indicating that the model is able to predict a La Niña event soon after initialization, but not its persistence in this case.

Turning to the second criterion, the model performs worse than for the first criterion in all cases except for the 1870 event. Figure 9 displays the results of the ensemble predictions evaluated based on whether the Niño 3 time series remains below the 2 °C threshold. The 1870 event is

Fig. 6 The root mean squared error averages over the entire ensemble of 2000 predictions for each of the events, with respect to the analog they were attempting to predict. The *bold line* depicts the mean while the *dashed lines* show one standard deviation above and below the mean



predicted very well and the 1890 event is only slightly less well-predicted in this case. In the case of the 1932 event, an El Niño event produced early in the predicted timeseries crosses the threshold, preventing most analogs from making a correct prediction according to this criterion. The prediction fails altogether in the 1948 case.

Table 2 lists the percentage of correct predictions according to the second criterion over the entire ensemble for each of the events. These values paint a similar picture to Fig. 9. The 1932 event shows the biggest difference between the two criteria: while its cool mean state is somewhat predictable, its lack of large El Niño events is not.

When both criteria are considered together, 68 and 64% of the full 2000-member ensemble correctly predict the 8-year 1870 event and the 7-year 1890 event. This number is lower than the corresponding values in both Tables 1 and 2, indicating that the two criteria do not always occur together, but is high enough that they can be considered successful predictions of both events. For the 1932 (8 years) and 1948 (10 years) events, the percentage that predict both criteria correctly are extremely low, at 13 and 14% respectively. In the 1932 case, it is clear from Tables 1 and 2 that the incorrect predictions of the second criterion are responsible for this low value, whereas in the case of 1948, it is simply a reflection of the poor performance of the ensemble forecast for this event.

In summary, the persistence of the cool mean state is predictable for 1870, 1890, and 1932, while the absence

of large El Niño events is not as predictable. Predictions of the cool mean state, however, have low confidence due to a large spread over the analogs; while this is partly a consequence of some analogs having higher predictability than others as seen in Fig. 7, it is also an indication that the analogs in the model are sensitive enough to initial conditions that predictions with high confidence may be difficult to achieve even with more sophisticated prediction schemes.

5 Forecasts of the decadal mean state in the ZC model

Given that the multi-year cool states have some predictability, and that the mean state of the Tropical Pacific shifts between warm and cool states on decadal timescales, we next perform a set of forecast experiments in order to determine whether it is possible to predict when the model is about to shift into a cooler or warmer decadal mean state. In Sect. 4, experiments were performed with the objective of testing whether a small perturbation to the initial state of the model was enough to prevent the system from taking a known trajectory—that of a persistent cool state. Since the analogs were selected by correlating with the cool periods, the entire model atmosphere–ocean state at the beginning of the analogs were effectively constrained by the information that the Niño 3 index of the analog would be in a cool mean state for a few years into the future. In the following

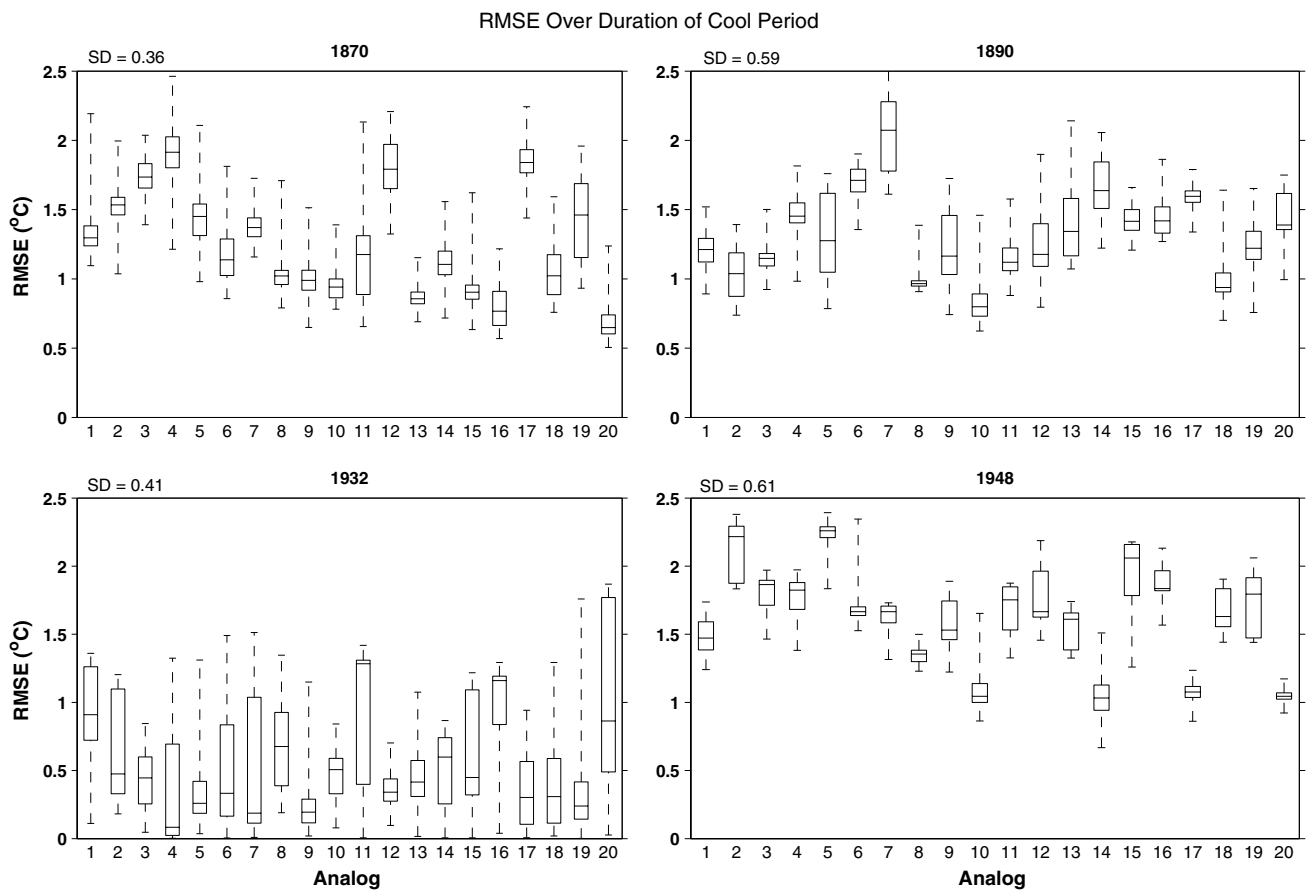


Fig. 7 The RMSE of the predictions with respect to the analog they were attempting to predict over the duration of the cool period. The *boxes* show the quartiles of the distribution, with the whiskers extending to the extreme values. The analogs are ranked from *left to right*

in descending order of their correlation coefficient with the observations. The standard deviation of the observed time series (SD) is indicated in each plot

experiments, we instead test whether it is possible to make forecasts of changes to the decadal mean state without including any information from the future, by choosing analogs based on correlations only with years prior to a shift in the mean state. In other words, these can be thought of as experiments performed in order to determine whether the model trajectory in the years leading up to the shift in the mean state contains enough information to predict the change in the decadal mean state.

We performed three hindcast experiments based on analogs to known decadal-scale shifts in the state of the Tropical Pacific: a shift from a warm to a cool state (“cool shift”) in 1943, a shift from a cool to a warm state (“warm shift”) in 1976, and a “neutral” period with no shift in the mean state centered on 1903. These three periods have been studied in earlier work (Karspeck et al. 2004) that examined their predictability. In each case, we selected the twenty 15-year segments of the ZC model output having the highest correlation coefficients with the 15 years prior to the year of the shift (but not including the shift itself) as

analog¹. For the 1943 cool (1976 warm) shift, the additional condition that the analog to the 15 years prior to the shift had a warm (cool) mean state was imposed; that is, a 15-year segment qualified as having a warm (cool) mean state if its mean was in the upper (lower) two quintiles of the distribution of all 15-year-means from the full 100,000-year run of the ZC model. The model was initiated in January of the 15th year of the segment, with perturbations added to the SST field using the same method as in Sect. 4, for an ensemble of 100 different simulations for each of the sets of 20 analogs, and run for 15 years.

We classified the results based on the quintiles from the distributions of 15-year means over the entire 100,000-year run of the ZC model as being correct, weakly correct, or

¹ The Karspeck et al. (2004) study, by contrast, used analogs that were based on a correlation with a 30-year period spanning the 15 years before and after the shift, thus including information from after the shift in the analog selection.

Fig. 8 The spread of the predictions over the 20 analogs for each case. On the y-axis is the percentage of cases (out of the 100 corresponding to each analog) that maintained a negative mean from the start date to the beginning of the year indicated on the x-axis. The colored solid lines show the median, the dashed lines show the quartiles, and the gray lines depict the individual analogs. The dotted red lines indicate the end of each cool period in the observations

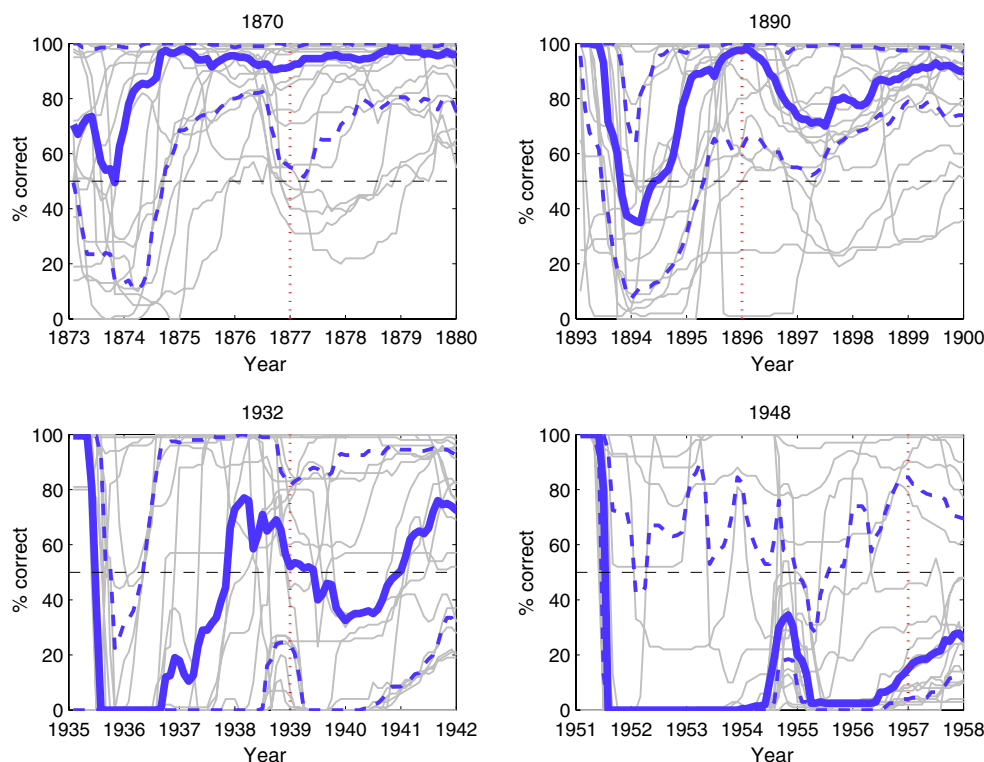


Table 1 Predictions of a negative mean over various periods from the full ensemble of 2000 experiments for each of the events, with the standard deviation of the prediction

Event	4 years	5 years	6 years	7 years	8 years	9 years	10 years
% With a negative mean							
1870	71 ± 2.0	57 ± 0.1	79 ± 0.7	87 ± 0.6	79 ± 0.4	82 ± 0.4	88 ± 0.4
1890	91 ± 2.1	42 ± 2.0	63 ± 0.9	81 ± 2.3	72 ± 2.2	76 ± 2.2	81 ± 0.7
1932	63 ± 0.9	20 ± 2.1	44 ± 2.1	55 ± 1.7	52 ± 1.7	45 ± 2.1	55 ± 1.5
1948	95 ± 1.6	26 ± 1.9	28 ± 2.1	30 ± 1.8	32 ± 1.8	30 ± 2.0	36 ± 2.1

The standard deviations are estimated using 1000 randomly-sampled subsets (with replacement) of 100 ensemble members each. The values in bold highlight the length of the cool period in the observations

wrong, based on the scheme used in Karspeck et al. (2004). For a warm (cool) shift, the prediction was correct if the mean of the predicted 15-year period was in the uppermost (lowermost) quintile. It was weakly correct if the predicted 15-year mean was in the second-highest (-lowest) quintile, and wrong otherwise. For the neutral state, the results were correct if the predicted 15-year mean fell within the middle quintile, wrong if it fell in either of the extreme quintiles, and weakly correct otherwise. The results of this classification are shown in Table 3.

Given that the probability of falling within each of the quintiles is 20% by pure chance, the distribution by chance between correct, weakly correct, and wrong is 20, 20 and 60% for the warm and cool shifts, and 20, 40 and 40% for the neutral shift. The neutral state is the best-predicted of

the three, with 32% of the predictions falling within the correct category, compared to 20% expected by chance. The warm shift hindcasts also perform better than chance, with 29% of the predictions being correct. However, the cool shifts are predicted worse than by pure chance with 73% of the predictions falling into the wrong category. This implies that the model state at the time of a cool shift is extremely sensitive to initial conditions, leading to predictions of a warm or neutral shift instead when perturbations are added. These results are consistent with those of Karspeck et al. (2004), who found shifts to a warm state to have higher predictability than shifts in the opposite direction in this model.

We examine these hindcasts more closely in Fig. 10, which shows the temporal evolution of the spread of the

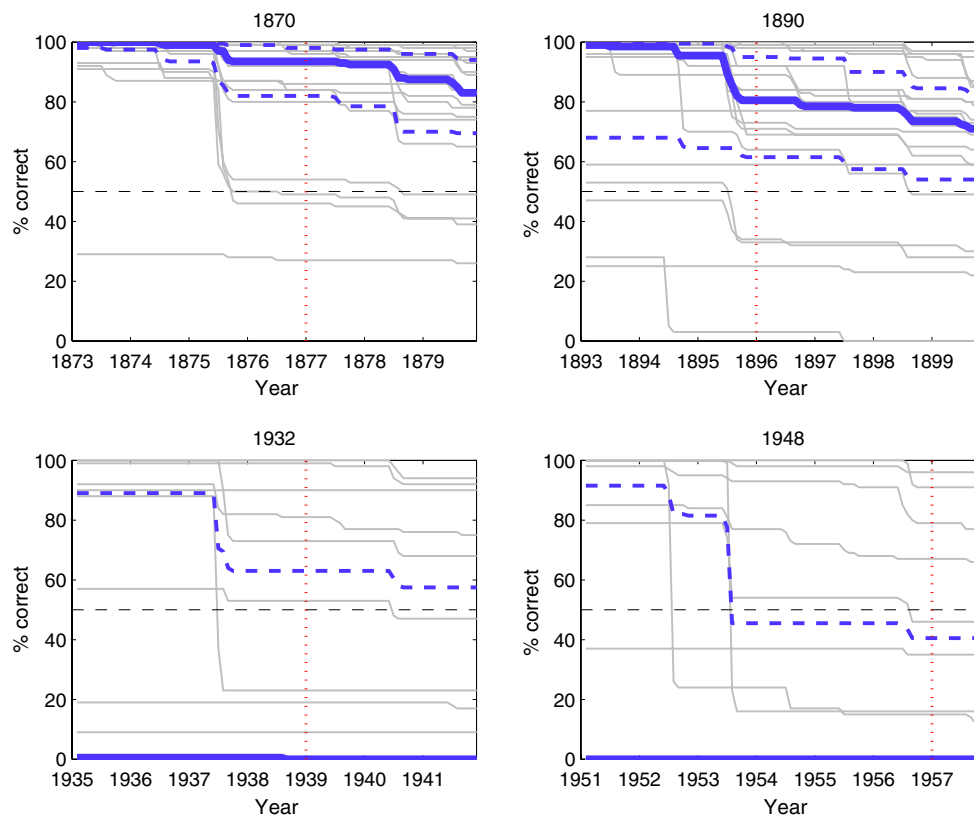


Fig. 9 As in Fig. 8, evaluating predictions based on whether the predicted Niño 3 time series remained below 2 °C (correct) or not (incorrect)

Table 2 As in Table 1, for predictions of the Niño3 index remaining below 2 °C

Event	4 years	5 years	6 years	7 years	8 years	9 years	10 years
% Below 2 °C							
1870	95 ± 2.3	95 ± 2.3	93 ± 2.3	84 ± 2.3	84 ± 2.3	83 ± 2.2	79 ± 2.2
1890	84 ± 0.9	83 ± 1.2	80 ± 1.2	72 ± 2.1	71 ± 1.9	69 ± 1.9	66 ± 1.9
1932	33 ± 0.2	33 ± 0.2	33 ± 0.2	27 ± 0.2	27 ± 0.2	27 ± 0.2	26 ± 0.2
1948	35 ± 1.6	35 ± 1.6	31 ± 1.6	25 ± 1.5	24 ± 1.5	24 ± 1.5	22 ± 1.4

The values in bold highlight the length of the cool period in the observations

Table 3 Classification of the hindcast predictions into correct, weakly correct, and wrong categories based on the quintiles of the distribution of 15-year means over the entire 100,000-year model simulation

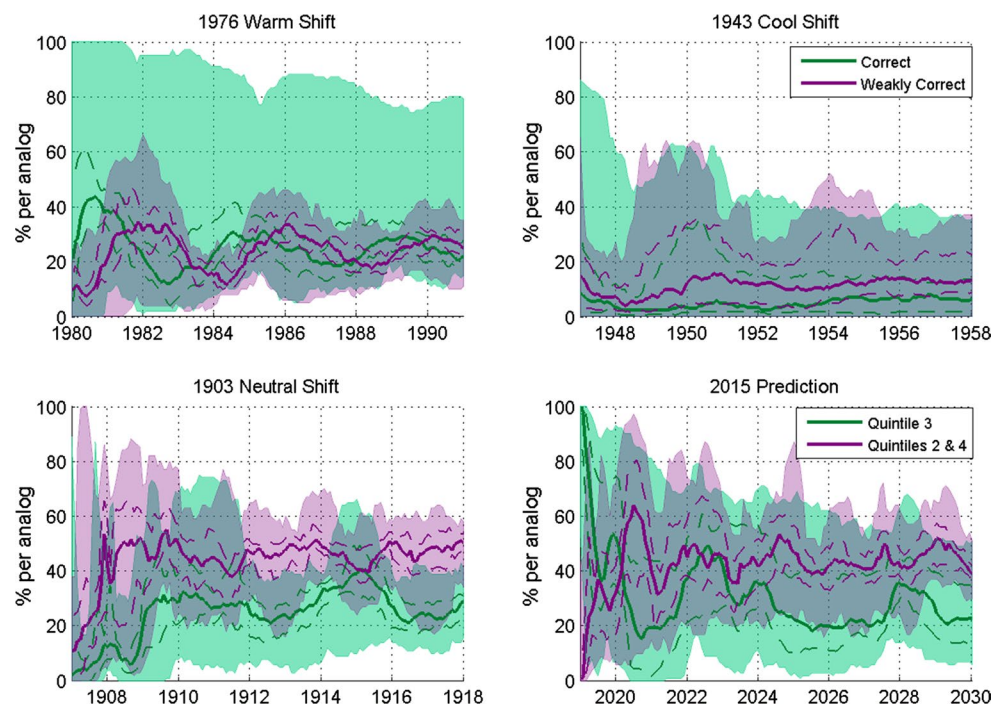
Year of the shift	Correct (%)	Weakly correct (%)	Wrong (%)
1976 (warm)	29.10 ± 2.22	24.70 ± 1.37	46.20 ± 1.84
1943 (cool)	9.85 ± 0.67	16.60 ± 0.84	73.55 ± 1.08
1903 (neutral)	32.45 ± 1.88	46.80 ± 1.73	20.75 ± 0.81

The standard deviations indicated were estimated as in Tables 1 and 2

number of correct and weakly correct predictions across the 20 analogs. Here, we use quintiles as described above for the length of time from the beginning of the prediction

ranging from 4 to 15 years (e.g., the ensemble spread 5 years into the prediction was classified using the quintiles of the distribution of 5-year means from the long run). In the case of warm and cool shifts, the percentage of correct and weakly correct predictions depicted should be compared against the “pure chance” prediction of 20%. Based on this metric, a correct prediction of the warm mean is made consistently after 1984 (i.e., 8 years into the prediction) as both the median percentage of correct predictions and weakly correct predictions (bold green and purple lines, respectively) each remain above the 20% line. The range of predictions also shows that some of the analogs make a correct prediction over 70% of the time. By contrast, the cool shift is poorly predicted with both correct and weakly correct predictions remaining below 20% for

Fig. 10 The spread over analogs for predictions of the decadal mean classified as correct (*green*) and weakly correct (*purple*) according to the quintile scheme described in the text. *Solid lines* indicate the median, *dashed lines* indicate the *upper* and *lower* quartiles, and the translucent *shading* fills in the range between extremes. Note that the prediction by pure chance is 20% in all cases except for the weakly correct classification in the *lower two panels*, where it is 40%



the entire prediction. However, the spread over the analogs shows that there are some analogs that make a correct prediction, with the upper quartile of the weakly correct predictions being above 20%.

The third panel shows the same for the neutral shift. This is the best-predicted of all the cases with the median correct and weakly correct predictions remaining above 20 and 40% respectively. Note that in this case, the weakly correct prediction should be compared to the pure chance prediction of 40%. In all cases, the envelope of the predictions narrows over time, suggesting that there is more confidence in predictions made over a 15-year timescale than over a shorter timescale of 8–10 years. This narrowing of the range of predictions is distinct from the convergence of the long-term mean to a single value, as the classification uses quintiles that correspond to the length of time from the beginning of the prediction.

Finally, a fourth forecast experiment was run using the same method with analogs to the 1999–2014 period in order to predict whether there will be a shift from this cool state to a warmer state in the near future. This forecast was initialized in August 2015, as the major El Niño event of 2015–2016 was developing. The results are shown in the bottom-right panel of Fig. 10.

This ensemble is most heavily weighted towards a neutral mean in the future, with 34.5% of the predictions falling within the central quintile and 41.2% of the predictions in the second and fourth quintiles combined. This indicates

a high-confidence prediction of a shift to a neutral mean state. Given that the analogs used to make this prediction were originally in a cool state, with 17 out of the 20 in the lowest quintile, it is clear that this implies a shift to a warmer state from that of the analogs. Thus, this prediction can be interpreted as a warm shift, but with the resulting state maintaining a neutral mean.

Figure 11 displays the spread of the annual means predicted by the full ensemble, and confirms that the model prediction is of a positive mean, albeit of low magnitude. The entire ensemble correctly predicts an El Niño event in 2015–2016, with most ensemble members predicting a La Niña event immediately after, as is currently anticipated (International Research Institute for Climate and Society, 2016, available at <http://iri.columbia.edu/our-expertise/climate/forecasts/enso/current>). The model predicts that this La Niña event will last for 2 years, after which the spread among ensemble members grows larger. The thick red line shows the evolution of the median prediction of the mean value of Niño 3 with the averaging beginning at the start of the prediction. While there is significant variability over the years following 2017, the mean remains warm. The influence of the predicted El Niño event of 2015–2016 on the 15-year mean is not solely responsible for the warm state: while 72% of the ensemble members predict a warm 15-year mean, if the averaging is begun a year later to exclude the initial El Niño event, 58% of them still make this prediction.

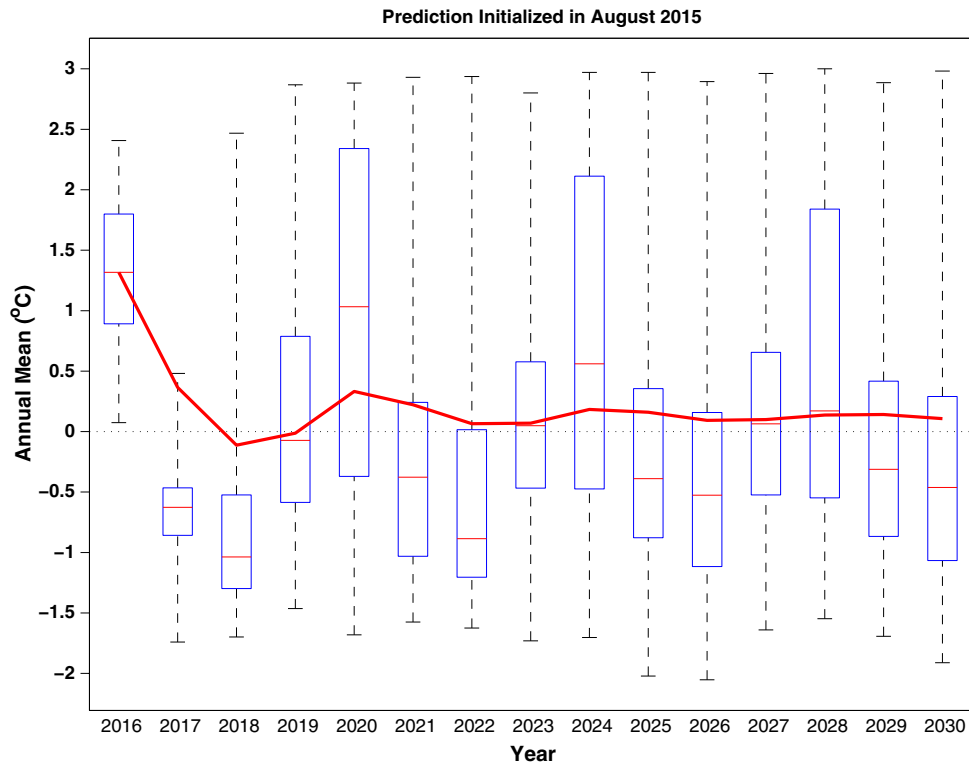


Fig. 11 The annual means predicted by the entire 2000-member ensemble, irrespective of analog. The *boxes* show the median, *upper* and *lower* quartiles; the *whiskers* extend to the extreme values. The *thick red line* displays the median among the ensemble members of

the long-term mean, where the averaging begins at the start of the prediction and ends in the year indicated on the x-axis. Note that the long-term model mean of $0.31\text{ }^{\circ}\text{C}$ has been subtracted from all values shown in this figure

6 Discussion and conclusion

Both the ZC and GFDL models are able to produce the type of quiescent, cool behavior of the Tropical Pacific that forces large-scale, prolonged droughts. The presence of these analogs in the models has implications for hypotheses regarding the origins of these cool periods in reality. Since both of the models used were unforced for the entire period, we can infer that the cool periods observed in nature were not necessarily a product of forcing by external factors such as aerosols or solar activity, allowing for the possibility that they are a feature of the natural, intrinsic variability of the climate system. The fact that the ZC model, which only simulates the tropical Pacific, is able to capture such behavior with as much skill as the GFDL model demonstrates that the addition of the Atlantic Ocean or the extratropics need not improve the simulation of these states, meaning that mechanisms originating in these parts of the world (Dong et al. 2006; Kang et al. 2014; Barnett et al. 1999; Vimont et al. 2003; Gu and Philander 1997) may not be necessary for their development. This points to the tropical Pacific itself being able to produce these cool periods independently, without requiring the influence of any other part of the globe.

The ZC model had great difficulty running for more than a few years without producing an El Niño event. However, the model is able to predict the negative mean temperature anomaly over multi-year periods; this extends to the multi-year timescale the limited predictability that Karspeck et al. (2004) found for decadal timescales. The existence of some predictability on these timescales also implies that long-term variability is not driven entirely by stochastic processes (Flugel and Chang 1999) in this model.

The model is able to make predictions of the cool mean state in all cases except for the 1948 event. This may indicate that the cool states in the model have more in common with those periods that take place immediately after large El Niño events (such as 1870, 1890, 1932, and 1999–2014) than the periods that are continuously cool, such as the 1950s. While the 1870 event has the highest correlations with the ZC model (up to 0.9), the distribution of the correlations with the best analog for the 1890 event has the highest number of segments with a correlation higher than 0.5. Therefore, it is likely that the predictability of the analogs of these two events arises for different reasons: the 1870 event is the one that most resembles the model's behavior, and the 1890 event, while not quite as well-correlated with its analogs, has analogs that are representative of some

recurring state of the model. This suggests that the model's own version of the cool periods may be quite predictable using more elaborate methods than those in this study.

Although the procedure for selecting analogs could easily be improved, the lack of improvement in forecast skill with an increase in the correlation coefficient of the analogs (seen in Fig. 7) suggests that using higher-quality analogs based on correlation is unlikely to significantly enhance the forecast skill. It is therefore likely that information other than the Niño 3 index will be required to improve predictions. More sophisticated constraints than a simple correlation and restriction of the long-term mean value could be applied to the process of finding analogs in the Niño 3 time series; but more importantly, other information from the models could also be included in the analog selection process, such as thermocline depth, wind stress patterns, or spatial information regarding the SST field.

A more detailed characterization of the states produced by the model and identifying what features the cool periods have in common will also allow us to investigate the mechanisms that could drive the persistence of these La Niña-like states, understand the processes by which they develop, and determine the features that give rise to their predictability.

For decadal timescale shifts, given the limitations of the method and simplicity of the model used, it is quite remarkable that the shift to a warm mean state in 1976 could be predicted. This simple method used implies that a significant portion of the information required in order to make a forecast of a warm shift is contained within the Niño 3 index itself. The continuation of a neutral mean state is similarly predictable, also proving that the model does not always make a prediction of a warm shift.

The ensemble initialized in August 2015 produces a fairly high-confidence prediction of a moderate shift to a warmer state. A forecast using the same method in 2004 made the correct prediction of the continuation of the cool state until 2013 (Seager et al. 2004), showing that while the model is unable to make a skillful forecast of a shift to a cool state, it has the ability to predict a continuing cool state. This lends further confidence to the forecast of a shift now to a warmer state. The predicted annual means show significant interannual variability over the 15-year period, similar to warm periods that have been observed in the past. As this forecast was made using the unforced ZC model, it relies on the current cool mean state being a product of the internal variability of the tropical Pacific system, and not the result of anthropogenic forcing or variability in other parts of the world (which is not necessarily the case (Cane et al. 1997)). That the unforced ZC model is able to produce analogs to the current state of the system implies that it could be a product of natural variability alone.

The demise of the cool mean state in the tropical Pacific would have significant implications. One of these is the easing of the drier-than-normal conditions currently affecting the southwestern United States and other parts of the subtropics (Delworth et al. 2015), even as human-driven hydroclimate change advances (Seager et al. 2007). Another possible consequence is the end of the global warming “hiatus”, with surface temperature rise accelerating to a rate similar to that prior to 1998.

Acknowledgements N.R. is supported by the National Aeronautics and Space Administration (NASA) Headquarters under the NASA Earth and Space Science Fellowship Program—Grant NNX-14AK96H. R.S. is supported by National Oceanic and Atmospheric Administration award NA14OAR4310232. M.A.C. and D.E.L. are supported by the Office of Naval Research under the research grant MURI (N00014-12-1-0911). We would like to thank Stephen E. Zebiak for his assistance with setting up the Zebiak–Cane model.

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

- Adams J, Mann M, Ammann C (2003) Proxy evidence for an El Niño-like response to volcanic forcing. *Nature* 426:274–278
- Barnett TP et al (1999) Interdecadal interactions between the tropics and midlatitudes in the Pacific basin. *Geophys Res Lett* 26(5):615–618
- Cane MA, Zebiak SE, Dolan SC (1986) Experimental forecasts of El Niño. *Nature* 321(6073):827–832
- Cane M, Zebiak S, Xue Y (1995) Model studies of the long term behavior of ENSO. In: Martinson DG, Bryan K, Ghil M, Hall MM, Karl TR, Sarachik ES, Sorooshian S, Talley LD (eds) *Natural climate variability on decade-to-century time scales*. National Academy Press, Washington, DEC-CEN Workshop, Irvine, pp 442–457
- Cane MA et al (1997) Twentieth-century sea surface temperature trends. *Science* 275:957–960
- Chen X, Wallace JM (2015) ENSO-like variability: 1900–2013. *J Clim* 28:9623–9642
- Chen D et al (2004) Predictability of El Niño over the past 148 years. *Nature* 428:733–736
- Delworth TL et al (2006) GFDL's CM2 global coupled climate models: part I—formulation and simulation characteristics. *J Clim* 19(5):643–674
- Delworth TL et al (2015) A link between the hiatus in global warming and North American drought. *J Clim* 28:3834–3845
- Deser C, Alexander MA, Timlin MS (1996) Upper-ocean thermal variations in the north Pacific during 1970–1991. *J Clim* 9:1840–1855
- Dong B, Sutton R, Scaife A (2006) Multidecadal modulation of El Niño Southern oscillation variance by Atlantic Ocean Sea surface temperatures. *Geophys Res Lett* 33:L08705

- Easterling DR, Wehner MF (2009) Is the climate warming or cooling? *Geophys Res Lett* 36:L0806
- Emile-Geay J, Cane MA (2009) Pacific decadal variability in the view of linear equatorial wave theory. *J Phys Oceanogr* 39:203–219
- Flugel M, Chang P (1999) Stochastically induced climate shift of El Niño-Southern oscillation. *Geophys Res Lett* 26(16):2473–2476
- Graham NE (1994) Decadal-scale climate variability in the tropical and North Pacific during the 1970s and 1980s: observations and model results. *Clim Dyn* 10(3):135–162
- Gu D, Philander SGH (1997) Interdecadal climate fluctuations that depend on exchanges between the tropics and the extratropics. *Science* 235:805–807
- Herweijer C, Seager R (2008) The global footprint of persistent extra-tropical drought in the instrumental era. *Int J Climatol* 28:1761–1774
- Herweijer C, Seager R, Cook ER (2006) North American droughts of the mid to late nineteenth century: a history, simulation and implication for Mediaeval drought. *Holocene* 16(2):159–171
- Hoerling M, Kumar A (2003) The perfect ocean for drought. *Science* 299(5607):691–694
- International Research Institute for Climate and Society (2016) IRI ENSO Forecast. <http://iri.columbia.edu/our-expertise/climate/forecasts/enso/current>. Accessed 6 June 2016
- Kang I, No H-H, Kucharski F (2014) ENSO amplitude modulation associated with the mean SST changes in the tropical central Pacific induced by Atlantic multi-decadal oscillation. *J Clim* 27:7911–7920
- Kaplan A et al (1998) Analyses of global sea surface temperature 1856–1991. *J Geophys Res* 103(C9):18567–18589
- Karspeck A, Seager R, Cane M (2004) Predictability of tropical Pacific decadal variability in an intermediate model. *J Clim* 17:2842–2850
- Kosaka Y, Xie S-P (2013) Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature* 501:403–407
- Kug JS et al (2010) Warm pool and cold tongue El Niño events as simulated by the GFDL CM2.1 coupled GCM. *J Clim* 23:1226–1239
- Lyon B (2014) Seasonal drought in the greater horn of Africa and its recent increase during the March–May long rains. *J Clim* 27:7953–7975
- Lyon B, DeWitt D (2012) A recent and abrupt shift in the East African long rains. *Geophys Res Lett* 39(2):L02702
- Mantua NJ et al (1997) A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull Am Meteorol Soc* 23:1069–1079
- McGregor S, Holbrook NJ, Power SB (2007) Interdecadal sea surface temperature variability in the equatorial Pacific Ocean: part I—the role of off-equatorial wind stresses and oceanic Rossby waves. *J Clim* 20:2643–2658
- Meehl GA et al (2011) Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nat Clim Change* 1:360–364
- Newman MCGP, Alexander MA (2003) ENSO-forced variability of the Pacific decadal oscillation. *J Clim* 16(23):3853–3857
- Newman M et al (2016) The Pacific decadal oscillation, revisited. *J Clim* 29:4399–4427
- Okumura Y, Deser C (2010) Asymmetry in the duration of El Niño and La Niña. *J Clim* 23:5826–5843
- Schneider N et al (1999) Subduction of decadal North Pacific temperature anomalies: observations and dynamics. *J Phys Oceanogr* 23:1056–1070
- Seager R (2007) The turn of the century North American drought: global context, dynamics, and past analogs. *J Clim* 20:5527–5553
- Seager R (2015) Decadal hydroclimate variability across the Americas. In: Chang C-P, Ghil M, Latif M, Wallace JM (eds) *Climate change: multidecadal and beyond*. World Scientific Publishing, Singapore
- Seager R et al (2004) Predicting Pacific decadal variability. In: Wang C, Xie SP, Carton JA (eds) *Earth climate: the ocean-atmosphere interaction*. American Geophysical Union, Washington, pp 105–120
- Seager R et al (2005) Modeling of tropical forcing of persistent droughts and pluvials over Western North America: 1856–2000*. *J Clim* 18:4065–4088
- Seager R et al (2007) Model Projections of an imminent transition to a more arid climate in Southwestern North America. *Science* 316(5828):1181–1184
- Takahashi K, Montecinos A, Goubanova K, DeWitte B (2011) ENSO Regimes: reinterpreting the canonical and Modoki El Niño. *Geophys Res Lett* 38:L10704
- Timmermann A, Jin F-F (2002) A nonlinear mechanism for decadal El Niño amplitude changes. *Geophys Res Lett* 29(1):3.1–3.4
- Tziperman E, Stone L, Cane MA, Jarosh H (1994) El Niño Chaos: overlapping of resonances between the seasonal cycle and the Pacific ocean–atmosphere oscillator. *Science* 264:72–74
- Vimont DJ, Battisti DS, Hirst AC (2003) The seasonal footprinting mechanism in the CSIRO general circulation models. *J Clim* 16:2653–2667
- Vose R et al (1992) *The global historical climatology network: long-term monthly temperature, precipitation, and pressure data*. American Meteorological Society, Anaheim
- Wittenberg AT et al (2014) ENSO modulation: is it decadal predictability? *J Clim* 27:2667–2681
- Yang W, Seager R, Cane MA, Lyon B (2014) The East African long rains in observations and models. *J Clim* 27:7185–7202
- Zebiak S (1982) A simple atmospheric model of relevance to El Niño. *J Clim* 39:2017–2027
- Zebiak SE, Cane MA (1987) A model El Niño Southern oscillation. *Mon Weather Rev* 97(3):2262–2278
- Zhang Y, Wallace JM, Battisti DS (1997) ENSO-like interdecadal variability: 1900–1993. *J Clim* 10:1004–1020