A regime view of ENSO flavours through clustering in CMIP6 models

Pradeebane Vaittinada Ayar[1,2], David S. Battisti[3], Camille Li[4], Martin King[4], Mathieu Vrac[2], and Jerry Tjiputra[1]

[1]NORCE Norwegian Research Centre AS, Bjerknes Centre for Climate Research, Bergen, Norway
[2]Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL), CEA/CNRS/UVSQ, Université Paris-Saclay, Centre d'Etudes de Saclay, Orme des Merisiers, 91191 Gif-sur-Yvette, France
[3]Department of Atmospheric Sciences, University of Washington, Seattle, WA
[4]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway

Key Points:

- A clustering approach identifies two observed warm and cold ENSO regimes to which simulated regimes are matched and consistently evaluated.
- Over historical period, CMIP6 models well simulate ENSO patterns with discrepancies in terms of frequency, seasonality and persistence.
- Future evolution in terms of frequency, magnitude and variability depends on type of cold or warm ENSO regime.

Corresponding author: Pradeebane Vaittinada Ayar, `pradeebane@laposte.net`; `pradeebane@lsce.ipsl.fr`

Abstract

El Niño Southern-Oscillation (ENSO) flavours in the tropical Pacific are studied from a regime perspective. Five recurring spatial patterns or *regimes* characterising the diversity of ENSO are established using a clustering approach applied to the HadISST sea surface temperature (SST) anomalies. Compared to previous studies, our approach gives a monthly characterisation of the diversity of the warm and cold phases of ENSO established from observations but commonly applied to models and observations. Two warm (eastern and central El Niño), two cold (basin wide and central La Niña) and a neutral reference regimes are found. Simulated SST anomalies by the models from the latest Coupled Model Intercomparison Project (CMIP6) are then matched to these reference regimes. This allows for a consistent assessment of the skill of the models in reproducing the reference regimes over the historical period and the change in these regimes under the high-warming Shared Socio-economic Pathway (SSP5.8.5) scenario. Results over the historical period show that models simulate well the reference regimes with some discrepancies. Models simulate more intense and spatially extended ENSO patterns and have issues in capturing the correct regime seasonality, persistence and transition between regimes. Some models also have difficulty simulating the frequency of regimes, the eastern El Niño regime in particular. In the future, both El Niño and central La Niña regimes are expected to be more frequent accompanied with a less frequent neutral regime. The central Pacific El Niño and La Niña regimes are projected to increase in amplitude and variability.

Plain Language Summary

A heuristic definition to characterise the diversity of sea surface temperature spatial patterns or *regimes*, typical of the El Niño-Southern Oscillation (ENSO) and common to observation and climate model simulations, is established here. Using this approach, we found from the observations two warm (eastern and central El Niño), two cold (basin wide and central La Niña) and a neutral reference ENSO regimes. State-of-the-art climate models are generally able to simulate the spatial patterns of these observed five ENSO regimes to some extent. Models simulate overly intense and spatially extended ENSO patterns and have issues in capturing the correct regime seasonality, persistence and transition between regimes. Under the business as usual future scenario, the model projections indicate that eastern and central El Niño and central La Niña regimes are expected to be more frequent accompanied with a less frequent neutral regime. The central Pacific El Niño and La Niña regimes are projected to increase in amplitude and variability.

1 Introduction

El Niño-Southern Oscillation (ENSO) is the leading mode of interannual climate variability (see, e.g., Rasmusson & Carpenter, 1982; Zhang et al., 1997; X. Chen & Wallace, 2015, and references therein). ENSO is a true mode of the coupled atmosphere-ocean system in the tropical Pacific (see Zebiak and Cane (1987) and the review papers by Neelin et al. (1998) and Battisti et al. (2019) and references therein): without the Southern Oscillation variability, there would be no El Niño or La Niña events, and vice versa. Owing to the slow decay rate of the ENSO mode, the state of ENSO is predictable up to a year in advance.

ENSO causes seasonal temperature and precipitation anomalies on a global scale by way of oceanic and atmospheric teleconnections associated with, respectively, changes in the wind stress acting on the ocean and changes in the location of precipitation in the tropical Pacific, (Trenberth et al., 1998; Davey et al., 2014; X. Chen & Wallace, 2015). As such, ENSO has nearly global impacts on agriculture (e.g., Phillips et al., 1998; Naylor et al., 2001; Iizumi et al., 2014), fisheries (e.g., Bertrand, 2020) and water resources (e.g., Hamlet & Lettenmaier, 1999; Poveda et al., 2001; Nicholas & Battisti, 2008). However, the impact of ENSO on the climate beyond the tropical Pacific depends greatly on subtle differences in patterns of sea surface temperature anomalies associated with each ENSO warm and cold event – the so-called different "flavours" of ENSO (K. Takahashi et al., 2011; Thomas et al., 2018; Vimont et al., 2022) – that are a result of the stochastic nature of the atmospheric forcing that provides the energy for ENSO (Vimont et al., 2003). ENSO also alters the global carbon cycle by dominating the year-to-year variability in global atmospheric carbon concentrations (P. J. Rayner et al., 1999). Roughly, land regions emit more CO_2 during El Niño and less CO_2 during La Niña (Betts et al., 2020). In the ocean, ENSO mostly affects the CO_2 fluxes in the tropical Pacific, which is the largest carbon outgassing system to the atmosphere, but with an anomaly signal that is the opposite of the land (Feely et al., 2006; T. Takahashi et al., 2009; Vaittinada Ayar et al., 2022).

ENSO events are diverse in terms of the magnitude, duration, and location of sea surface temperature (SST) anomalies (Capotondi et al., 2020). Among the well-known flavours of ENSO are warm (El Niño) events that tend to feature maximum warm anomalies in the far eastern equatorial Pacific and those that tend to have maximum amplitude in the central equatorial Pacific, and cold (La Niña) events that mostly have maximum amplitude in the central equatorial Pacific. That warm events can be more extreme than cold events stems from the non-linear relationship between thermocline displacements and SST anomalies in the eastern Pacific (Battisti et al., 2019).

In order to better consider ENSO diversity, K. Takahashi et al. (2011) introduced an approach that differentiates between central and eastern Pacific warm anomaly patterns in observations or models. It is based on the nonlinear relationship between the two leading empirical orthogonal functions (EOF) of tropical Pacific SST anomalies. K. Takahashi et al. rotated the first and the second principal component (PC1 and PC2) axes by 45° to introduce two indices E and C defined as: $E = \frac{PC1-PC2}{\sqrt{2}}$ and $C = \frac{PC1+PC2}{\sqrt{2}}$. They then showed that E and C represent, respectively, eastern and central Pacific warm events. E and C indices have been extensively used to study warm events in observations and in different generations of numerical climate models (see, Dommenget et al. (2013); K. Takahashi et al. (2011) for the Coupled Model Intercomparison Project Phase 3, CMIP3, Cai et al. (2018); Karamperidou et al. (2017) for CMIP5 and Fredriksen et al. (2020) for CMIP6). This approach allows a better characterisation of warm event diversity (Dommenget et al., 2013) and distinguishes climate models according to their ability to simulate this PC1/PC2 non-linearity (Dommenget et al., 2013; Cai et al., 2018). However, the SST patterns associated with EOF1 and EOF2 (from which PC1 and PC2 are derived to calculate E and C indices) can differ greatly between observations and models and between models (Cai et al., 2018). Indeed, the two model-specific leading EOFs of any given model do not necessarily capture the same SST variability as in observations, making comparisons difficult. Therefore, in order to consistently evaluate the diversity and asymmetry of ENSO events representation across models and observations, a reference framework that provides a common definition of ENSO events based on spatial SST anomaly patterns has to be established.

One approach to characterise modes of variability (for different climate variables) is through regime analysis which picks out recurrent spatio-temporal structures or regimes (for instance, seasonal North Atlantic atmospheric circulation or rainfall patterns associated with the North Atlantic oscillation) , in observations (Vautard, 1990; Yiou & Nogaj, 2004; Cassou, 2008; Vrac & Yiou, 2010; Vrac et al., 2014; Hertig & Jacobeit, 2014) and in climate models (Sanchez-Gomez et al., 2009; Fabiano et al., 2021; Breton et al., 2022). In this paper, a statistical regime analysis of SST anomalies over the tropical Pacific is performed to identify recurring spatial patterns (or regimes) typical of ENSO. To our knowledge, very few studies have used clustering approaches to analyse ENSO-associated

SST anomaly patterns in observations and none are applied to the models. Based on observations, Johnson (2013) used self organised maps to define nine ENSO patterns from November to February averaged SST anomaly and Su et al. (2018) defined 13 patterns using a K-means approach applied to zonally averaged SST anomalies. In Johnson (2013), the clustering was applied to a few dozen seasonally averaged maps which does not allow a description of the dynamics of ENSO. In Su et al. (2018), the number of regimes was increased until it was large enough to describe ENSO dynamics solely based on its spatial pattern. The objective of the present study is to provide a common definition of ENSO flavours based on observations that enables us to robustly study the dynamics and the variability of these flavours in observations and in CMIP6 models by characterising continuous monthly ENSO evolution rather than different types of warm (or cold) events. The novelty of this study is to provide such a definition based on clustering using a Gaussian mixture model (Pearson, 1894) which is a more flexible generalisation of k-means clustering that provides a data-driven method for identifying the appropriate number of regimes. From such defined ENSO regimes, various properties of each regime are examined, such as their frequency of occurrence, persistence, seasonal distribution and their regime transitions in both observations and in CMIP6 models (Eyring et al., 2016) over the historical period (1920-2014). The changes in the regimes under high-warming scenario in terms of occurrence, intensity and variability are also evaluated.

The paper is structured as follows. Section 2.1 details the datasets and pre-processing requirements for the analysis. Section 2.2 explains the methodology. The results regarding the reference observation-based ENSO regimes are presented in Section 3.1. Section 3.2 and 3.3 respectively describe the ability of the models to reproduce reference regimes and their future changes. Some discussions and conclusions are provided in Section 4 and 5.

2 Data and Methods

2.1 Data and Preprocessing

The analysis is conducted on monthly sea surface temperature (SST) extracted from the Met Office Hadley Centre HadISST observation-based gridded analyses from 1870 to 2021 (N. A. Rayner et al., 2003) at $1°\times1°$ spatial horizontal resolution and from an ensemble of 16 Earth system model (ESM) simulations from the Coupled Model Intercomparison Project 6 (CMIP6, Eyring et al., 2016, see Table 1). In this study, HadISST is considered as the reference observational data-set used to define reference ENSO regimes for evaluating the simulations. All simulations are regridded onto a regular $1°\times1°$ grid using bilinear interpolation provided by climate data operators (CDOs). In this study, analyses are conducted over the HadISST reference period 1920-2014. The starting year is set to 1920 due to observational data (ship records) in the equatorial East Pacific being very sparse before the 1920s which can impact ENSO variance (*i.e.,* Solomon & Newman, 2012). ENSO regimes simulated by the ESMs for the period 1850-2100 is examined, combining model output from the Historical simulations from 1850 to 2014 (which corresponds to the end of the reference period) high CO_2 Shared Socio-economic Pathway scenario (SSP5-8.5), from 2015 to 2100 (O'Neill et al., 2016).

ENSO regimes are usually defined using SST anomalies over the tropical Pacific. Our study is conducted on the anomalies over the Pacific domain between 20°S-20°N and from 140°E to the west coast of the Americas from the regridded data (see Panel 3 of Figure 2a) for the exact study area).

Monthly SST anomalies at each grid-point are computed by separately removing the trend of each calendar month time-series using a cubic smoothing spline (implemented by the function smooth.spline in R software; R Core Team, 2020) over the period 1870-2021 for HadISST and 1850-2100 for the model simulations. For instance, the non-linear

Table 1. List of the 16 CMIP6 models used in this study with the horizontal resolution of the ocean component, variant label, model and data references. Note that most of the models have irregular grids and the resolution quoted in the table are approximate.

CMIP6 Model Name	Horizontal Ocean Resolution (lon. by lat. in degree)	Variant Label	ESM Reference	Data
ACCESS-ESM1-5	$1°\times1°$	r1i1p1f1	Law et al. (2017)	Ziehn et al. (2019)
CanESM5	$1°\times1°$	r1i1p2f1	Swart et al. (2019a)	Swart et al. (2019b)
CESM2	$1.125°\times0.53°$	r10i1p1f1	Danabasoglu et al. (2020)	Danabasoglu (2019a)
CESM2-WACCM	$1.125°\times0.53°$	r1i1p1f1	Liu et al. (2019)	Danabasoglu (2019b)
CMCC-ESM2	$1°\times1°$	r1i1p2f1	Lovato et al. (2022)	Lovato et al. (2021)
CNRM-ESM2-1	.3°-1°	r1i1p1f2	Séférian et al. (2019)	Seferian (2018)
GFDL-CM4	$0.25°\times0.25°$	r1i1p1f1	Held et al. (2019)	Guo et al. (2018)
GFDL-ESM4	$0.5°\times0.5°$	r1i1p1f1	Dunne et al. (2020)	Krasting et al. (2018)
IPSL-CM6A-LR	.3°-1°	r1i1p1f1	Boucher et al. (2020)	Boucher et al. (2018)
MIROC-ES2L	$1°\times1°$	r1i1p1f2	Hajima et al. (2020)	Hajima et al. (2019)
MPI-ESM1-2-HR	$0.4°\times0.4°$	r1i1p1f1	Müller et al. (2018)	Jungclaus et al. (2019)
MPI-ESM1-2-LR	$1.5°\times1.5°$	r1i1p1f1	Mauritsen et al. (2019)	Wieners et al. (2019)
MRI-ESM2-0	$1°\times(0.3\text{-}0.5)°$	r1i2p1f1	Yukimoto, Kawai, et al. (2019)	Yukimoto, Koshiro, et al. (2019)
NorESM2-LM	$1°\times1°$	r1i1p1f1	Tjiputra et al. (2020)	Seland et al. (2019)
NorESM2-MM	$1°\times1°$	r1i1p1f1	Seland et al. (2020)	Bentsen et al. (2019)
UKESM1-0-LL	$1°\times1°$	r1i1p1f2	Sellar et al. (2019)	Tang et al. (2019)

trend of Januaries at a given grid-point is removed from the respective time-series comprising all January values. The degrees of freedom of the spline is set to 5 for a good compromise between the smoothness (smoothing parameter above 0.8) and the number of parameters (knots) of the spline used to estimate the trend for all tropical Pacific grid-points (Chap.10, Hastie & Tibshirani, 1990).

The Niño 3.4 index is also computed for HadISST and for each model. It corresponds to the standardised area-weighted mean SST anomaly over the Niño 3.4 region: 5°S-5°N × 190°-240°E. These anomalies are computed relative to the 1981-2010 climatology. The SST values are first detrended over the 1870-2021 (for HadISST) and 1850-2100 (for the CMIP6 models) period using the same cubic smoothing spline (degrees of freedom set to 5). Then, the Niño 3.4 index for each model is computed relative to the respective 1981-2010 climatology.

A principal component analysis (PCA) is applied to the reference SST anomalies from HadISST in order to reduce the dimension of the data while keeping most of the variability. SST anomalies are weighted by the square root of the cosine of the latitude to give equivalent weights to all grid-cells (Vrac et al., 2014). The first principal component (PC) accounts for more than 56% of the total SST anomaly variance while 14 are needed to retain 90%. In this study, the four leading PCs containing more than 78% of the total variance have been kept for clustering. This choice has been made based on the stability of the clustering performed on these four PCs, and further presented below.

This 4-dimensional (4-d) space defined from the four leading PCs of HadISST anomalies (sometimes referred to as the "phase space") are used for defining ENSO events in both the observation-based analyses and the ESM simulations. Monthly anomalies from the simulations are projected onto the four spatial patterns (also known as EOFs) associated with the HadISST PCs to obtain the four leading "pseudo-PCs" for each ESM. The term pseudo-PC is used to differentiate them from the actual PCs obtained from a PCA computed for each ESM. Using the same HadISST-based phase space for all data sets allows for a consistent comparison of the regime patterns. Indeed, performing PCA separately for each model simulation would add a complicating factor to the analysis: for instance, the spatial pattern associated with the 4^{th} PC in HadISST could be associated with the 5^{th} PC of a model, a mismatch which would penalize the model in terms of performance.

2.2 ENSO Regimes Definition

In this section, the methodology to define the regimes associated with ENSO is described. Our approach consists of clustering the 4-d time series of PCs, representing monthly HadISST SST anomalies to define the observation-based reference ENSO regimes that are used as benchmark regimes to evaluate the models. Our clustering approach is based on a Gaussian mixture model (GMM, Pearson, 1894; Peel & McLachlan, 2000). It relies on the fact that any probability density function (pdf) f can be approximated by a weighted sum of K Gaussian pdfs f_k $(k = 1, \cdots, K)$:

$$f\,(\,x\,) = \sum_{k=1}^{K} \pi_k\, f_k(\,x\,;\,\alpha_k\,), \tag{1}$$

where α_k corresponds to the parameters (mean μ_k and covariance matrix Ω_k) of pdf f_k and π_k is the mixture ratio, also referred to as the prior probability. The parameters α_k and π_k are to be estimated. Then, each of the K estimated Gaussian pdfs characterises one cluster, in the sense that each cluster C_k is supposed to be generated from one specific density function f_k. In this study, GMM is preferred to k-means due to a key limitation of k-means: all clusters are equal in size (or volume) and spherical (*i.e.*, all clusters have the same diagonal covariance matrix Ω, so that the cluster assignment is made solely based on the distance to the cluster center, which can lead to statistically suboptimal splits. The GMM is more flexible because it accounts for both variances and covariances in the assignment process (Rust et al., 2010). The GMM result is thus able to accommodate clusters of variable size as well as intra-cluster correlations much better than k-means.

The estimation of the GMM parameters, μ_k, Ω_k and π_k is performed iteratively using the Expectation Maximization (EM, Dempster et al., 1977) algorithm by maximizing the likelihood (Fraley & Raftery, 2002). The parameters are initialized by the result of a model-based hierarchical agglomerative clustering. The result is a tree-like structure, which proceeds from n clusters containing one month each to one cluster containing all n month as object clusters are successively merged. This provides the basis for an educated initialisation of the EM algorithm for any number of mixture components (*i.e.*, Gaussian pdfs) and parametrisations of the component covariance matrices and helps to avoid a local maximum when optimising the likelihood function. Scrucca and Raftery (2015) provide a thorough description of the initialisation.

EM is based on the principle that the π_k is calculated when knowing α_k and vice-versa, thus optimizing successively and iteratively both. To be more specific, after the initialization (iteration 0) of the parameters α_k^0, μ_k^0 and Ω_k^0, each iteration i consists of the following two steps:

1. Expectation-step (or E-step) estimates the posterior probability τ_k^i (update of π_k^i) that the the 4-d data x_m for month m belongs to cluster C_k:

$$\tau_k^i(x_m) \;=\; \frac{\pi_k^i\, f_k\,(\,x_m\,,\,\alpha_k^i\,)}{\sum_{k=1}^{K} \pi_k^i\, f_k\,(\,x_m\,,\,\alpha_k^i\,)}. \tag{2}$$

2. The Maximization-step (or M-step) uses the posterior probabilities to improve the estimates of GMM parameters (iteration $i+1$):

$$\pi_k^{i+1} \;=\; \frac{1}{n} \sum_{m=1}^{n} \tau_k^i(x_m), \tag{3}$$

$$\mu_k^{i+1} \;=\; \frac{1}{n\,\pi_k^{i+1}} \sum_{m=1}^{n} x_m\, \tau_k^i(x_m), \tag{4}$$

$$\Omega_k^{i+1} = \frac{1}{n\,\pi_k^{i+1}} \sum_{m=1}^{n} \tau_k^i(x_m)\,(\,x_m - \mu_k^{i+1}\,)'\,(\,x_m - \mu_k^{i+1}\,), \tag{5}$$

where n is the number of months.

To summarise, the EM algorithm iteratively repeats (i) E-step estimating the posterior probabilities that the x_m belongs to cluster C_k from the updated parameters of the GMM and (ii) M-step estimating the GMM parameters from the updated posterior probabilities.

Finally, each cluster C_k is defined, according to the principle of posterior maximum:

$$C_k = \{\,x_m\,;\,\pi_k\,f_k\,(\,x_m\,;\,\alpha_k\,) \geq \pi_j\,f_j\,(\,x_m\,;\,\alpha_j\,)\,,\,\forall\,j = 1, \ldots K\}. \tag{6}$$

In other words, a cluster contains all monthly data whose probability of belonging to that cluster is maximised.

The freedom of EM in the definition of the regimes depends on the number K of clusters and on the constraints applied to the covariance matrices Ω_k (Fraley & Raftery, 2002). EM is performed several times with different constraints of the GMM covariance structure (see, Fraley & Raftery, 2002; Dempster et al., 1977) and several numbers K of clusters. Hence, in practice, several GMMs are fitted and it is needed to select the "best" one. This is typically a "model selection" problem. The Bayesian Information Criterion (BIC) is a traditional tool in statistics to perform such a task (Schwarz, 1978). The BIC is used for model selection and helps to prevent overfitting by introducing penalty terms for the complexity of the GMM (*i.e.*, the number of parameters). Hence, minimizing the BIC achieves a good compromise between keeping the model simple and providing a good representation of the data. The BIC is given by:

$$BIC\,(\,K\,) = p\;log\,(\,n\,) - 2\;log\,(\,L\,), \tag{7}$$

where K is the number of clusters, L the likelihood of the parametrized mixture model, p the number of parameters of the GMM to estimate, and n the size of the sample (*i.e.,* total number of months from January 1920 to December 2014, which is 1140 months).

The clustering described above is performed using the R package 'Mclust' (Scrucca & Raftery, 2015).

A different approach is used to assign each month in the model data to a specific regime. The EM algorithm is not applied, but 4-d representation of monthly SST anomalies (pseudo-PCs from 1850-2100) of each model is associated with the most appropriate HadISST regime based on the principle of posterior maximum (see Eq. 6). Thus, the regimes are consistently defined for all simulations in the sense that, in the following, ENSO regimes in the models actually represent similar regime determined from HadISST SST anomalies. In addition, the variability in the clustering itself as a possible source of noise is ruled out. Such defined regimes are used to compare the regime patterns and their temporal properties across different model simulations within a common reference framework. In practice, the common reference is ensured by computing τ_k^i from Eq. 2 using the GMM parameters estimated from HadISST (hence the common framework) but using pseudo-PCs from each model.

3 Results

3.1 Reference HadISST ENSO Regimes

The optimal number K of ENSO regimes that best describe SST anomalies spanned by the four leading PCs was determined using the clustering approach described in section 2.2. In order to get a robust number of regimes, a bootstrap-like procedure has been implemented. The EM algorithm used to define the clusters (regimes) has been applied 250 times to a sub-sample of the total set containing 75% of the data randomly selected (*i.e.* without replacement) and the BIC has been computed for each K from 2 to 10 for each sub-sample. BIC values are presented as violin plots in Figure 1. The fraction of total draws that results in $K \in [2, \cdots, 10]$ clusters is given in the insert; for example, 58% of the 250 sub-samples show $K = 5$ is the optimal number of clusters.

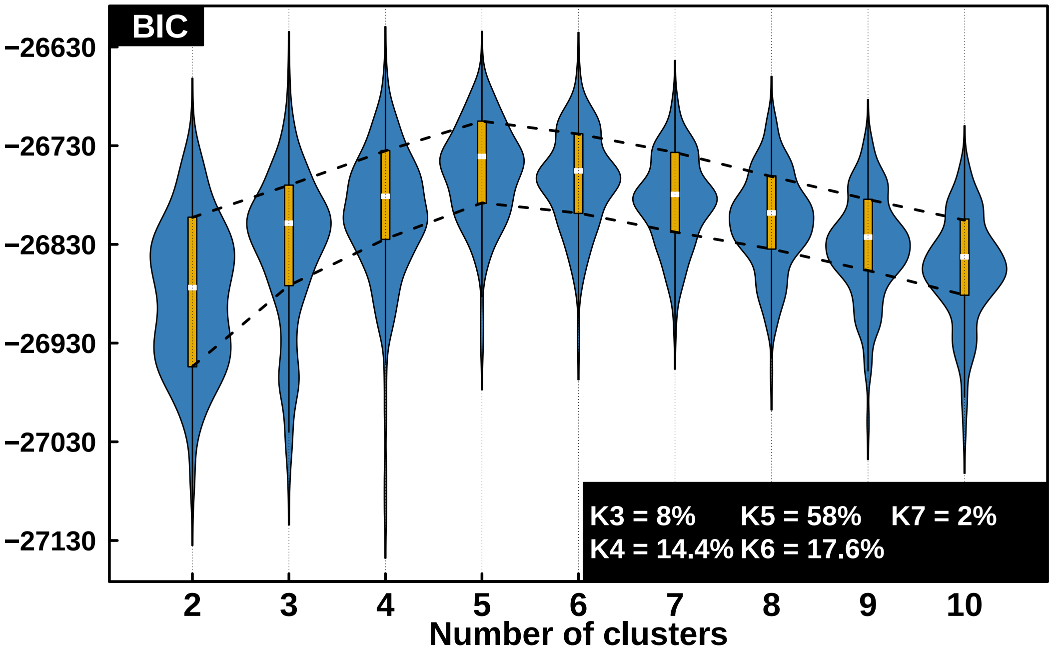

Figure 1. Violin plots represent BIC values as a function of K obtained by applying the EM algorithm 250 times to sub-samples of the total set containing 75% of the data randomly selected. Yellow boxes indicate BIC inter-quartile range and the median is indicated by white dots. The BIC is computed for each K from 2 to 10 for each sub-sample. The ratio (in %) of how often a given value of K is selected as optimal is also given in the bottom.

The sensitivity of the clustering results to the number of PCs has been tested (not shown). Results for higher numbers of PCs from the bootstrap procedure yields unclear results in terms of optimal number of clusters (usually higher than five), with the additional clusters not describing to known ENSO phases. This explains our choice of 4 PCs for the clustering.

Figure 2 a) represents the average HadISST pattern of the five reference regimes determined with the EM algorithm. Two La Niña regimes (basin-wide La Niña BW-LN, central La Niña C-LN), two El Niño regime (central El Niño C-EN, eastern El Niño E-EN) and one Neutral regime are obtained. BW-LN is the most frequent (13.3%) La Niña configuration showing strong negative SST anomalies covering a large portion of the tropical Pacific. C-LN shows negative anomalies more circumscribed to the equatorial area with positive anomalies in the southeastern part of the domain. Both La Niña regimes have similar ranges of intensity with similar average Niño 3.4 indices (see Fig. 2 b). C-EN is the most frequent El Niño regime with strongest positive SST anomalies close to the equator. E-EN is the most intense regime with large positive anomalies in the eastern Pacific. Similar results are obtained from the clustering obtained over a shorter period (1950-2014) and from JRA-55 reanalyses over the 1958-2019 period (see supplementary Figure S1; Kobayashi et al. (2015) and Harada et al. (2016)).

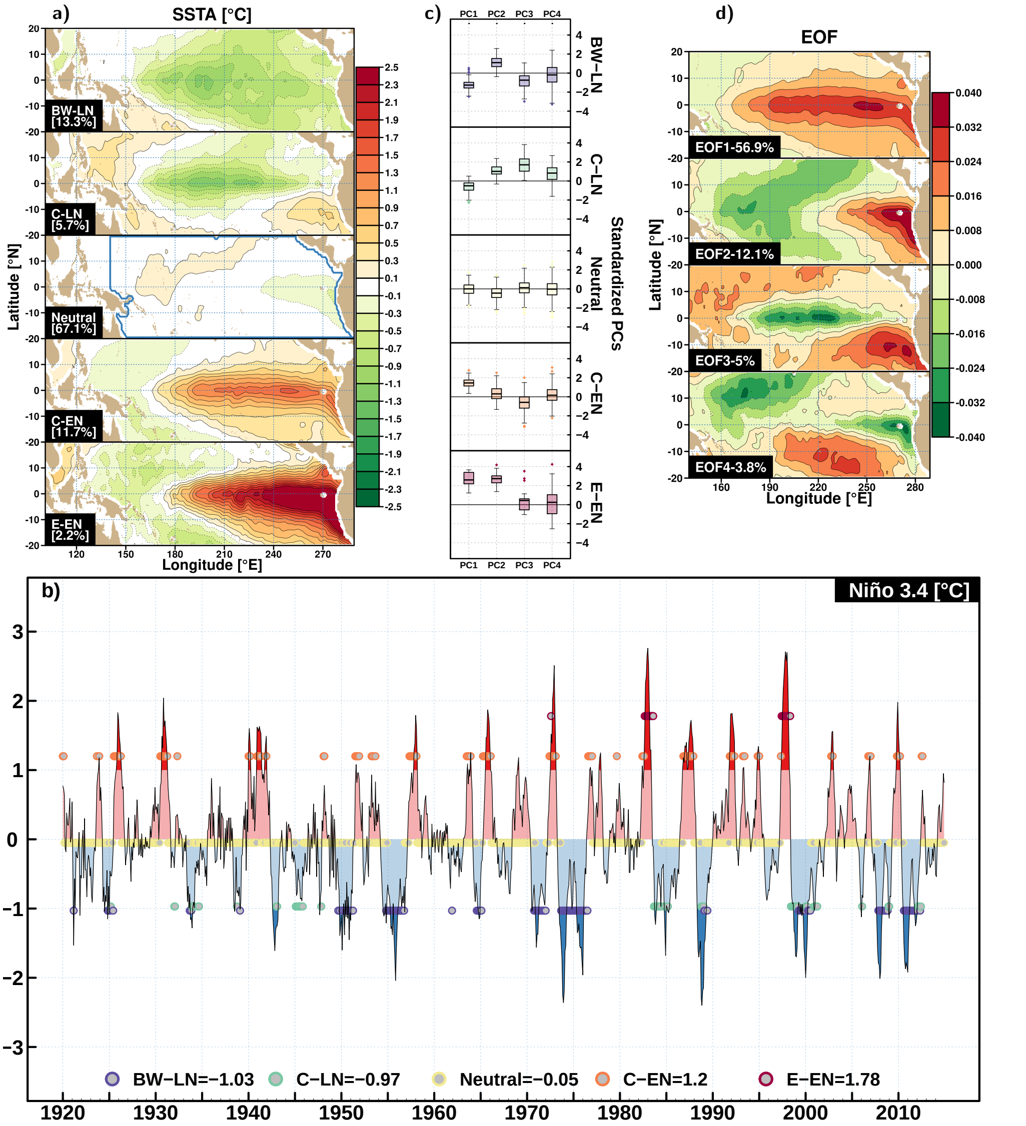

Figure 2. a) Maps of the five ENSO regimes in observations defined by EM. Colours correspond to average SST anomaly within a regime in °C. The frequency (in %) of occurrence of each regime is given in the bottom left corner of each panel. The blue contour in the Neutral panel indicates the area used to perform the clustering. b) Monthly Niño 3.4 index time series (solid line with red or blue shading when Niño 3.4 is positive or negative). The coloured dots show the assigned regime for each month with the vertical position indicating the average Niño 3.4 value of that cluster (given at the bottom, in °C). c) Boxplots showing the distributions of the four standardised PCs within each regime. Boxes indicate inter-quartile range, whiskers indicate 1.5 times the inter-quartile range from the box and the dots are the values beyond that range and the middle bar the median of the PCs over the 1920-2014 historical period. d) Spatial patterns associated with the four leading obtained from HadISST SST anomaly over the 1920-2014 period. The fraction of total SST variance explained (in %) by each EOF is indicated in the lower left corner of each panel.

The time series of the Niño 3.4 index and the cluster assigned to each month are shown in Fig. 2 b), which depicts that El Niño and La Niña events are well captured by the cluster index. For example, the cluster E-EN corresponds to the strong El Niño events (*e.g.,* 1972-73, 1982-83, 1997-98, K. Takahashi et al., 2011; Ren et al., 2018). Central Pacific El Niño events (1986-87, 1991-92, 1994-1995, 2002-03, 2004-05, and 2009-10) are consistent with cluster C-EN. Similarly, the BW-LN regime contains strong La Niña events (*e.g.,* 1954-56, 1973-74, 1975-76, 1988-89, 1998-2000, 2007-08, Ren et al., 2018, and references therein). C-LN rather corresponds to moderate La Niñas, a regime that is in transition from an extreme El Niño (in 1983 and 1998) to an extreme La Niña.

To identify which of the four leading PCs are the most important for each regime, boxplots of their PC distributions are shown in Fig. 2c) and the spatial patterns associated with those PCs are given in Fig. 2d). The warm ENSO patterns are mainly determined by PC1 and PC2 with PC1 dominating for C-EN. This is quite straightforward since PC1 and PC2 are is associated with central Pacific positive pattern (EOF1) and to a West-East dipole (EOF2). Although cold patterns are partly explained by PC1 and PC2 (with almost the same contributions for both), PC3 and PC4 are indispensable for capturing them. In particular, these latter PCs are needed to differentiate BW-LN from the C-LN regimes. C-LN regime has a strong positive contribution from PC3 (EOF3) and moderate a positive contribution from PC4. While BW-LN has a moderate negative contribution from PC3 and no contribution from PC4.

In the next section, consistency in the pseudo-PC weighting across nearly all the models and observations is shown, especially for the two La Niña patterns and the C-EN pattern (see supplementary Figure S2). This indicates that models are able to simulate regime patterns that are similar to those in the observations, and that by projecting model data onto the observed EOFs, temporal information (about *e.g.*, pattern frequencies and probabilities of transition) can be extracted from the models and compared to those in observations. This also advocates for our approach instead of using the dataset specific EOFs.

3.2 Model Evaluation over 1920-2014

ENSO regimes from CMIP6 models are evaluated relative to the reference regimes (HadISST) in terms of spatial patterns, frequency of occurrence of each ENSO regime, the average persistence within each regime (defined as average duration in months a model remains in each regime from the moment that model enters it), and the transition probability from one regime to another.

First, the ability of each model to reproduce the reference patterns is assessed by associating pseudo-PCs from the models with the most appropriate reference regime. Supplementary Figure S3 shows spatial patterns of the ENSO regimes obtained for CMIP6 models and HadISST over the historical period. Interestingly, every model is able to reproduce patterns resembling the reference regimes in terms of spatial distribution and intensity of SST anomalies. In particular, the asymmetry and the diversity of ENSO event spatial patterns in the reference regimes are well reproduced in the CMIP6 models. However, there are some notable differences: the extrema in regimes are usually more intense and spatially broader (for BW-LN, C-EN and E-EN) in the models than in the observations. The extrema of the E-EN regime in the models are not located as far east as in the E-EN regime in HadISST. SST anomalies patterns are also zonally more extended in the models compared to the patterns in the observations and extend too far west (all except the neutral regime). Figure 3 presents the Taylor diagram for the average SST anomalies of each ENSO regime in the 1920-2014 historical period. Taylor diagrams are used to evaluate the agreement between average simulated and reference regime patterns. They summarise three statistics comparing simulated grid point 'centered' values ('centered' means that the spatial average is subtracted from each grid-point value) to a reference value (represented by the red diamonds and lines): 1) the Pearson correlation coefficient measuring the 'similarity' between pairs of centered simulated and reference values is given by the azimuthal position; 2) the centered root mean square error between the mean centered values of the observations and the simulation is given by the green curves; 3) the standard deviation of simulated and observed pattern values are proportional to the radial distance from the origin (for more details, see Taylor, 2001). Therefore the closer a simulation marker is to the reference one (red diamond), the better is the model.

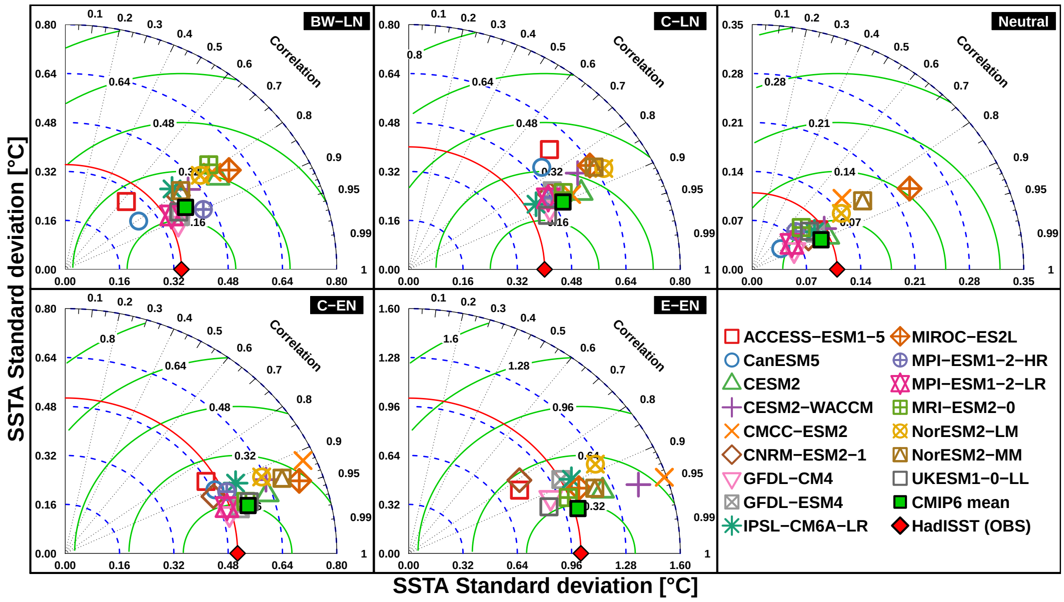

Figure 3. Taylor diagrams for each of the regime patterns from each CMIP6 model and observations (HadISST) over the 1920-2014 period. Each coloured marker refers to one climate model. Red diamonds and red curves indicate the spatial standard deviation of the clusters obtained from the observations.

For each regime, all models show similar spatial patterns as HadISST (spatial correlation typically between 0.8 to 0.9) but with amplitudes that vary greatly across models. Note that the E-EN regime shows greater differences between models and observations, and accordingly, has a larger centered root mean square error. Models are then ranked based on their metric performance depicted in the Taylor diagrams. The models are first ranked according to each regime and all ranks are then added (the smaller the sum, the better the model) to obtain the rank reported in the first column of Table 2. GFDL-CM4, UKESM1-0-LL, GFDL-ESM4, MPI-ESM1-2-LR and CNRM-ESM2-1 are the top five ESMs for the spatial representation of ENSO. The CMIP6 ensemble mean is ranked between models 1 and 2.

The frequency of occurrence of each ENSO pattern over the historical period is shown in Figure 4. This varies from one model to another but it roughly agree with the regime frequency in the observations. In particular, the models feature an E-EN regime that occurs less frequently than the C-EN regime, and a C-LN regime that occurs less frequently than the BW-LN regime. However, a few models do not (CanESM5, and MPI models) or too rarely (CNRM-ESM2, UKESM and GFDL models) simulate the E-EN pattern, or produce too evenly distributed regime frequencies (CESM2-WACCM and NorEMS2-MM). In order to rank the models, the absolute value of relative frequency bias (in %) is computed for each regime (see supplementary Fig. S4 for actual and absolute bias).

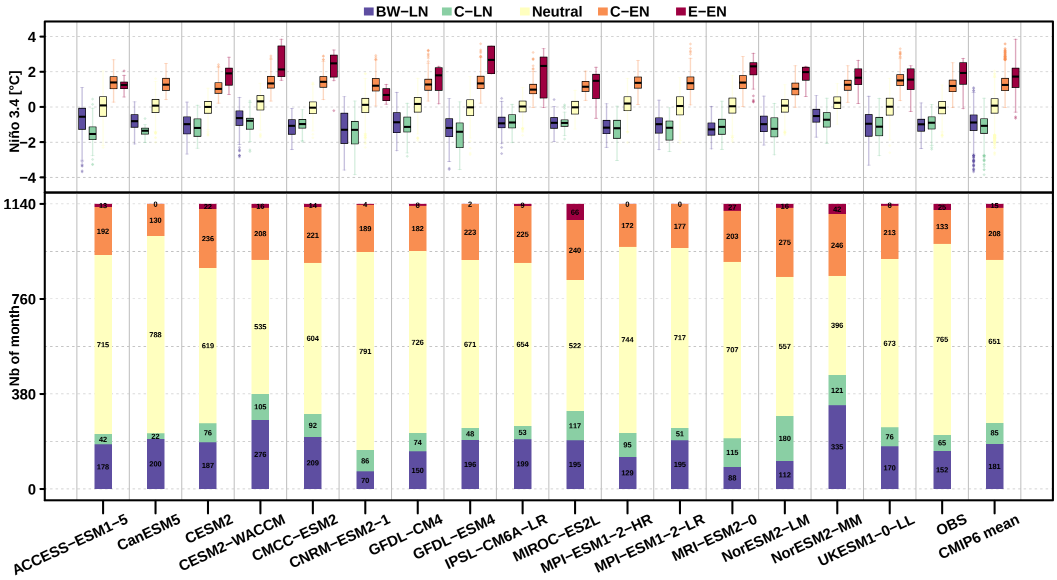

Figure 4. (Bottom) Barplots of each ENSO regime frequency for all the models and HadISST (observations) over the 1920-2014 period. (Top) The boxplots above indicate the Niño 3.4 index distribution for each model and regime.

Relative frequency biases are larger for C-LN and E-EN, which is expected given their lower occurrence frequency. The frequency bias for each pattern is then combined to produce the "average frequency bias" metric reported in column 2 of Table 2 alongside their corresponding ranks. GFDL-CM4, CESM2, ACCESS-ESM1-5, UKESM1-0-LL and MRI-ESM2-0, and are the top five ESMs for the frequency representation of ENSO regimes. The CMIP6 mean is positioned between models 3 and 4.

ENSO events generally peak during boreal winter. Figure 5 depicts the monthly ratio (in %) of how each regime is distributed throughout the year. In the observations, the Neutral pattern occurs more often outside the winter months while C-EN ad C-LN show higher frequencies during the winter. In contrast, BW-LN and E-EN seem to be quite evenly distributed throughout the year. In the models, the seasonality is generally consistent with HadISST for the Neutral regime and the two La Niña regimes BW-LN and C-LN, but the models do not produce the marked seasonality in the (most frequent) El Niño regime, C-EN. This is consistent with previous studies showing the inability of CMIP6 (and also CMIP3 and CMIP5) models to correctly simulate ENSO peaking in winter (see, H.-C. Chen & Jin, 2021, and references therein). The corresponding Taylor diagram is given in Fig. S5 of the supplementary material. Correlations do not exceed 0.6 for any model meaning that the seasonal variation of pattern occurrences is not well represented in the models.

The average persistences of the observed ENSO regimes are 4.7, 3.2, 7, 4.2 and 8.3 months for, respectively the BW-LN, C-LN, Neutral, C-EN and E-EN regimes. Supplementary Fig. S4 gives the persistence bias in the models. Models are either over- or under-estimating the persistence in the BW-LN (from -2 up to 3 month), C-LN regimes (± 2 months) and the Neutral regime (± 2 months). Persistences of C-EN regime are rather over-estimated (up to 2 months). For E-EN regime, whose frequency is under-estimated by the models, the persistence is also widely under-estimated. Similar to the frequency, the absolute persistence bias is computed (see Fig. S4) for each model and the average is reported with their rank in column 3 of Table 2. The top five models are CESM2-WACCM,

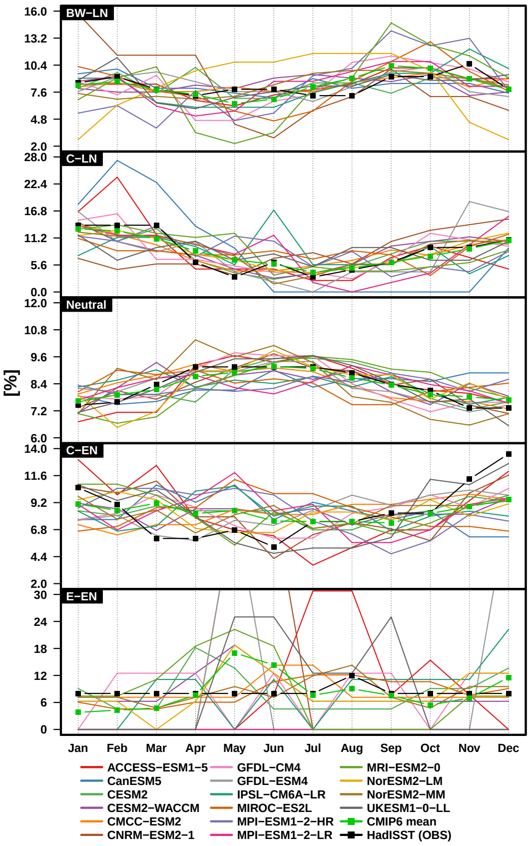

Figure 5. Monthly occurrence ratio (%) for all regimes and all the models and HadISST over the 1920-2014 period.

MIROC-ES2L, MRI-ESM2-0, GFDL-CM4 and NorESM2-MM. On average, the CMIP6 ensemble achieves better performance than any individual model.

Figure 6a) shows the month-to-month transition diagram from one reference ENSO regime to another. The probability of remaining in any given regime ranges from 69 to 92%, which is higher than any transition. The second most favoured transition for BW-LN, C-LN and C-EN is towards the Neutral regime (resp. at 12, 23 and 22%). For the E-EN regime, the second transition is towards C-LN (8%) which interestingly happened after the very strong El Niño events of 1982-83 and 1997-98. There is no direct transition towards the Neutral regime. Direct transitions from either La Niña regime to C-EN and between La Niña regimes are rare.

Transition probability matrices for each model and for observations are given in supplementary Figure S6. The Taylor diagram in Fig. 6b built from those matrices compares the ability of the models to reproduce the transitions of the reference regimes. The poorest performing models tend to underestimate the persistence of E-EN and transition too frequently from E-EN to C-EN (lower right corner of the matrices in Fig. S6), mostly due to the low frequency or the absence of occurrence of the E-EN regime. Models are ranked according to their transition behaviour based on the Taylor diagram in column 4 of Table 2. The top five models are MRI-ESM2-0, CMCC-ESM2, CESM2-WACCM, CESM2 and MIROC-ES2L. CMIP6 mean is positioned between model 7 and 8.

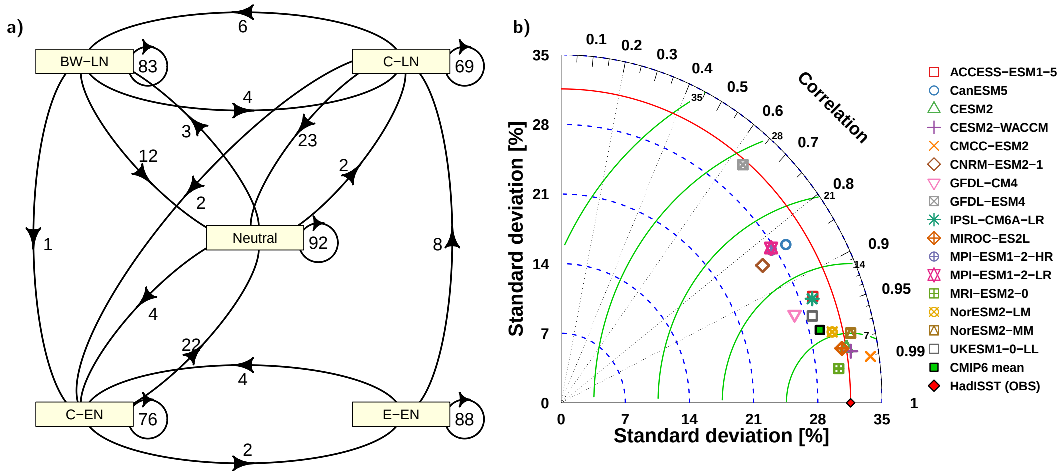

Figure 6. a) Transition diagram from one regime to another obtained for HadISST; values are the transition probability (in %). The probability of remaining in a regime is noted by the circled values (in %). b) The Taylor diagram evaluating the regime transition probabilities in the CMIP6 models compared to the regime transition probabilities in the observations.

Table 2. Model rank for each of four metrics based on model bias (given in parenthesis) or Taylor diagram. The top five models according to each metric are bolded. The overall rank is calculated by adding the rank according to each metric (given in parenthesis in column 5). The top five models according to the overall rank are highlighted in grey. The position of CMIP6 ensemble average is given in the last row.

	Spatial pattern	Average frequency absolute relative bias (%)	Average persistence absolute bias (month)	Transition probability	Overall rank (total)
ACCESS-ESM1-5	11	**3 (30.3%)**	15 (2.24)	11	9 (40)
CanESM5	12	9 (40.61%)	9 (1.97)	13	14 (43)
CESM2	7	**2 (29.7%)**	8 (1.96)	**4**	4 (21)
CESM2-WACCM	10	13 (53.1%)	**1 (1.38)**	**3**	5 (27)
CMCC-ESM2	15	10 (42%)	16 (2.79)	**2**	14 (43)
CNRM-ESM2-1	**5**	11 (43.2%)	14 (2.2)	12	12 (42)
GFDL-CM4	**1**	**1 (25%)**	**4 (1.66)**	9	1 (13)
GFDL-ESM4	**3**	12 (45.4%)	13 (2.17)	16	16 (44)
IPSL-CM6A-LR	6	8 (39.4%)	10 (1.97)	10	6 (34)
MIROC-ES2L	16	15 (76.9%)	**2 (1.41)**	**5**	8 (38)
MPI-ESM1-2-HR	9	7 (38.7%)	11 (2.02)	14	11 (41)
MPI-ESM1-2-LR	**4**	6 (37.8%)	12 (2.11)	15	7 (37)
MRI-ESM2-0	8	**5 (37.4%)**	**3 (1.45)**	**1**	2 (17)
NorESM2-LM	14	14 (74.6%)	7 (1.85)	7	12 (42)
NorESM2-MM	13	16 (81.5%)	**5 (1.71)**	6	9(40)
UKESM1-0-LL	**2**	**4 (33.8%)**	6 (1.84)	8	3 (20)
CMIP6 mean	1-2	3-4 (31.7%)	<1 (1.15)	7-8	2 (11-14)

3.3 Future Changes

The changes in the regime frequencies under a high-warming future scenario are analysed. As described in section 2.2, the frequency of the model regimes is obtained by matching the pseudo-PC of each model to the most appropriate reference regime. Thus, changes in regime frequency in the models are not artifacts of potential changes in the spatial patterns of regimes with global warming. Figure 7 shows the ENSO regime frequency over the 1965-2014 historical and 2051-2100 future periods.

The most consistent result is the projected decrease in the BW-LN regime (16 out of 16 models). In contrast, the other La Niña regime (C-LN) is expected to occur more frequently in the future for 12 out of 16 models. Similarly, C-EN and E-EN frequency

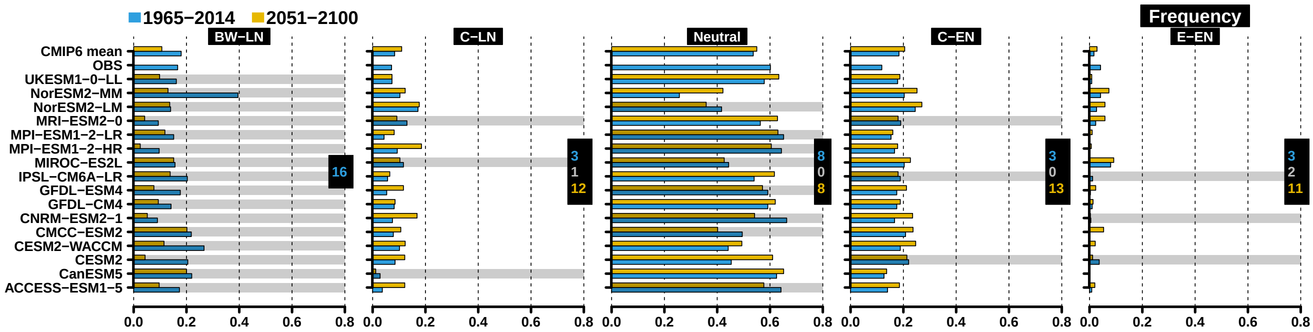

Figure 7. Regime frequencies over the 1965-2014 historical (blue) and the 2051-2100 future (yellow) periods. Grey shading designates the models with lower regime frequency in the future compared to historical period. The number of models with lower, equal and higher occurrence in the future is given in blue, grey and yellow, respectively, for each regime.

is also expected to increase in the future for the majority (13 and 11) of the models. For the Neutral regime, there is no clear consensus with half of the models projecting increased frequency in the future. The decrease in frequency of BW-LN cold events and increase in frequency of C-EN and E-EN warm events cannot be a consequence of the mean warming trend since the latter has been removed by detrending the model output.

Another way to investigate regime frequency days is through continuous long-term trends in both HadISST and CMIP6 model simulations (respectively over the 1920-2014 and 1850-2100 periods). The linear trends are estimated from the 30-year running mean of the regime frequency time series (see supplementary Figure S7). Figure 8 presents the sign of significant linear trends of ENSO regime frequencies. A trend is considered significant at the 95% confidence level ($\alpha = 0.05$) based on a t-test on the null hypothesis that there is no trend (slope is equal to 0, estimated with *lm* function; R Core Team, 2020). This trend analysis shows that the frequencies of E-EN, C-EN and C-LN regimes are projected to increase significantly in respectively 12, 12 and 13 models, by the end of the 21$^{\text{st}}$ century. This is consistent with their higher occurrences in the future period shown in Fig. 7 and the historical trends of the reference regimes (Figure S7). The Neutral regime frequencies shows a significant decreasing trend in 14 models and in observations while BW-LN is projected to significantly decrease in 8 and increases in 5 of the models. The same trend results are obtained using a non-parametric trend test (*e.g.,* the Theil-Sen test, not shown).

Figure 9 shows the median and the standard deviation of the Niño 3.4 index, within each cluster, for the reference and model regimes over the 1965-2014 historical and the 2051-2100 future periods. The C-LN and C-EN clusters are associated with more intense SSTA in the future for respectively 14 and 11 models with a larger median Niño 3.4 index. For BW-LN and E-EN, the results are mixed with, respectively, 9 and 8 of the models projecting more intense patterns. In terms of variability, the BW-LN, C-LN, Neutral and C-EN regimes are expected to show increased variability with, respectively, 11, 12, 14 and 11 with higher intra-regime Niño 3.4 standard deviation in the future. Given the low frequency of the E-EN no consistent conclusion can be drawn for the change in Niño 3.4 variability for that regime.

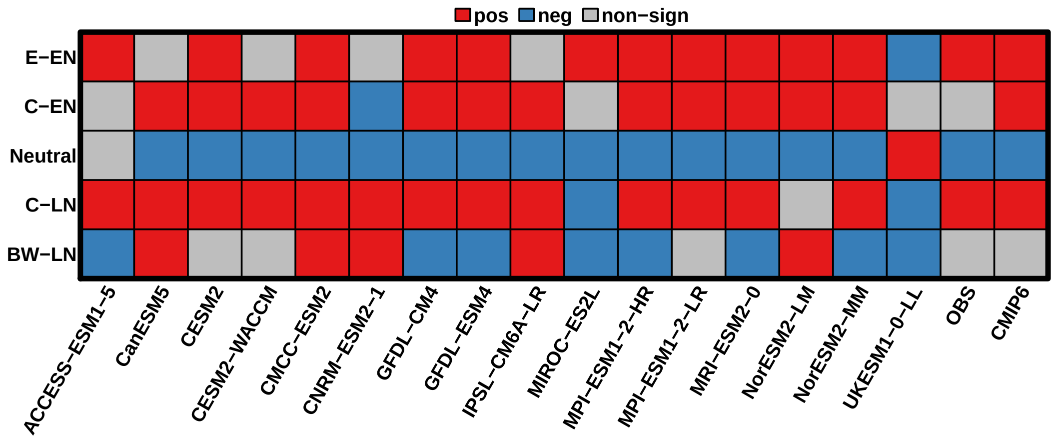

Figure 8. ENSO regime frequency trends over 1850-2100 for CMIP6 models and 1920-2014 of the observations. Significant positive (negative) trends are given in red (blue) and non significant trends are given in grey. The CMIP6 ensemble mean trends are also given.

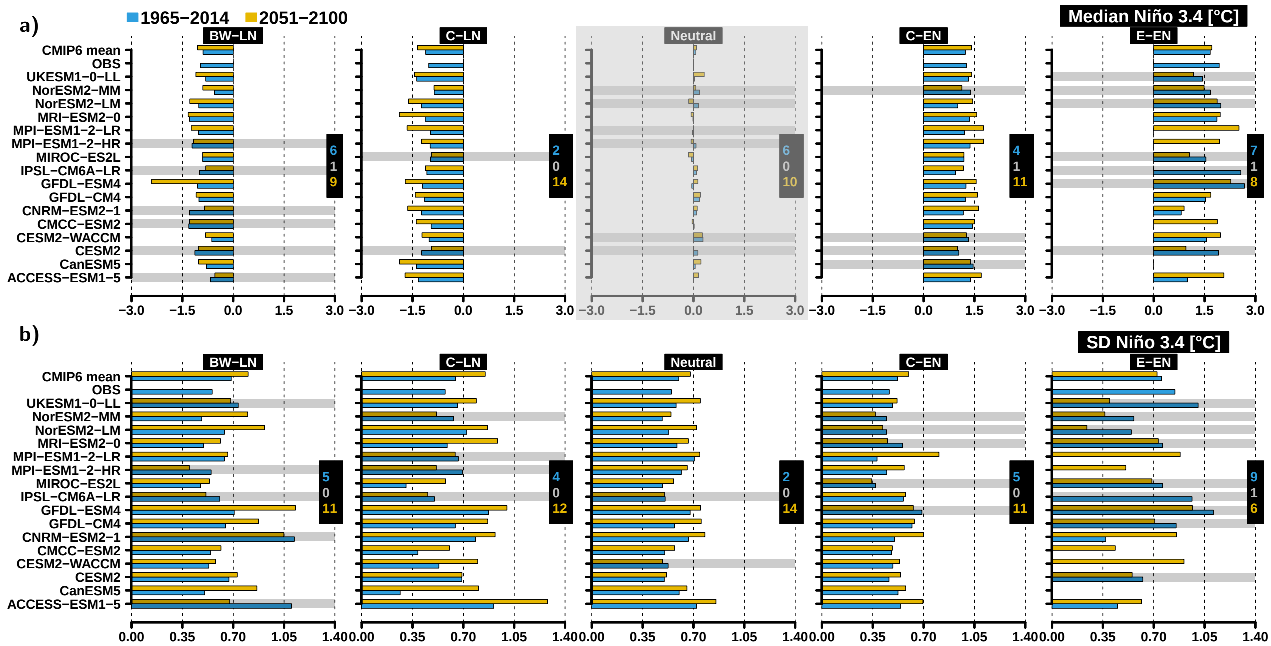

Figure 9. Median a) and standard deviation b) of Niño 3.4 for each regime over the 1965-2014 historical (blue) and the 2051-2100 future (yellow) periods. Grey shading designates the models with smaller median and standard deviation in the future compared to the historical period. The number of models with smaller, equal and larger Niño 3.4 statistics in the future are given in blue, grey and yellow, respectively, for each regime. The Neutral panel is greyed out for the median because it is not meaningful and hence is not considered.

4 Discussion

4.1 Evolution of ENSO in the Historical period in observations

Previous studies of ENSO in observations and in climate models make use of traditional metrics such as variance in a "Niño.x" index (Rasmusson & Carpenter, 1982), x being the region over which SST anomalies are averaged, or variance in the indices E and C based on the first two PCs of tropical Pacific SST anomalies (K. Takahashi et al., 2011). Such metrics presume all phases of ENSO-related variability are captured by one or two patterns of SST variability that are independent of the phase of ENSO; for example regression of SST upon a Niño.x index yields a single representative pattern of variability for all phases of ENSO. Similarly, regression against E and C indices requires patterns of La Niña variability to be identical to patterns of El Niño variability. Unlike the traditional metrics of ENSO variability, the GMM clustering identifies five patterns of SST variability that capture the well-known differences in SST anomalies associated with the observed El Niño events and La Niña events, including the different amplitudes and structures of eastern vs. central Pacific El Niño events and the different La Niña events.

The inadequacy of the E and C indices in representing either type of La Niña event – or even central Pacific El Niño events – is already evident from the different locations of the SST extrema in La Niña event and El Niño events (cf. the top two panels of Fig. 2a) to the bottom two panels) as well as from the weighting of the PCs that comprise these regimes/phases of ENSO (Fig. 2c)). In terms of the observations, the bottom panels of Figure 2 a) and c) show the rare far eastern Pacific warm events E-EN (i.e., eastern Pacific El Niños) are well characterized by a combination of PC1 and PC2. If we reverse the sign of the second PC in Fig. 2d) to conform with the convention adopted in K. Takahashi et al. (2011), where E = (PC1-PC2)/$\sqrt{2}$, a direct correspondence to the large positive values of the E index associated with these events is found. In the central Pacific, warm events are mostly captured by the first PC1 of tropical Pacific SST, while PC2 mainly contributes to cold central Pacific events (see top two panels of Fig. 2 a) and c)). The two combine to explain the skill of the C index C=(PC1+PC2)/$\sqrt{2}$ in representing central Pacific SST variability. Our results also show that PC3 and PC4 contribute importantly to central Pacific SST variability, but are not accounted for in studies that characterize ENSO variability by the E and C indices (*e.g.*, K. Takahashi et al., 2011; Geng et al., 2022).

An added value of our approach is that the use of four EOFs allows a more comprehensive characterisation of ENSO-related SST variability, including that which contributes to variability in the traditional indices of ENSO variance (e.g., Niño3, Niño3.4 and Niño4). In particular, it allows one to characterise both warm and cold ENSO regimes and the transitions between them, suggesting that the PCs describe the continuous nuances in ENSO monthly evolution rather than distinct types of warm (or cold) events.

4.2 Comparison of ENSO variability over the Historical period simulated by the CMIP6 models to that observed

Applying the GMM clustering to the detrended output from each of 16 CMIP6 models for the Historical period (i.e., the models forced by the observed anthropogenic and natural forcing from 1920-2014) shows that the models, in general, reproduce the observed ENSO-related SST variability. However, there are some discrepancies between the observed and simulated regimes, including:

- models generally show broader and more intense ENSO patterns that extend too far west compared to those observed;
- a few models do not (CanESM5, and MPI models) or too rarely (CNRM-ESM2, UKESM and GFDL models) simulate the large amplitude eastern Pacific El Niño (E-EN) events, or produce too evenly distributed regime frequencies (*e.g.* CESM2-WACCM and NorEMS2-MM models);
- models generally feature central Pacific El Niño (C-EN) and La Niña (C-LN) events that are too frequent and have too large amplitudes compared to those observed;
- the strong seasonality of the central Pacific El Niño (C-EN) regime, which accounts for the overwhelming majority of El Niño events, is not captured in the models;

- the persistence of the central Pacific El Niño (C-EN) regime is overestimated while the persistence of the eastern Pacific El Niño (E-EN) regime is underestimated in the models;
- transitions between the regimes in the models are largely similar to those in observations except for the models having too rare or no eastern Pacific (E-EN) events.

When considering that models with ENSO properties in the historical period close to the observed ones are better in projecting potential future ENSO changes, our methods and metrics will help to identity the models that provide more reliable projections.

4.3 ENSO Regime Changes in the Future

GMM clustering of the CMIP6 model output for the end of the 21st Century under a high emission scenario reveals the following changes in ENSO-related SST variability:

- the higher amplitude La Niña (BW-LN) regime is in general projected to become less frequent but there is no consensus in terms of changes in magnitude;
- there is a strong consensus among the models that the central Pacific, moderate La Niña (C-LN) regime will become more frequent (significantly) and more intense and variable;
- the Neutral regime will become significantly less frequent and more variable in the future;
- the moderate El Niño (C-EN) regime will be significantly more frequent and is projected to become more intense and variable in the majority of the models;
- the strong El Niño (E-EN) regime is projected to become significantly more frequent but there is no consensus on how the magnitude changes.

Previous studies have reported that CMIP6 models project an increase in SST variance in the eastern tropical Pacific in the 21st Century compared to the 20th Century (*e.g.,* the AR6 WG1 IPCC, 2021). Using traditional Niño.x indices, Cai et al. (2022) and Maher et al. (2023) find enhanced ENSO variability over the course of the 21st Century compared to the end of the 20th Century (although the increase in variance is much smaller than the bias in the variance in the typical CMIP6 model). Geng et al. (2022) find increased variance in the E index, representing a pattern of SST anomaly in the far eastern equatorial Pacific, by the first half of 21st Century relative to the 20th Century. To understand the mechanisms responsible for the increase in ENSO variance, Maher et al. (2023) focused on changes in the mean state SST gradient, while Geng et al. (2022) argued for the importance of nonlinearity in the Bjerknes feedback in the models, which is absent in observations (e.g., Battisti et al., 2019, Fig. 8-15 and references therein).

Our results extend the findings of these studies to show that, compared to the 20th Century, there is a statistically significant increase in the frequency of occurrence of the common C-EN and rare E-EN patterns in most of the models in the 21st Century (Fig. 8). There is also a statistically significant increase in occurrence of C-LN events in most of the models at the expense of a decrease in occurrence of BW-LN events. Since (unlike in observations) many models feature a stronger cold anomaly in the C-LN pattern compared to that in BW-LN, the changes in the frequencies of cold BW-LN and C-LN patterns and the warm C-EN pattern (Fig. 7) act together to increase the total variance of SST in the eastern tropical Pacific projected in the 21st Century compared to the 20th Century. This effect is further amplified by the projected increase in the amplitude of the C-LN and C-EN patterns in the 21st Century, measured by the contributions of the patterns to Niño3.4 (Fig. 9b). It remains unclear to us to what extent the large bias in the amplitude of the models' central Pacific El Niño and La Niña events in the Historical simulations (Fig. 9b) jeopardizes the projections of increasing variance in the ENSO-related SST anomalies in the eastern Pacific.

4.4 Projected changes in ENSO variability in relation to projected changes in the SST mean state

Cai et al. (2021) report that the *amplitude* of eastern Pacific El Niño events in CMIP5/6 models increases in the 21^{st} Century compared to the 20^{th} Century, which would lead to changes in atmospheric teleconnections (because eastern Pacific El Niño events cause a greater eastward displacement in the centroid of precipitation, which is climatologically centered over the maritime continent, than central Pacific El Niño events). By first removing the simulated forced trends in SST, however, we find no systematic future response in the amplitude of the simulated east Pacific El Niño events in the CMIP6 models (8 models show an increase, while 7 show a decrease). Hence, the increase in amplitude of eastern Pacific El Niño events reported in Cai et al. (2021) must be due to the change in the simulated mean state SST, which features more warming along the equator in the eastern equatorial Pacific than in the western equatorial Pacific *i.e.*, a decrease in the climatological east-west SST gradient in the region. However, the observed long-term (e.g., 50-70 years) trend in SST along the equator shows the opposite warming pattern compared to both historical simulations and future projections: more warming in the western equatorial Pacific and little, if any, warming in the eastern Pacific, resulting in an increase in the mean climatological zonal SST gradient. There is increasing evidence that the observed trend in the equatorial Pacific SST gradient is indeed the response to anthropogenic forcing (*e.g.*, Seager et al., 2019; Dong et al., 2022; Wills et al., 2022) and the projected trend in the gradient (with more warming in the eastern than in the western equatorial Pacific) is a result of biases in the simulated mean state climate that are common to almost all the climate models (e.g., double ITCZs, too weakly stratified Southern Ocean). Should the observed trend in the zonal SST gradient indeed be the forced response, teleconnections of El Niños in the 21^{st} Century will become more like those seen during central Pacific El Niños than during the eastern Pacific El Niños.

5 Conclusions

In this study, we use a Gaussian Mixture Model (GMM) for clustering tropical Pacific SST anomalies to document the evolution of ENSO-related variance in observations, to evaluate the fidelity of ENSOs simulated by climate models in the 20^{th} Century, and to assess how ENSO changes in future projections. The clustering is performed on the first four PCs of monthly tropical Pacific SST anomalies. Before performing the clustering, the observed and projected long-term trends in the mean state is removed so that we can identify changes in the character of ENSO variability on interannual and shorter time scales. Compared to the more common k-means clustering (which is a particular case of GMM clustering; see Fabiano et al., 2021, and the references therein), GMM clustering differs in that the identification of the number of clusters K is probabilistic (see, Eq. 2 and 6), and there are fewer restrictions on the covariance matrices (see Eq. 5). Our choice of this particular GMM approach over k-means is motivated by two factors:

1. The number of clusters is fully data-driven without *a priori* knowledge about the clusters themselves.
2. More flexibility in the covariance matrices allows for very different shapes and sizes of clusters, which is beneficial given the diversity of ENSO patterns.

The GMM-derived clusters defined here are better able to represent the diversity of ENSO, including extreme or rare events.

Clustering provides more ways to characterize the observed ENSO-related variability than do traditional metrics, which typically assume that all phases of ENSO are represented by just one or two set SST anomaly patterns. Such metrics include a host of Niño.x indices (Rasmusson & Carpenter, 1982) and indices of E and C (K. Takahashi et al., 2011) based on the variability of the first two PCs of tropical Pacific SST anomalies. The GMM clustering approach used here instead identifies five "regimes" of SST anomalies' variability that are able to recover the (well-known) patterns of SST anomalies associated with La Niña and El Niño events and allows for a quantitative analysis of the frequency of occurrence and typical duration of each regime, as well as the likelihood of transition from one regime to another. Together these regimes present a more nuanced and demanding yardstick than traditional metrics of ENSO variability for measuring the fidelity of ENSOs simulated by the models in the modern climate and how they change in simulations of future climates.

The diversity of observed ENSO events is well captured by our GMM-based clusters of observed SST anomalies in the HadISST dataset (see Fig. 2) and in the JRA-55 (see supplementary Fig. S1). GMM clustering results in an assignment of each month's SST anomaly pattern to one of five possible regimes: including two El Niño regimes (a strong Eastern Pacific one, E-EN) and a more frequent moderate central Pacific one, C-EN); two La Niña regimes (a more frequent long lasting La Niña covering almost the whole Pacific domain, BW-LN, that includes the strongest La Niña events and a central La Niña, C-LN); and a Neutral pattern showing light to very tenuous SST anomalies.

The GMM clusters capture and quantify essential, well-known differences between El Niño and La Niña events, including differences in the magnitude and spatial patterns of SST anomalies, in the duration of cold and warm regimes, and in the seasonality of the regimes. The inadequacy of the traditional Niño.x indices and the E and C indices to distinguish either type of La Niña event from either type of El Niño event – and for the the C index to represent central Pacific El NIño events – is evident from the well-known differences in the location of the SST extrema in La Niña event and El Niño events (c.f. the top two panels of Fig. 2a) to the bottom two panels) and by the weighting of the EOFs that are required to describe these regimes/phases of ENSO (Fig. 2c).

Applying the GMM clustering to the detrended output from each of 16 CMIP6 models for the Historical period shows that the models, in general, reproduce the observed ENSO-related SST variability. Some notable discrepancies between observed and model regime statistics that are common to the models include central Pacific El Niño (C-EN) events that last too long and eastern Pacific El Niño (E-EN) that are too short and have too weak amplitudes compared to those observed (for a more complete list, see section 4.2).

By the end of the 21$^{\mathrm{st}}$ Century, ENSO-related variability in the CMIP6 models under a high emission scenario features several notable changes (see section 4.3) including an increase in the frequency and amplitude of central Pacific El Niño (C-EN) events in the majority of the models, and an increase in the frequency of eastern Pacific El Niño (E-EN) events (but with no consensus on whether the amplitude will change). As mentioned in section 4.3, it remains unclear to us to what extent the large bias in the amplitude of the central Pacific El Niño and La Niña events in the Historical simulations jeopardizes the projections of increasing variance in the ENSO-related SST anomalies in the eastern Pacific over the 21$^{\mathrm{st}}$ Century.

Although not pursued here, note that the ENSO phases, their probablistic duration and the transition frequencies between phases can also be used to make operational forecasts of the state of ENSO. Specifically, the value of the pseudo-PCs up to the time of the forecast initialization can be computed from SST and Eq. 2 can then be used to forecast the most probable ENSO regime that will develop (in particular, the type of El Niño or La Niña).

Understanding the effects of ENSO changes locally is important to anticipate future changes in weather conditions and the consequences for nature and society. Future studies could further investigate local implications of the different ENSO regimes. One way to do that would be to define regimes accounting for local-scale meteorological patterns (e.g., precipitation, wind speed) and large-scale patterns (*e.g.*, SST, see Vrac & Yiou, 2010) at the same time.

Open Research Section

The HadISST analysis SST product used in this study is accessible from their Web site at `https://www.metoffice.gov.uk/hadobs/hadisst/data/HadISST_sst.nc.gz`. The CMIP6 data used in the analysis were obtained from `https://esgf-node.llnl.gov/search/cmip6`.

Acknowledgments

All computations and figures are made using the R free software (R Core Team, 2020). PV and JT acknowledge funding from the Research Council of Norway (COLUMBIA-275268, CE2COAST-318477, and EASMO-322912).

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

References

Battisti, D. S., Vimont, D. J., & Kirtman, B. P. (2019). 100 Years of progress in understanding the dynamics of coupled atmosphere/ocean variability. *Meteorological Monographs*, *59*, 8.1–8.57. doi: 10.1175/AMSMONOGRAPHS-D-18-0025.1

Bentsen, M., Olivié, D. J. L., Seland, y., Toniazzo, T., Gjermundsen, A., Graff, L. S., ... Schulz, M. (2019). *NCC NorESM2-MM model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.8040` doi: 10.22033/ESGF/CMIP6.8040

Bertrand, A. (2020). *El Niño Southern Oscillation (ENSO) effects on fisheries and aquaculture* (No. 660).

Betts, R. A., Burton, C. A., Feely, R. A., Collins, M., Jones, C. D., & Wiltshire, A. J. (2020). ENSO and the Carbon Cycle. In (p. 453-470). In El Niño Southern Oscillation in a Changing Climate (eds M.J. McPhaden, A. Santoso and W. Cai). doi: 10.1002/9781119548164.ch20

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., ... Cheruy, F. (2018). *IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.5195` doi: 10.22033/ESGF/CMIP6.5195

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., ... Vuichard, N. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002010. doi: 10.1029/2019MS002010

Breton, F., Vrac, M., Yiou, P., Vaittinada Ayar, P., & Jézéquel, A. (2022). Seasonal circulation regimes in the north atlantic: Towards a new seasonality. *International Journal of Climatology*, *42*(11), 5848-5870. doi: https://doi.org/10.1002/joc.7565

Cai, W., Ng, B., Wang, G., Santoso, A., Wu, L., & Yang, K. (2022). Increased

ENSO sea surface temperature variability under four IPCC emission scenarios. *Nature Climate Change*, *12*(3), 228–231. doi: 10.1038/s41558-022-01282-z

Cai, W., Santoso, A., Collins, M., Dewitte, B., Karamperidou, C., Kug, J.-S., ... Zhong, W. (2021). Changing El Niño–Southern Oscillation in a warming climate. *Nature Reviews Earth & Environment*, *2*(9), 628–644. doi: 10.1038/s43017-021-00199-z

Cai, W., Wang, G., Dewitte, B., Wu, L., Santoso, A., Takahashi, K., ... McPhaden, M. J. (2018). Increased variability of eastern Pacific El Niño under greenhouse warming. *Nature*, *564*(7735), 201–206. doi: 10.1038/s41586-018-0776-9

Capotondi, A., Wittenberg, A. T., Kug, J.-S., Takahashi, K., & McPhaden, M. J. (2020). Enso diversity. In (pp. 65–86). In El Niño Southern Oscillation in a Changing Climate (eds M.J. McPhaden, A. Santoso and W. Cai). doi: 10.1002/9781119548164.ch4

Cassou, C. (2008). Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature*, *455*, 523–527. doi: 10.1038/nature07286

Chen, H.-C., & Jin, F.-F. (2021). Simulations of enso phase-locking in cmip5 and cmip6. *Journal of Climate*, *34*(12), 5135 - 5149. doi: 10.1175/JCLI-D-20-0874.1

Chen, X., & Wallace, J. M. (2015). ENSO-like variability: 1900–2013. *Journal of Climate*, *28*(24), 9623–9641. doi: 10.1175/JCLI-D-15-0322.1

Danabasoglu, G. (2019a). *NCAR CESM2 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.7627` doi: 10.22033/ESGF/CMIP6.7627

Danabasoglu, G. (2019b). *NCAR CESM2-WACCM model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.10071` doi: 10.22033/ESGF/CMIP6.10071

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., ... Strand, W. G. (2020). The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. doi: https://doi.org/10.1029/2019MS001916

Davey, M., Brookshaw, A., & Ineson, S. (2014). The probability of the impact of ENSO on precipitation and near-surface temperature. *Climate Risk Management*, *1*, 5–24. doi: 10.1016/j.crm.2013.12.002

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22. doi: 10.1111/j.2517-6161.1977.tb01600.x

Dommenget, D., Bayr, T., & Frauen, C. (2013). Analysis of the non-linearity in the pattern and time evolution of El Niño southern oscillation. *Climate dynamics*, *40*(11), 2825–2847. doi: 10.1007/s00382-012-1475-0

Dong, Y., Armour, K. C., Battisti, D. S., & Blanchard-Wrigglesworth, E. (2022). Two-way teleconnections between the southern ocean and the tropical pacific via a dynamic feedback. *Journal of Climate*, *35*(19), 6267 - 6282. doi: 10.1175/JCLI-D-22-0080.1

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., ... Zhao, M. (2020). The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2019MS002015. doi: 10.1029/2019MS002015

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. doi: 10.5194/gmd-9-1937-2016

Fabiano, F., Meccia, V. L., Davini, P., Ghinassi, P., & Corti, S. (2021). A regime view of future atmospheric circulation changes in northern mid-latitudes. *Weather and Climate Dynamics*, *2*(1), 163–180. doi: 10.5194/wcd-2-163-2021

Feely, R. A., Takahashi, T., Wanninkhof, R., McPhaden, M. J., Cosca, C. E., Sutherland, S. C., & Carr, M.-E. (2006). Decadal variability of the air-sea CO_2 fluxes in the equatorial Pacific Ocean. *Journal of Geophysical Research: Oceans*, *111*(C8). doi: 10.1029/2005JC003129

Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association*, *97*(458), 611–631. doi: 10.1198/016214502760047131

Fredriksen, H.-B., Berner, J., Subramanian, A. C., & Capotondi, A. (2020). How does el niño–southern oscillation change under global warming—a first look at cmip6. *Geophysical Research Letters*, *47*(22), e2020GL090640. doi: https://doi.org/10.1029/2020GL090640

Geng, T., Cai, W., Wu, L., Santoso, A., Wang, G., Jing, Z., ... others (2022). Emergence of changing Central-Pacific and Eastern-Pacific El Niño-Southern Oscillation in a warming climate. *Nature Communications*, *13*(1), 6616. doi: 10.1038/s41467-022-33930-5

Guo, H., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., ... Zhang, R. (2018). *NOAA-GFDL GFDL-CM4 model output historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.8594` doi: 10.22033/ESGF/CMIP6.8594

Hajima, T., Abe, M., Arakawa, O., Suzuki, T., Komuro, Y., Ogura, T., ... Tachiiri, K. (2019). *MIROC MIROC-ES2L model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.5602` doi: 10.22033/ESGF/CMIP6.5602

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., ... Kawamiya, M. (2020). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geoscientific Model Development*, *13*(5), 2197–2244. doi: 10.5194/gmd-13-2197-2020

Hamlet, A. F., & Lettenmaier, D. P. (1999). Columbia River streamflow forecasting based on ENSO and PDO climate signals. *Journal of water resources planning and management*, *125*(6), 333–341. doi: 10.1061/(ASCE)0733-9496(1999)125:6(333)

Harada, Y., Kamahori, H., Kobayashi, C., Endo, H., Kobayashi, S., Ota, Y., ... Takahashi, K. (2016). The JRA-55 Reanalysis: Representation of Atmospheric Circulation and Climate Variability. *Journal of the Meteorological Society of Japan. Ser. II*, *94*(3), 269-302. doi: 10.2151/jmsj.2016-015

Hastie, T., & Tibshirani, R. (1990). *Generalized Additive Models.* Chapman and Hall. Retrieved from `http://books.google.co.uk/books?id=qa29r1Ze1coC`

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., ... Zadeh, N. (2019). Structure and Performance of GFDL's CM4.0 Climate Model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691-3727. doi: 10.1029/2019MS001829

Hertig, E., & Jacobeit, J. (2014). Variability of weather regimes in the north atlantic-european area: past and future. *Atmospheric Science Letters*, *15*(4), 314–320. doi: 10.1002/asl2.505

Iizumi, T., Luo, J.-J., Challinor, A. J., Sakurai, G., Yokozawa, M., Sakuma, H., ... Yamagata, T. (2014). Impacts of El Niño - Southern Oscillation on the global yields of major crops. *Nature Communications*, *5*(1), 3712. doi: 10.1038/ncomms4712

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis,* *M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]* (Vol. In Press) [Book]. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. doi: 10.1017/9781009157896

Johnson, N. C. (2013). How many enso flavors can we distinguish? *Journal of Climate*, *26*(13), 4816 - 4827. doi: 10.1175/JCLI-D-12-00649.1

Jungclaus, J., Bittner, M., Wieners, K.-H., Wachsmann, F., Schupfner, M., Legutke, S., ... Roeckner, E. (2019). *MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.6594` doi: 10.22033/ESGF/CMIP6.6594

Karamperidou, C., Jin, F.-F., & Conroy, J. L. (2017). The importance of ENSO nonlinearities in tropical pacific response to external forcing. *Climate Dynamics*, *49*(7), 2695–2704. doi: 10.1007/s00382-016-3475-y

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... Takahashi, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5-48. doi: 10.2151/jmsj.2015-001

Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., ... Zhao, M. (2018). *NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.8597` doi: 10.22033/ESGF/CMIP6.8597

Law, R. M., Ziehn, T., Matear, R. J., Lenton, A., Chamberlain, M. A., Stevens, L. E., ... Vohralik, P. F. (2017). The carbon cycle in the Australian Community Climate and Earth System Simulator (ACCESS-ESM1) – Part 1: Model description and pre-industrial simulation. *Geoscientific Model Development*, *10*(7), 2567–2590. doi: 10.5194/gmd-10-2567-2017

Liu, S.-M., Chen, Y.-H., Rao, J., Cao, C., Li, S.-Y., Ma, M.-H., & Wang, Y.-B. (2019). Parallel Comparison of Major Sudden Stratospheric Warming Events in CESM1-WACCM and CESM2-WACCM. *Atmosphere*, *10*(11). doi: 10.3390/atmos10110679

Lovato, T., Peano, D., & Butenschön, M. (2021). *Cmcc cmcc-esm2 model output prepared for cmip6 cmip historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.13195` doi: 10.22033/ESGF/CMIP6.13195

Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., ... Navarra, A. (2022). Cmip6 simulations with the cmcc earth system model (cmcc-esm2). *Journal of Advances in Modeling Earth Systems*, *14*(3), e2021MS002814. doi: 10.1029/2021MS002814

Maher, N., Wills, R. C. J., DiNezio, P., Klavans, J., Milinski, S., Sanchez, S. C., ... Wu, X. (2023). The future of the El Niño–Southern Oscillation: using large ensembles to illuminate time-varying responses and inter-model differences. *Earth System Dynamics*, *14*(2), 413–431. doi: 10.5194/esd-14-413-2023

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., ... Roeckner, E. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO_2. *Journal of Advances in Modeling Earth Systems*, *11*(4), 998-1038. doi: 10.1029/2018MS001400

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., ... Marotzke, J. (2018). A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *Journal of Advances in Modeling Earth Systems*, *10*(7), 1383-1413. doi: 10.1029/2017MS001217

Naylor, R. L., Falcon, W. P., Rochberg, D., & Wada, N. (2001). Using El Niño/Southern Oscillation climate data to predict rice production in Indonesia. *Climatic Change*, *50*(3), 255–265. doi: 10.1023/A:1010662115348

Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F.-F., Wakata, Y., Yamagata, T., & Zebiak, S. E. (1998). ENSO theory. *Journal of Geophysical Research-Oceans*, *103*(C7), 14261-14290. doi: 10.1029/97JC03424\%7D

Nicholas, R. E., & Battisti, D. S. (2008). Drought recurrence and seasonal rainfall prediction in the Rio Yaqui Basin, Mexico. *Journal of Applied Meteorology and Climatology*, *47*(4), 991–1005. doi: 10.1175/2007JAMC1575.1

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... Sanderson, B. M. (2016). The scenario model intercomparison project (scenariomip) for cmip6. *Geoscientific Model Development*, *9*(9), 3461–3482. doi: 10.5194/gmd-9-3461-2016

Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London. A*, *185*, 71–110. doi: 10.1098/rsta.1894.0003

Peel, D., & McLachlan, G. J. (2000). Robust mixture modelling using the t distribution. *Statistics and computing*, *10*(4), 339–348. doi: 10.1023/A:1008981510081

Phillips, J. G., Cane, M. A., & Rosenzweig, C. (1998). ENSO, seasonal rainfall patterns and simulated maize yield variability in Zimbabwe. *Agricultural and Forest Meteorology*, *90*(1), 39–50. doi: 10.1016/S0168-1923(97)00095-6

Poveda, G., Jaramillo, A., Gil, M. M., Quiceno, N., & Mantilla, R. I. (2001). Seasonally in ENSO-related precipitation, river discharges, soil moisture, and vegetation index in Colombia. *Water Resources Research*, *37*(8), 2169–2178. doi: 10.1029/2000WR900395

R Core Team. (2020). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rasmusson, E. M., & Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Monthly Weather Review*, *110*(5), 354–384. doi: 10.1175/1520-0493(1982)110⟨0354:VITSST⟩2.0.CO;2

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., ... Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, *108*(D14). doi: 10.1029/2002JD002670

Rayner, P. J., Enting, I. G., Francey, R. J., & Langenfelds, R. (1999). Reconstructing the recent carbon cycle from atmospheric CO_2, δ13C and O_2/N_2 observations. *Tellus B: Chemical and Physical Meteorology*, *51*(2), 213-232. doi: 10.3402/tellusb.v51i2.16273

Ren, H.-L., Lu, B., Wan, J., Tian, B., & Zhang, P. (2018). Identification standard for enso events and its application to climate monitoring and prediction in china. *Journal of Meteorological Research*, *32*(6), 923–936. doi: 10.1007/s13351-018-8078-6

Rust, H. W., Vrac, M., Lengaigne, M., & Sultan, B. (2010). Quantifying differences in circulation patterns based on probabilistic models: Ipcc ar4 multimodel comparison for the north atlantic. *Journal of Climate*, *23*(24), 6573 - 6589. doi: https://doi.org/10.1175/2010JCLI3432.1

Sanchez-Gomez, E., Somot, S., & Déqué, M. (2009). Ability of an ensemble of regional climate models to reproduce weather regimes over europe-atlantic during the period 1961–2000. *Climate Dynamics*, *33*(5), 723–736. doi: 10.1007/s00382-008-0502-7

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461–464. doi: 10.1214/aos/1176344136

Scrucca, L., & Raftery, A. E. (2015). Improved initialisation of model-based clustering using gaussian hierarchical partitions. *Advances in data analysis and classi-*

fication, *9*(4), 447–460. doi: 10.1007/s11634-015-0220-z
Seager, R., Cane, M., Henderson, N., Lee, D.-E., Abernathey, R., & Zhang, H. (2019). Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nature Climate Change*, *9*(7), 517–522. doi: 10.1038/s41558-019-0505-x
Seferian, R. (2018). *CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.4068` doi: 10.22033/ESGF/CMIP6.4068
Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., … Madec, G. (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4182-4227. doi: 10.1029/2019MS001791
Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., … Schulz, M. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, *13*(12), 6165–6200. doi: 10.5194/gmd-13-6165-2020
Seland, Ø., Bentsen, M., Olivè, D. J. L., Toniazzo, T., Gjermundsen, A., Graff, L. S., … Schulz, M. (2019). *NCC NorESM2-LM model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.8036` doi: 10.22033/ESGF/CMIP6.8036
Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., … Zerroukat, M. (2019). UKESM1: Description and Evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513-4558. doi: 10.1029/2019MS001739
Solomon, A., & Newman, M. (2012). Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record. *Nature Climate Change*, *2*(9), 691–699. doi: 10.1038/nclimate1591
Su, J., Lian, T., Zhang, R., & Chen, D. (2018, jun). Monitoring the pendulum between el niño and la niña events. *Environmental Research Letters*, *13*(7), 074001. Retrieved from `https://dx.doi.org/10.1088/1748-9326/aac53f` doi: 10.1088/1748-9326/aac53f
Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., … Sigmond, M. (2019b). *CCCma CanESM5 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/10.22033/ESGF/CMIP6.3610` doi: 10.22033/ESGF/CMIP6.3610
Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., … Winter, B. (2019a). The Canadian Earth System Model version 5 (CanESM5.0.3). *Geoscientific Model Development*, *12*(11), 4823–4873. doi: 10.5194/gmd-12-4823-2019
Takahashi, K., Montecinos, A., Goubanova, K., & Dewitte, B. (2011). Enso regimes: Reinterpreting the canonical and modoki el niño. *Geophysical Research Letters*, *38*(10). doi: https://doi.org/10.1029/2011GL047364
Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., … de Baar, H. J. (2009). Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(8), 554-577. doi: https://doi.org/10.1016/j.dsr2.2008.12.009
Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., … Jones, C. (2019). *MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from `https://doi.org/`

10.22033/ESGF/CMIP6.6113 doi: 10.22033/ESGF/CMIP6.6113

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183-7192. doi: 10.1029/2000JD900719

Thomas, E., Vimont, D., Newman, M., Penland, C., & Martínez-Villalobos, C. (2018). The role of stochastic forcing in generating ENSO diversity. *Journal of Climate*, *31*, 9125-9150. doi: 10.1175/JCLI-D-17-0582.1

Tjiputra, J. F., Schwinger, J., Bentsen, M., Morée, A. L., Gao, S., Bethke, I., ... Schulz, M. (2020). Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2). *Geoscientific Model Development*, *13*(5), 2393–2431. doi: 10.5194/gmd-13-2393-2020

Trenberth, K. E., Branstator, G. W., Karoly, D., Kumar, A., Lau, N.-C., & Ropelewski, C. (1998). Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *Journal of Geophysical Research: Oceans*, *103*(C7), 14291–14324. doi: 10.1029/97JC01444

Vaittinada Ayar, P., Bopp, L., Christian, J. R., Ilyina, T., Krasting, J. P., Séférian, R., ... Tjiputra, J. (2022). Contrasting projections of the enso-driven co_2 flux variability in the equatorial pacific under high-warming scenario. *Earth System Dynamics*, *13*(3), 1097–1118. Retrieved from https://esd.copernicus.org/articles/13/1097/2022/ doi: 10.5194/esd-13-1097-2022

Vautard, R. (1990). Multiple weather regimes over the north atlantic: Analysis of precursors and successors. *Monthly weather review*, *118*(10), 2056–2081. doi: 10.1175/1520-0493(1990)118⟨2056:MWROTN⟩2.0.CO;2

Vimont, D. J., Newman, M., Battisti, D. S., & Shin, S.-I. (2022). The Role of Seasonality and the ENSO Mode in Central and East Pacific ENSO Growth and Evolution. *Journal of Climate*, *35*(11), 3195–3209. doi: 10.1175/JCLI-D-21-0599.1

Vimont, D. J., Wallace, J. M., & Battisti, D. S. (2003). The Seasonal Footprinting Mechanism in the Pacific: Implications for ENSO. *Journal of Climate*, *16*(16), 2668–2675. doi: 10.1175/1520-0442(2003)016⟨2668:TSFMIT⟩2.0.CO;2

Vrac, M., Vaittinada Ayar, P., & Yiou, P. (2014). Trends and variability of seasonal weather regimes. *International Journal of Climatology*, *34*(2), 472–480. Retrieved from http://dx.doi.org/10.1002/joc.3700 doi: 10.1002/joc.3700

Vrac, M., & Yiou, P. (2010). Weather regimes designed for local precipitation modeling: Application to the Mediterranean basin. *J. Geophys. Res.*, *115, D12103*. doi: 10.1029/2009JD012871

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., ... Roeckner, E. (2019). *MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.6595 doi: 10.22033/ESGF/CMIP6.6595

Wills, R. C. J., Dong, Y., Proistosecu, C., Armour, K. C., & Battisti, D. S. (2022). Systematic climate model biases in the large-scale patterns of recent sea-surface temperature and sea-level pressure change. *Geophysical Research Letters*, *49*(17), e2022GL100011. doi: https://doi.org/10.1029/2022GL100011

Yiou, P., & Nogaj, M. (2004). Extreme climatic events and weather regimes over the North Atlantic: When and where? *Geophys. Res. Lett.*, *31*, L07202. doi: 10.1029/2003GL019119

Yukimoto, S., Kawai, H., Koshiro, T., OSHIMA, N., YOSHIDA, K., URAKAWA, S., ... ISHII, M. (2019). The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *Journal of the Meteorological Society of Japan*, *97*(5), 931-965. doi: 10.2151/jmsj.2019-051

Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., ...

Adachi, Y. (2019). *MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.6842 doi: 10.22033/ESGF/CMIP6.6842

Zebiak, S. E., & Cane, M. A. (1987). A model El-Niño Southern Oscillation. *Monthly Weather Review*, *115*, 2262-2278. doi: 10.1175/1520-0493(1987)115⟨2262:AMENO⟩2.0.CO;2

Zhang, Y., Wallace, J. M., & Battisti, D. S. (1997). ENSO-like interdecadal variability: 1900-93. *Journal of Climate*, *10*, 1004-1020. doi: 10.1175/1520-0442(1997)010%3C1004:ELIV%3E2.0.CO;2%7D

Ziehn, T., Chamberlain, M., Lenton, A., Law, R., Bodman, R., Dix, M., . . . Druken, K. (2019). *CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP historical.* Earth System Grid Federation. Retrieved from https://doi.org/10.22033/ESGF/CMIP6.4272 doi: 10.22033/ESGF/CMIP6.4272