Identifying Quasi-periodic Variability Using Multivariate Empirical Mode Decomposition: a Case of the Tropical Pacific

Lina Boljka[1], Nour-Eddine Omrani[1], and Noel S. Keenlyside[1]

[1]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway

Correspondence: Lina Boljka (lina.boljka@uib.no), Nour-Eddine Omrani (noureddine.omrani@uib.no)

Abstract. ~~Tropical Pacific is home to climate variability on different timescales, including El Niño Southern Oscillation (ENSO) – one of the most prominent quasi-periodic modes of variability in the Earth's climate system. It is a coupled atmosphere-ocean mode of variability with a 2-8-year-timescale and oscillates between a warm (El Niño) and a cold (La Niña) phase~~A variety of statistical tools have been used in climate science to gain a better understanding of the climate system's variability on various temporal and spatial scales. However, ~~the dynamics of ENSO is complex, involving a variety of spatial and temporal scales as well as their interactions, which are not necessarily well understood. We~~ these tools are mostly linear, stationary or both. In this study, we use a recently developed nonlinear and nonstationary multivariate timeseries analysis tool – multivariate empirical mode decomposition (MEMD)~~ to revisit quasi-periodic variability within ENSO~~. MEMD is a powerful tool for objectively identifying (intrinsic) timescales of variability within a given ~~system. We apply it~~ spatio-temporal system without any timescale pre-selection. Additionally, a red noise significance test is developed to robustly extract quasi-periodic modes of variability. We apply these tools to reanalysis and observational data ~~as well as to climate model output (NorCPM1). Observational/reanalysis data reveal~~ of the tropical Pacific. This reveals a quasi-periodic variability in the tropical Pacific on timescales ∼~~2-4.5 years. We then test different conceptual oscillator models from literature and find that ∼2-4.5-year variability can be related to ENSO~~1.5-4.5 years, which is consistent with El Niño Southern Oscillation (ENSO) – one of the most prominent quasi-periodic modes of variability in the Earth'~~s~~s climate system. The approach successfully confirms the well known out-of-phase relationship of tropical Pacific mean thermocline depth with sea-surface-temperature in the eastern tropical Pacific (recharge-discharge ~~and simplified West-Pacific conceptual oscillator models. The latter has not been considered before and it occurs only on this timescale, however it is not necessarily well represented in NorCPM1. Additionally, the ∼2-4.5-year variability in ENSO can be 'predicted' up to ∼18 months ahead, while predicting the full ENSO amplitude remains challenging.~~ process). Furthermore, we find a co-variability between zonal wind stress in the western tropical Pacific and the tropical Pacific mean thermocline depth, which only occurs on the quasi-periodic timescale. MEMD coupled with a red noise test can therefore successfully extract (nonstationary) quasi-periodic variability from the spatio-temporal data, and could be used in the future for identifying potential (new) relationships between different variables in the climate system.

1 Introduction

25 ~~The dynamics of the tropical Pacific is typically characterised by ocean-atmosphere interaction, whereby atmospheric changes in winds can lead to changes in the distribution of warm and cold waters in the ocean that in turn impact the atmosphere. The variability in the tropical Pacific occurs on various timescales – from subseasonal to multidecadal (e.g., Maloney et al., 2008; Enfield and M. One of the most prominent features in the tropical Pacific is the El Niño Southern Oscillation (ENSO), which is a~~ The climate system is a highly complex system consisting of variability across many different timescales (e.g., Baede et al., 2001)

30 . Among these, quasi-periodic ~~phenomenon occurring on (inter-annual) timescales of 2-8 years (e.g., Philander, 1990; Wang and Fiedler, 20. ENSO events are characterized by warming sea surface temperatures (SSTs) during the development of El Niño (warm phase) and cooling of SSTs afterwards leading into La Niña (cold phase).~~ patterns of variability are important as they can be leveraged for medium-to-long range predictions (e.g., Mariotti et al., 2018, 2020; L'Heureux et al., 2020). Statistical tools have been particularly helpful in identifying quasi-periodic variability.

35 ~~ENSO is associated with major changes in the precipitation distribution in the region (increased precipitation typically follows warmer SSTs; e.g., Wang and Fiedler 2006; Dai and Wigley 2000) as well as in the distribution of warm and cold waters in the upper ocean (warmer SSTs are associated with a deeper thermocline in the ocean). This has strong relevance for the society (e.g., Wang and Fiedler, 2006; Santos, 2006; Lam et al., 2019) through modulation of food production (fishery, agriculture), weather-related disasters, and economy (e.g., Cashin et al., 2017; Guimarães Nobre et al., 2019).~~

40 ~~Because of its impacts, modelling and understanding of ENSO have been important topics for decades. The first models that captured crucial dynamics and were able to skilfully predict ENSO were developed in the 1980's (Quinn, 1974a, b; Cane et al., 1986; Zebia, and ENSO predictions have been continuously improving since (L'Heureux et al., 2020). To improve (long-term) predictions of ENSO and its impacts, it is important to study ENSO variability and associated physics. Indeed, ENSO can be described by several different processes and conceptual models that can yield~~ The following are typical statistical tools used for exploring the

45 patterns of variability on different temporal and/ ~~explain prediction skill on longer timescales.~~

~~First, one can use conceptual models based on red noise arguments, where relatively slow ocean variability is an 'integral' response to stochastic higher-frequency atmospheric (e.g., "weather") variability (following, e.g., Hasselmann, 1976; Frankignoul and Hass. While this is an integral part of the dynamical coupling between atmosphere and ocean, Clement et al. (2011) argued that the thermodynamic air-sea coupling (via thermally coupled Walker mode) also matters. Also, atmospheric 'wind-forcing' can help~~

50 ~~triggering ENSO events, however the timescale of this 'forcing' is still debated (e.g., Roulston and Neelin, 2000; Capotondi et al., 2018) and its efficiency in triggering ENSO may depend on the background state (e.g., Lopez et al., 2013; Lopez and Kirtman, 2014; Timmermann. Nonetheless, these processes give the system some persistence and thus can be related to prediction skill on longer timescales.~~

~~Second, as ENSO is a~~ or spatial scales (Broomhead et al., 1987; Hasselmann, 1988; Penland and Sardeshmukh, 1995; Ghil et al., 2002;
55 : Fourier transform (FFT), wavelet transform, principal component analysis (PCA), (multi-channel) singular spectrum analysis ((M)SSA), principal oscillation patterns (POPs), linear inverse model (LIM) or even nonlinear Laplacian spectral analysis. However, these methods do not necessarily have a multivariate extension, and/or are either stationary, linear or both. Additionally, some methods, such as (M)SSA, require specification of a window over which quasi-periodic ~~event (e.g., Wang et al., 2017) a part of its prediction skill likely comes from its internal quasi-periodic variability (e.g., Ghil and Jiang, 1998). Thus, ENSO's~~

60 ~~quasi-oscillatory dynamics has been studied extensively in the past, and the theories typically consist of positive (e.g., Bjerknes) and negative feedbacks between the atmosphere and ocean. The Bjerknes feedback (Bjerknes, 1969) refers to any decrease (increase) of trade winds that leads to reduced (enhanced) ocean upwelling (downwelling) and thus warming (cooling) in the eastern tropical Pacific leading to reduced (enhanced) zonal SST- and pressure-gradients, which in turn reinforce the initial increase (decrease) of the trade winds. This feedback alone would result in continuous warming (cooling) in the~~

65 ~~eastern tropical Pacific, therefore negative feedbacks are necessary for quasi-oscillatory behaviour in the eastern tropical Pacific (e.g., Wang et al., 2017).~~ signals are sought, making them somewhat subjective. These can be drawbacks when studying complex, inherently nonlinear, and nonstationary systems, such as the climate system.

~~To describe the interplay between the (Bjerknes) positive and negative feedbacks, several conceptual oscillator models have been proposed (e.g., Wang, 2018, see also section ??): (i)~~ Multivariate empirical mode decomposition (MEMD) addresses

70 these drawbacks as it is an analysis tool that is entirely data adaptive, and is designed to extract *nonlinear* and *nonstationary* signals. MEMD is a generalisation of the empirical mode decomposition (EMD; Huang et al., 1998) to multivariate datasets of more than two timeseries (Rehman and Mandic, 2010). EMD is a 1-D timeseries analysis tool that is based on Hilbert transform and takes advantage of the instantaneous frequency, allowing a 'local' extraction of modes of variability. Each mode that the EMD extracts consists of two elements: (1) typical timescale of the mode, i.e., average instantaneous frequency

75 of the ~~delayed oscillator (e.g., Suarez and Schopf, 1988; Battisti and Hirst, 1989) that included reflected Kelvin waves at the western ocean boundary as a negative feedback; (ii) the recharge-discharge oscillator (e.g., Jin, 1997a, b; Burgers et al., 2005) that included Sverdrup transport as a discharge/recharge processes (negative feedback); (iii) the advective-reflective oscillator (e.g., Picaut et al., 1997; Wang, 2001a) that included advection and reflection of Rossby waves at the eastern ocean boundary as a negative feedback~~ mode; ~~and (iv) the Western Pacific oscillator (e.g., Weisberg and Wang, 1997; Wang et al., 1999) that~~

80 ~~included interactions with the West-Pacific wind-forced Kelvin waves as a negative feedback. These oscillators can all be viewed as part of the unified oscillator (Wang, 2001a, 2018, see also section ??).~~ 2) the timeseries of the mode. Beyond this, MEMD extracts timescales *common* to *all* input timeseries (i.e., synchronises signals) and provides modes (patterns) of variability according to these timescales.

~~The above oscillators describe different processes that can lead to an ENSO event. These processes generally involve~~

85 ~~changes in the thermocline depth, surface wind stress, and SST anomalies. Note that seasonal cycle can strongly affect ENSO-related oscillations, since positive feedbacks are generally stronger in boreal autumn and negative feedbacks are stronger in boreal spring, i.e., it helps phase-locking of ENSO (e.g., Stein et al., 2010; Wengel et al., 2018). However, ENSO also has many flavours and can, for example, occur in the Eastern or Central Pacific (EP and CP ENSO, respectively; e.g., Kao and Yu 2009; Singh and Delcroix 2013; Zhang et al. 2019), and can occur on different timescales as well, e.g.: (i) a quasi-biennial~~

90 ~~(QB) ENSO with a ~2 year timescale and (ii) a low-frequency/quasi-quadrennial (LF/QQ) ENSO with a ~4-year timescale (e.g., Jiang et al., 1995; Allan, 2000; Kim et al., 2003; Keenlyside et al., 2007; Bejarano and Jin, 2008; Jajcay et al., 2018; Froyland et al., .~~ Despite their appeal, MEMD and EMD have hardly been used in climate research. EMD and its 1-D extension Ensemble EMD (EEMD; Wu and Huang 2009) have been used for smoothing, filtering, extracting trends, variability, and testing for red noise distribution of climate data (e.g., Duffy, 2004; Wu et al., 2007; Franzke, 2009; Lee and Ouarda, 2011; Qian et al., 2011; Franzke and Wool

95 . MEMD has only been used for an idealised analysis of atmosphere-ocean coupling strength (Alberti et al., 2021). Moreover, neither MEMD or EMD have been used for extracting quasi-periodic modes of climate variability.

Due to a multitude of factors impacting ENSO (described above) , its predictability remains challenging (e.g., Fedorov et al., 2003; L'Heu . This is due to unpredictable atmospheric (stochastic) forcing, different local and remote factors that impact it A major challenge in applying MEMD in climate analysis is that no statistical null hypothesis test for red-noise has been developed.

100 When applied to climate data, MEMD can reveal many modes that are consistent with red (or white) noise. In particular, sea surface temperature (SST) exhibit a red spectrum, because it represents the 'integral' response of the ocean to stochastic higher-frequency atmospheric (e.g., "weather", "white noise") variability (e.g., Hasselmann, 1976; Frankignoul and Hasselmann, 1977) . However, there are also patterns of quasi-periodic variability (e.g., inter-basin interactions) , and different spatial and temporal scales involved. All these processes impact the initial conditions, which ultimately determine the ENSO onset/magnitude and

105 thus its prediction. Presently, we are able to predict ENSO reasonably well 6 months ahead, and in some (rare) special cases even up to 2 years ahead (e.g., Chen et al., 2004; Park et al., 2018; L'Heureux et al., 2020). Some further improvements of ENSO predictability are also possible by utilising machine learning and neural network algorithms (e.g., Ham et al., 2019, 2021; Dijkstra et . However, to improve and extend ENSO's (general) prediction range, further physical understanding of this phenomenon is needed, and with it also model improvements (e.g., McPhaden, 2015; L'Heureux et al., 2020)repeating every 2-8 years) that

110 reflect more complex climate dynamics. To identify such quasi-periodic variability in the climate system, any analysis would first require a significance test that could distinguish this variability from the red noise. In other words, we seek spectral peaks (modes of variability) that pass the red-noise threshold (Gilman et al. 1963; Madden and Julian 1971; Bretherton et al. 1999, see also section 4, Appendix B2).

The main aim of this studyis to explore intrinsic In this study, we combine the MEMD method with a red noise test
115 (sections 3, 4) to robustly detect quasi-periodic variability of the tropical Pacific (and thus ENSO) on different timescales and thereby revisit the conceptual oscillator models of ENSO. This is done first by objectively splitting the tropical Pacific variability(using reanalysismodes of variability. Thus, MEMD becomes well-suited for analysing nonlinear and nonstationary climate data. It also has the advantage of *objectively* detecting intrinsic timescales *without pre-selecting* or/observational products) into different timescales (identify intrinsic variability). This then allows identification of potential oscillations and

120 their timescales, as well as the physical mechanisms that contribute to the tropical Pacific (and thus ENSO) variability on the quasi-periodic timescales (followingand filtering for a frequency band (as in, e.g., Jin 1997a; Wang 2001a. To achieve this, many different methods have been used inthe pastMSSA), and without any periodic signal or basis of functions specifications (as in, e.g., multi-channel singular spectrum analysis(MSSA), principal oscillation patterns (POPs) , linear inverse model (LIM) (Broomhead et al., 1987; Hasselmann, 1988; Penland and Sardeshmukh, 1995; Ghil et al., 2002). However, these methods can

125 be linear, stationary or both (Huang et al. 1998; Ghil et al. 2002; see also section 3), which can be a constraint for studying inherently nonlinear and non-stationary systems, such as the climate systemFourier transform and wavelet analysis). Since MEMD and (especially) its combination with a red noise test is a new tool in climate science, we test it on a known example of quasi-periodic variability. Namely, we analyse the tropical Pacific atmosphere-ocean variability to extract the El Niño Southern

Oscillation (ENSO). In the future, this tool may be used in other spatio-temporal applications, where quasi-periodic variability has not yet been identified.

~~Therefore, in this study we employ a recently developed nonlinear and nonstationary method for multivariate-timeseries filtering/analysis, called Multivariate Empirical Mode Decomposition (MEMD; Rehman and Mandic 2010) . This method is a multivariate extension of the Empirical Mode Decomposition (EMD; Huang et al. 1998) and can identify common timescales across different timeseries and multi-dimensional fields (for details see section 3 and Appendices A, B). This method objectively extracts intrinsic modes of variability in the tropical Pacific without a pre-selection of timescales, and can be used for identifying potential quasi-oscillatory behaviour, as well as the physics related to the tropical Pacific (and ENSO) variability on different timescales (more in section 5, and Appendix B2). This knowledge can then be used for constructing conceptual oscillator models, as well as for further understanding of model biases, ENSO prediction, ENSO teleconnections across scales, and other processes (see sections ?? and 7).~~ ENSO is a quasi-periodic phenomenon occurring on (interannual) timescales of 2-8 years (e.g., Philander, 1990; Wang and Fiedler, 2006; Timmermann et al., 2018). ENSO events are characterized by warming tropical Pacific SST during the development of El Niño (warm phase) and cooling of SST afterwards leading into La Niña (cold phase). These events are typically characterised by ocean-atmosphere interaction, whereby atmospheric changes in winds can lead to changes in the distribution of warm and cold waters in the ocean that in turn impact the atmosphere. ENSO exhibits significant non-linearities, with marked skewness and phase locking to the seasonal cycle (e.g., Stein et al., 2010; Dommenget et al., 2013) . It is also non-stationary (e.g., Crespo et al., 2022; Fedorov and Philander, 2000). This complex coupled dynamics is different from red noise and is therefore the focus of the present study.

The manuscript is structured as follows. Section ~~?? provides data and methods used in this study (Appendices A and B provide further detail son methodology~~ 2 provides description of data used; section 3 describes the EMD and MEMD implementation (Appendix A provides further details); section ~~5 explores the physical mechanisms (and conceptual oscillator models)~~ 4 provides description of the red noise test (Appendix B provides further details); sections 5, 6 identify modes of variability in the tropical Pacific and explore the physical mechanisms relevant on different timescales ~~with a focus on the timescales that are quasi-oscillatory; section ?? compares climate model data with reanalyses/observations and tests predictability of the quasi-oscillatory mode (see also Appendix C). .~~ Conclusions are given in section 7.

2 Data~~and Methods~~

2.1 ~~Data~~

~~In this study we~~ We focus on the ~~variability of ENSO using MEMD algorithm (described below). For this, we~~ intrinsic variability of the tropical Pacific and analyse monthly mean data of ~~four different variables~~ three different variables relevant for atmosphere-ocean exchange: sea surface temperature (SST) from HadISST observational dataset (Rayner et al., 2003), surface zonal wind stress (τ_x) and thermocline depth (i.e., the depth of the $20°C$ isotherm) both from SODA2 ocean-reanalysis dataset (Carton and Giese, 2008). These three variables are used to asses the ENSO dynamics as these quantities typically play a leading role in the onset and decay of ENSO events, and are typically used in the oscillator models that explain ENSO

dynamics ~~(see section ?? and, e.g., Wang, 2018)~~(see section 6 and, e.g., Wang, 2018). The data are analysed in the tropical Pacific (110°E - 65°W, 25°S - 25°N) over the period 1871-2010 for which all datasets are available. Note that surface wind stress and the ocean subsurface data reconstructions in the 19th and early 20th century are less reliable than in the late ~~20th~~ 20th century due to sparser data coverage, thus the results presented here are only as accurate as these reconstructions can be (Wittenberg, 2004; Crespo et al., 2022). We use the early record data here ~~as the (simple) prediction model used in section ?? requires as much data as possible, and the early record does not affect the overall results of this study. In fact, results of the recent decades are consistent with previous studies (see below)~~ to show the nonstationarity/nonlinearity of the tropical Pacific variability over the last ~140 years. The identification of the two main quasi-oscillatory modes (section 5) is also not affected by the inclusion of this data.

The MEMD analysis (~~described below~~section 3) is performed on all fields simultaneously with SST at the highest resolution (1° in latitude and longitude), whereas thermocline depth and τ_x (both 9° resolution in longitude, 5° resolution in latitude) have much lower resolution. This gives greater weight to SST data in the analysis, and less towards the other variables, such that the mode does not change significantly by adding other variables in the analysis (i.e., results below for the ~~SSTs~~ SST are similar whether we use ~~SSTs~~ SST alone or together with other fields). This ~~is because here~~ choice is made because we are primarily interested in the quasi-periodic behaviour in ~~SSTs, since SSTs~~ SST, since SST are used for defining ENSO (e.g., in Niño3 ~~region). SSTs~~ region—see Table 1). SST are also typically smoother than other fields~~ – this~~. This is especially true for wind stress, which is strongly affected by the "noisy" atmospheric variability. This ensures that modes that emerge from MEMD analysis are representative of quasi-periodic variability in ~~SSTs~~ SST (and thus ENSO), while the rest of the variables are "enslaved" to SST variability. ~~Thus, the other variables can help explaining the dynamics of ENSO (SST variability) on a specific timescale~~The other variables are added to the MEMD input data to describe the climate dynamics involved in the quasi-periodic SST variability, e.g., ENSO (section 6). Note that MEMD can be sensitive to input data, thus we must carefully consider the input data structure (relevant to a specific study).

~~In section ?? below we also analyse model data, namely the first ensemble member (i.e., r1i1p1f1; other ensemble members were qualitatively similar) of the NorCPM1 historical simulation (Bethke et al., 2019, 2021). As in observational data, we use model's τ_x, thermocline depth, and SSTs in the MEMD analysis. For consistency, we use the time period 1871-2010 in the model as well.~~

~~While the SSTs~~While the SST, thermocline depth, and τ_x play an important role in the ENSO dynamics, it is specific regions (see Table 1) that ~~are more relevant for the oscillator models (e.g., Wang, 2001a, 2018; Burgers et al., 2005). Namely,~~ have historically been analysed in more detail, for example, in conceptual oscillator models (e.g., Jin, 1997a; Wang, 2018). The timeseries in the specific regions are thus used to assess if MEMD modes on ENSO timescales are consistent with physics described by conceptual oscillator models (section 6). Additionally, analysing such timeseries helps a simpler visualisation of temporal evolution of different variables. Thus, we average τ_x over Niño4 and Niño5 regions separately, thermocline depth over Niño6 region (off-equatorial thermocline depth) and over the tropical Pacific (Pacific mean), and ~~SSTs~~ SST over Niño3 region (again, see Table 1; see also Fig. 3 in Wang et al. 1999). Note that regional averages (Table 1) are computed after the MEMD analysis is performed (for details see section 3).

Table 1. Tropical Pacific regions used for computing timeseries (see text for details). The right-most column lists the variables that are averaged over specified regions.

Region	Latitude Range	Longitude Range	variable
Niño3	5°S - 5°N	150°W - 90°W	SST
Niño4	5°S - 5°N	160°E - 150°W	zonal wind stress
Niño5	5°S - 5°N	120°E - 140°E	zonal wind stress
Niño6	8°N - 16°N	140°E - 160°E	thermocline depth
Pacific mean	5°S - 5°N	120°E - 90°W	thermocline depth

Before performing the analysis, we detrend the data and remove its seasonal cycle, which is done the following way (cf., de la Cámara et al., 2019). First, we calculate 30-year means centered on every 10th year for each individual month. This yields one value for each individual month every 10 years. Then ~~,~~ we interpolate between these values (of every 10th year) to obtain yearly time series, again for each individual month. This yields a smooth seasonal cycle that includes a trend and seasonal cycle for every month in the record. Detrended and deseasonalized data are then computed as the difference between the original monthly timeseries and the smooth trend/seasonal cycle. We do this at every grid point and for every variable separately.

This is done to avoid domination of the seasonal cycle or trend in the statistical analysis below, even though the ~~method presented below~~ MEMD can generally extract nonlinear trends by itself. Note that this means that we cannot assess the impact of long-term variability or seasonal cycle on ENSO variability in this study, but the latter may still be present indirectly ~~as it helps phase-locking of ENSO~~, as ENSO is phase-locked to the seasonal cycle (e.g., Stein et al., 2010; Wengel et al., 2018).

2.1   ~~(Multivariate) Empirical Mode Decomposition~~

3   (Multivariate) Empirical Mode Decomposition

~~To analyse the ENSO dynamics and relevant data, we use Multivariate Empirical Mode Decomposition (MEMD). This method was first introduced by Rehman and Mandic (2010) as a multivariate extension of the~~

3.1   MEMD description

We employ MEMD to *objectively* identify intrinsic modes of variability in *nonlinear* and *nonstationary* spatio-temporal data. To understand MEMD, it is easier to first consider the simpler implementation of the 1-D version, Empirical Mode Decomposition (EMD ~~; also called Hilbert-Huang transform). The EMD was introduced by Huang et al. (1998)as an alternative method for time-filtering of the one-dimensional~~), as outlined by Huang et al. (1998):

(i) Local minima and maxima of the input (1-D) timeseries ~~that is entirely data adaptive, nonlinear, and nonstationary, and thus more appropriate for analysing nonlinear and/or nonstationary data – the main advantage of this method. This~~

~~For simplicity, consider first the EMD method (i.e., 1-D version) as it has a relatively simple implementation (for details see Huang et~~
~~: (i) first, we identify local minima and maxima of the input (1-D) timeseries and create an envelope~~ see black solid line
in Fig. 1) are identified.

(ii) Envelopes are created by interpolating between ~~the~~ subsequent maxima (upper envelope, shown as grey dotted line in
Fig. 1) and between subsequent minima (lower envelope~~); (ii) then we obtain an average envelope from the~~, represented
by grey dashed line in Fig. 1).

(iii) An average envelope is obtained by taking the mean of the upper and lower ~~envelope and subtract it from the timeseries~~
~~data; (iii) the subtracted data(i.e.,~~ envelopes (depicted by the red solid line in Fig. 1).

(iv) The average envelope is subtracted from the original timeseries data.

(v) The subtracted data, i.e., the original data minus the average envelope~~) become~~, represent the first mode of variability
~~with the signal of the highest frequency~~ and typically correspond to the highest-frequency signal in the dataset~~, whereas~~
~~.~~ However, the average envelope can be ~~analysed further ; (iv) repeat steps~~ further analysed.

(vi) Steps (i)-~~(iii) for~~v) are repeated on the average envelope until only a trend ~~(residual) remains, i.e., until a condition of~~
~~at least 2~~ or residual component remains. This occurs when we can no longer find at least two extrema in the dataset~~can~~
~~no longer be satisfied . The~~, which is a condition that needs to be satisfied by EMD's modes of variability.

The resulting modes of variability (i.e., ~~timeseries)obtained through this process are called~~ the output of the EMD analysis),
represented by their respective timeseries, are known as intrinsic mode functions (IMFs)~~, and their instantaneous timescale~~.
The timescale of each IMF is characterised by the time-lapse between two subsequent extrema~~(and IMF's mean timescale is~~
~~an average over the instantaneous values). Each IMF also~~, and the mean timescale of an IMF represents the average over all
time-lapses within its timeseries.

Additionally, each IMF has to satisfy two criteria: (a) the number of extrema and the number of zero-crossings differs at
most by one; and (b) the mean value of the envelope of the IMF is zero. Note that the procedure from (i) to (iv) does not

Schematic of EMD envelope & signal extraction
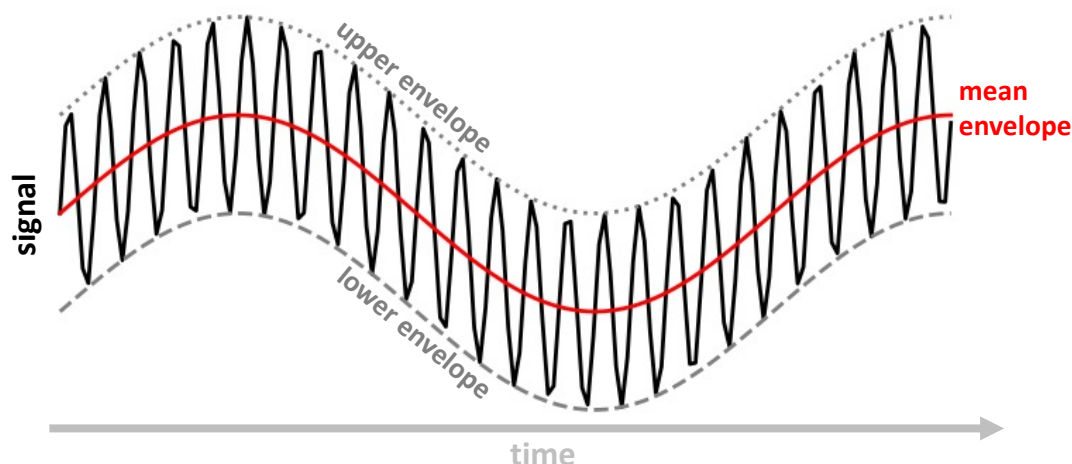
Figure 1. Schematic for obtaining average envelope during EMD process. Black line shows a simple input signal that is a sum of two sine waves (a low frequency and a high frequency wave). Grey dotted and grey dashed lines show upper and lower envelope, respectively. Red line shows average envelope, which represents the low frequency of the input signal. If we remove the average envelope from our input data, we obtain the high frequency signal (e.g., the first mode of EMD).

250   necessarily satisfy (a)-(b) immediately, thus an additional sifting process (typically iterative) is used that requires a stopping criteria to ensure physical meaning of the IMFs. The stopping criteria can be based on the standard deviation of each IMF, on the maximum number of iterations, etc., which set tolerance and confidence limits for the IMF (for details see Huang et al., 1998; Rilling et al., 2003; Huang et al., 2003).

~~The MEMD method~~ Ultimately, this process extracts modes of variability that consist of (1) typical timescale of the mode
255   and (2) its timeseries. The modes are automatically ordered from highest/shortest (mode 1) to lowest/longest (last mode) frequency/period. Note that modes that emerge are largely independent with only small correlations between them.

MEMD (Rehman and Mandic, 2010) is a generalisation of the EMD to multivariate datasets of more than two timeseries (for bivariate and trivariate data separate methods exist; Rilling et al. 2007; Rehman and Mandic 2010). The method solves a similar problem as in (i)-~~(iv) but~~ vi) but the mean envelope is computed as (Rehman and Mandic, 2010) "an integral of
260   all the envelopes along multiple directions in an N-dimensional space" (i.e., on an N-Sphere~~, which~~). This is much more complex, but the basic idea remains similar to the 1-D method, and the method retains similar stopping criteria for the sifting process. For further details on and visualisations of the method the reader is referred to ~~Rehman and Mandic (2010).~~ Rehman and Mandic (2010); Alberti et al. (2021).

The MEMD ultimately extracts timescales *common* to *all* input timeseries (i.e., synchronises signals; Rehman and Mandic 2010
265   ) and provides multivariate IMFs (i.e., the outputs of MEMD method) describing those timescales. The timescales of IMFs are

then consistent across the input timeseries — to visualise this, see supplementary Figs. S6-S10, Tables S1-S2, and section S.3 for an idealised example.

As with all statistical methods, it is important to be aware of the drawbacks associated with the (M)EMD (e.g., Stallone et al., 2020). Similar to other timeseries-filtering methods, (M)EMD can encounter issues at the edges of the timeseries, which can result
270 in "travelling waves" and thus unrealistic peaks in the timeseries (e.g., Stallone et al., 2020). Another common challenge with (M)EMD is the mixing of modes (see below) where a single genuine mode may be split into multiple modes if inappropriate parameters are chosen (e.g., Huang et al., 1999, 2003; Stallone et al., 2020). However, this issue can also arise when there is no clear timescale-separation. To address these challenges, it is crucial to test the physical relevance of the modes of variability identified using (M)EMD and to ensure convergence and stability of the modes through different parameter sweeps that are
275 related to the stopping criteria (see section 3.2). Note that different parameter sweeps may be relevant for different applications.

3.2   MEMD parameter sweep

The code for the ~~method~~ MEMD is freely available on Github (https://github.com/mariogrune/MEMD-Python-; similarly for the EMD discussed above: https://github.com/laszukdawid/PyEMD), and the user ultimately only decides about the stopping
280 criteria~~, which are here set to their~~. These are set to the "fix_h" parameter, following Huang et al. (2003) who suggest that limiting iterations yields better-behaved IMFs than other stopping criteria. ~~Here, we~~ We limit the number of iterations to 15 (parameter "n_iter" is 15), though other values were tested and a range for "n_iter" around ~~10-30~~ 5-30 yielded similar results, suggesting some convergence for the *significant* modes of variability ~~.~~ (see below).

Note that at higher/lower frequencies we find mode-mixing in our MEMD analysis where timescales are not clear (here,
285 this occurs on timescales shorter than about 8 months ~~), especially~~ and longer than about 700 months), and also with larger number of iterations. These modes are not detected as different from red-noise (see sections 4, 5 for more details). However, the significant modes of variability on interannual timescales that are of interest here are largely unaffected by this ~~.~~ (section 5).

~~The MEMD ultimately extracts timescales *common* to *all* input timeseries (i.e., synchronises signals) and provides IMFs~~
290 ~~according to these timescales, which are consistent across the input timeseries (to visualise this, see below and supplementary Figs. S7-S11, Tables S1-S2, and section S.3 for an idealised example). This is an additional property of MEMD relative to its 1-D version (EMD). Note that the use of IMFs obtained via MEMD analysis will depend on application—here, we analyse principal components of a 3-D field to extract quasi-oscillatory modes of variability (the whole procedure is described below and in Appendices A, B). In this study, we use the MEMD method because of its nonlinear and nonstationary properties,~~
295 ~~and its property to detect quasi-periodic signals without pre-selecting (filtering for) a frequency band (a constraint with, e.g., singular-spectrum analysis), and without any periodic signal or basis of functions specifications (a problem with, e.g., Fourier transform and wavelet analysis).~~

3.3 Obtaining modes of variability via MEMD

Even though EMD and its 1-D extension Ensemble EMD (EEMD; Wu and Huang 2009) have been applied in climate science in various applications, e.g., for smoothing, filtering, extracting trends, variability, and testing for red noise distribution of climate data (e.g., Duffy, 2004; Wu et al., 2007; Franzke, 2009; Lee and Ouarda, 2011; Qian et al., 2011; Franzke and Woollings, 2011; Franzke, , it has not been explicitly used for extracting quasi-periodic signals. Moreover, the MEMD has only been applied to an analysis of the atmosphere-ocean coupling strength (Alberti et al., 2021) in climate science, which was done in a more idealised setting from the present study – used to analyse ENSO dynamics. Therefore, (in addition to analysis of ENSO dynamics) we also perform extensive analysis of the method itself and compare it to the basic band-pass filtering (5th order Butterworth filter) and to Fourier transform analysis. This shows that MEMD's results are consistent with other methods, but can also extract modes of variability in a more objective way (see below and Appendices A, B).

As with all statistical methods, also (M)EMD has drawbacks that we must be aware of (e.g., Stallone et al., 2020). Because it is a statistical method we must always test if the modes we find using (M)EMD are physical, and also check for convergence (stability) of the modes using different parameter sweeps (as mentioned above). Much like other timeseries-filtering methods, also (M)EMD has issues at the edges of the timeseries, which can lead to "travelling waves" and thus unrealistic peaks in the timeseries (e.g., Stallone et al., 2020). There is also a common issue of mixing modes (as mentioned above) where one (real) mode is split into two or more modes if we choose "wrong" parameters (e.g., Huang et al., 1999, 2003; Stallone et al., 2020), but this is also an issue when there is no clear timescale-separation. We have tested different sweeps of parameters and have ultimately decided on the ones specified above, as they suggested some stability and the results were realistic and comparable to other methods (note that for other applications different parameter sweeps may yield better results). Therefore, we continue from hereon with the description of implementation of the MEMD method and later we use it to understand the ENSO dynamics on different timescales.

Obtaining modes of variability via MEMD

As mentioned in Sect. ??, we use mentioned in section 2, we use 3-D data relevant for ENSO dynamics, i.e., $\mathbf{A}(t,y,x)$ (with t as time $\mathbf{A}(L_t, L_y, L_x)$ (with L_t length of time dimension, L_y length of latitude dimension, L_x length of longitude dimension, y as latitude , x as longitude , $\mathbf{A}$ as the selected variable(s)/field(s); bold letters represent matrices) , two- or three-dimensional arrays) as basis for the MEMD analysis. We first , which is done the following way.

(i) First, we remove the smooth trend/seasonal cycle (described above; see also de la Cámara et al., 2019) from $\mathbf{A}$ to get a temporal anomaly, $\mathbf{A}'$. Since here

(ii) We divide data by their standard deviations (σ; done separately for each variable and at each grid point).

(iii) Since we use more than one variable (i.e., SSTsSST, surface wind stress, thermocline depth) in the analysis, we concatenate (denoted with $\oplus$) the different variables in their spatial dimensions, i.e., we get $\mathbf{A}'(t, y = y_1 + y_2 + ..., x = x_1 + x_2 + ...)$ $\mathbf{A}'(L_t, L_y = L_{y1} \oplus L_{y2} \oplus ..., L_x = L_{x1} \oplus L_{x2} \oplus ...)$ (with subscripts 1,2,... representing spatial dimensions of the differ-

330   ent variables). ~~Note that we also divide data by their standard deviations (σ; done separately for each variable and at each grid point) before concatenating them. Then~~

(iv) Then, we reduce the dimensionality by computing spatial patterns (empirical orthogonal functions, EOFs) and their timeseries (principal components, PCs) via singular value decomposition (SVD). ~~We~~

(v) Finally, we only retain the first 20 PCs that explain the majority of the variance in the field $\mathbf{A}'$. The PCs are then
335   ~~ultimately~~ used as input data for the MEMD algorithm ~~as described in Appendix A in detail. The modes that emerge from the MEMD analysis can become spatio-temporal modes of variability (IMFs) by reconstructing the fields from PCs and EOFs (~~(for further details of the whole procedure see Appendix A)~~for each IMF. Here, we identify 22~~.

MEMD analysis identifies 21 IMFs that are ordered by frequency from the highest (IMF1) to ~~lowest (IMF22~~the lowest (IMF21) with the last ~~22nd~~ 21st mode typically representing a trend, which in our case was already removed (see above).
340   Namely, we find 21 potential intrinsic timescales within the tropical Pacific, i.e., common to all input PCs. This means that we obtain 21 IMFs for each PC-timeseries, i.e., $\mathrm{PC}_m(L_t) = \sum_s \mathrm{IMF}_s(\mathrm{PC}_m(L_t))$ with s IMF-number and m PC-number, where s^{th} IMF of each PC_m has the same timescale (see Table S2). Since we initially obtained PCs via SVD analysis and we thus have corresponding EOF patterns, we can then reconstruct spatio-temporal patterns of variability for each field (i.e., SST, wind stress and thermocline depth) from PCs and EOFs. As the interest here is in the variability on different timescales, i.e., for each
345   (s^{th}) IMF separately, we can perform this reconstruction for each IMF across all 20 PCs/EOFs, i.e.,

$$\mathrm{IMF}_s^{spatial}(L_t, L_y, L_x) \simeq \sum_m \mathrm{IMF}_s(\mathrm{PC}_m(L_t))\mathrm{EOF}_m(L_x, L_y). \tag{1}$$

Please recall that input data for MEMD analysis were divided by σ, thus for variability of a field in its original units we need to multiply spatio-temporal IMFs by σ.

To compute an index, such as eastern Pacific SST (Niño3), we can average over a ~~x-y~~latitude-longitude region (Table 1)
350   from spatio-temporal IMFs (Eq. 1) to obtain timeseries of, e.g., eastern Pacific SST (Niño3) for each IMF separately ~~(~~ see supplementary Figs. ~~S7-S11 for timeseries~~ S6-S10 for timeseries of different variables from Table 1, and Table S1 for their timescales~~)~~. This yields ~~the IMFs of an index~~ an index for each IMF (e.g., $\mathrm{IMF}_s(\mathrm{SST}$ (Niño3)~~)~~) with s IMF-number) corresponding to an equivalent index computed from input data ($\mathbf{A}'$). The latter is approximately the same as the sum of indices computed from all IMFs, e.g., $\mathrm{SST}(\mathrm{Niño3})(L_t) \simeq \sum_s \mathrm{IMF}_s(\mathrm{SST}(\mathrm{Niño3})(L_t))$.

355   ## 4   Statistical significance test for climate

Once we have computed the IMFs~~from the input data~~, we need to test if they are statistically significant. The importance of each IMF can be assessed by computing variance explained of each IMF relative to the input field (e.g., retaining those IMFs that explain more than 0.1% variance) or through other significance tests (e.g., white noise test; Appendix B1; Wu and Huang, 2004). ~~As the interest in this study is in potential oscillatory behaviour in the ENSO region, we seek IMFs that pass~~ We develop

360 a test for variability that can be distinguished from red-noise, as appropriate for studying climate variability (see Introduction).

The red noise test can be performed in different ways (see also section 5.1): (i) we can choose an index of interest and perform a red noise ~~threshold, since the monthly sea surface temperature data typically follow a red noise distribution. The red noise test is described in Appendix B2, and the results are shown in Fig. 2 (incl. a white noise test for comparison). The~~
365 ~~timescales of significant modes, identified via~~ test on its IMFs (see below; Fig. 2); (ii) we can use PC1's IMFs (more objective; Fig. S2); (iii) we can test for quasi-periodic variability at each grid point (section 5.1, Fig. 5); or other. Thus, significance of modes is only relevant for the index/region used in the red noise test calculation.

Here, we first test for potential quasi-periodic variability using SST timeseries that are relevant for (eastern Pacific) ENSO variabilty, i.e., eastern Pacific SST (Niño3) from input data and corresponding eastern Pacific SST (Niño3) from spatio-temporal
370 IMFs (as described in section 3; $\text{IMF}_s(\text{SST (Niño3)})$). Then, we compute power spectrum for each timescale/period (i.e., each IMF) by obtaining average squared amplitude of each (s^{th}) IMF (E_s), i.e., $E_s = (\sum_j [\text{IMF}_s(j)]^2)/L_t$ (with L_t length of the timeseries, j each timestep). Average timescale (T_s) of each IMF is computed from instantaneous frequency using Hilbert transform (see text around Eq. B4 in Appendix B1), which yields similar results to computing time lapse between two extrema in the timeseries. E_s is then plotted against T_s (here we use their logarithmic values) to yield the power spectrum plot (blue
375 dots in Fig. 2).

The shape of the red noise fit (red solid line in Fig. 2~~, fall well within the significant timescales identified via a typical power spectrum (via Fourier transform) analysis~~), or rather red noise spectrum for every s^{th} IMF, can be computed from the lag-one autocorrelation of the input data (e.g., eastern Pacific SST (Niño3)) as (cf., Gilman et al. 1963; Kolotkov, D. Y. et al. 2016; for full derivation and further discussion see Appendix B)

380
$$E_s^{red}(\nu_s) = \frac{1-r^2}{1-2r\cos 2\pi\nu_s + r^2}(\nu_s\alpha_s - \nu_s/\beta_s). \tag{2}$$

E_s^{red} is a theoretical estimate of the (mean) energy of the red noise of sth IMF, $\nu = 1/t$ is frequency, t is time, r is lag-1 auto-correlation of input data of eastern Pacific SST (Niño3), $\beta_s = \sqrt{\nu_s/\nu_{s+1}}$ and $\alpha_s = \sqrt{\nu_{s-1}/\nu_s}$, and subscript s represents the sth IMF of frequency ν_s (ordered from highest to lowest frequency).

E_s^{red} is then scaled so that the variance of the red noise fit is identical to the variance of the spectra computed from the
385 spatio-temporal IMFs (E_s; Eq. (B9); see also, e.g., Madden and Julian 1971; Bretherton et al. 1999). The 95% confidence curve (red dashed line in Fig. ~~B3 (for more~~ 2) is computed by multiplying the red noise fit by χ^2 statistic for a one-sided p-value of 0.05. The number of degrees of freedom for s^{th} IMF is computed as $\text{DoF}_s = E_s L_t(1-r^2)/(1+r^2)$ (for further details see Appendix B2~~). -; see also Bretherton et al. 1999; Wu and Huang 2004; Kolotkov, D. Y. et al. 2016)~~. Note that the same procedure (as described above) can be applied to any other timeseries/IMFs.
390 ~~The results from Appendix B2 suggest that there are one to two significant modes of variability in the ENSO region, namely the 12~~
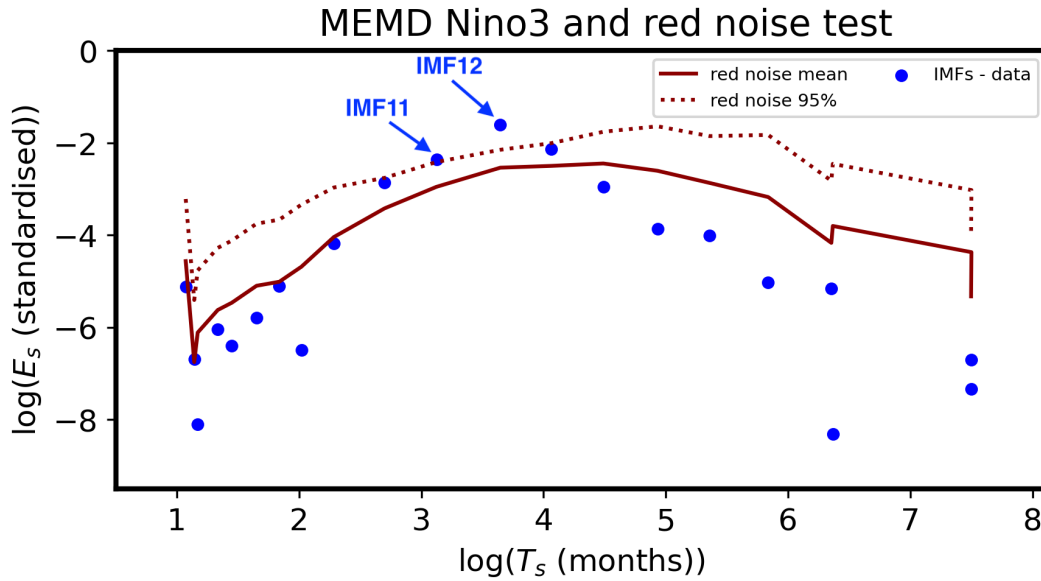
Figure 2. Red noise significance test for spatio-temporal IMFs (from MEMD) via SST averaged over the Niño3 region (for details see text and Appendices A,B). Each blue dot represents average squared amplitude (E_s; Eq. B1) and average period (T_s; in months; for periods of IMFs see Table S1) of each IMF that we have identified within our timeseries/data. Average period is computed from instantaneous frequency obtained via Hilbert transform (see text under Eq. B4). For visualisation purposes we obtain a natural logarithm of both average squared amplitude ($\log_e(E_s)$; y-axis) and average period ($\log_e(T_s)$; x-axis) of each IMF and plot them as a scatter plot (blue dots). Note that the logarithms of periods ($\log_e(T_s)$) are ordered from shortest period (highest frequency; IMF1; left-most blue dot) to longest period (trend; IMF21; right-most blue dot). Red solid line represents the theoretical red spectrum fit (Eqs. 2, B7-B9), red dotted line represents the one-tailed 95th percentile confidence level (via χ^2-test). For further descriptions of the figure see text.

5 Tropical Pacific modes of variability

As MEMD in conjunction with a red-noise test has not been applied in climate science before, we perform extensive analysis of the method itself (in addition to analysis of ENSO dynamics - see below). Then, we compare it to the basic band-pass filtering (5th ~~and 13th mode (IMF12, IMF13) . Recall that any potential oscillatory timescales longer than 30-years were removed via detrending mentioned above, and are thus not considered here. The~~ order Butterworth filter) and to Fourier transform analysis to ensure consistency with other methods. Please recall that while other spectral methods often require prior knowledge about the spatial/temporal patterns of interest in order to construct appropriate indices, the MEMD method allows for the objective extraction of significant patterns and modes of variability from data without pre-existing knowledge (see also below and Appendices A, B).

5.1 Significant modes of variability

We identify two significant modes of variability ~~(i.e., IMF12 and IMF13) in the ENSO region have~~ in the eastern Pacific: the 11th IMF (IMF11) and 12th IMF (IMF12) (pointed out with blue arrows in Fig. 2) with average timescales of ~~about 22 and 34 months (i.e.,~~ ~23 and ~39 months (~2-3 years; see also Table S1), respectively~~, and are therefore~~ . Their timescales fall well within the ~~the~~ typical ENSO timescale range (2-8 years)~~. Such~~, as can be identified via a typical power spectrum analysis (Appendix B2, Fig. B3). These timescales (with their uncertainty ranges; see below) are consistent with ~~both QBand~~ previously identified quasi-biennial (QB, ~2 years) and low-frequency/quasi-quadrennial (LF/QQ~~ENSO (e.g., Jiang et al., 1995; Jajcay et al., 2018) . We find that the~~, ~4 years) ENSO modes (e.g., Jiang et al., 1995; Allan, 2000; Kim et al., 2003; Keenlyside et al., 2007; Bejarano and Jin~~.~~~

The quasi-periodic mode of variability with ~~34~~ 39 months average period (~~IMF13) is much clearer (than~~ IMF12) is more statistically significant (than IMF11). This can be seen in different ways. First, considering the Niño3 index (i.e., eastern Pacific SST in Niño3 region), IMF11 (blue dot) lies only slightly above the 95% red noise threshold (red dashed line), whereas IMF12 ~~), which is consistent with 'more periodic'LF/QQ ENSO (Jiang et al., 1995). Spatial pattern of the significant mode with 34 months timescale (IMF13) is provided in Fig~~lies well above the threshold (Fig. 2). Second, considering the PC1's IMFs (i.e., before we do spatio-temporal reconstruction of IMF-data; Fig. S2), only IMF12 is found to be significant, despite the generally good agreement between the Niño3 index and PC1 of the tropical Pacific SST (e.g., Ashok et al., 2007). Finally, we can perform a red noise test on each grid point of the tropical Pacific (similarly as for specific timeseries shown above) and plot spatial structure of SST together with the significance test (section 6, Figs. 5, ~~which shows that it indeed resembles ENSO spatial pattern, and is further discussed in section 5. Note that the spatial pattern of the 22 months mode (S3). This~~ shows that none of the grid points of spatio-temporal IMF11 pass the red noise threshold (Fig. S3), but many grid points of the spatio-temporal IMF12 ~~; see Fig. S2 in supplement) looks qualitatively similar, but it is only weakly significant.~~ are significant (not-shaded in grey in Fig. 5).

Nevertheless, even though IMF11 is marginally statistically significant, the MEMD analysis suggests that there may be two quasi-periodic modes of interannual SST variability in the tropical Pacific region. This agrees with previous results that have identified two statistically significant modes of variability with specific frequencies and distinguished by their dynamics. Note, we also find that the two modes of variability are well separated (in terms of timescale) from the other modes and from each other, as shown by Fig. 2. This ensures no mode-mixing on the quasi-periodic timescales.

On longer timescales, we do not find any behaviour that would be discernible from red noise, suggesting that the lower-frequency range of ENSO (timescales longer than ~4.5 years) ~~(Allan, 2000; Jajcay et al., 2018, e.g.,)~~ (e.g., Allan, 2000; Jajcay et al., 2018) is better represented by red noise~~and likely less predictable than QB or (especially) LF/QQ ENSO.~~

~~Even though the results from Appendix B2 provide 'significant' IMFs, MEMD remains a statistical method and thus further analysis is required to identify whether its modes have any physical meaning. If IMFs are physical, any inferred information could be used for improving our climate models, long-term prediction, understanding the teleconnections, the underlying physics/variability of the field we are interested in (e.g., ENSO) , etc~~ Recall, however, that any potential oscillatory behaviour on timescales longer than 30-years was removed via detrending. Similarly, we do not find any quasi-periodic modes of variability on shorter timescales. Also, on these longer (> 700 months) and shorter (< 8 months) timescales we find some

mode mixing. This is ~~addressed in the following sections (5, ??) through the analysis of ENSO dynamics and its conceptual oscillator models in observations and a climate model.~~ likely a consequence of inputting several different variables (i.e., SST, τ_x, thermocline depth) into MEMD algorithm, where each variable can have different timescales represented. For example, wind stress can be much noisier than SST (or different parts of the tropical Pacific have different variabilities) and can thus lead to identification of several high frequency modes of variability, resulting in mode-mixing (see somewhat overlapping blue dots in Fig. 2).

6 ~~Variability and dynamics of ENSO~~

5.1 Timeseries of significant modes of variability

~~In the previous section~~ Above we have established that the Niño3 index in the tropical Pacific exhibits two quasi-periodic modes of variability with average periods ∼2-3 years (via MEMD analysis~~; note that from hereon we only consider the results from the MEMD analysis). The~~). We now analyse the associated timeseries of the two IMFs (~~IMF12 and IMF13) are shown in Fig. 3 (red dotted and dashed lines respectively), along with their sum (black dashed line) and~~IMF11 and IMF12) and compare them with a band pass (~~17-52.5~~16-53 months) filtered Niño3 index (~~black solid line).~~

Fig. 3~~shows that the~~).

The period/frequency of the two modes is not constant (i.e., varies in time~~; see below), thus~~). Thus, we also specify a range of periods/frequencies for the two modes. The mean periods of ~~IMF12,13~~IMF11,12 with their 'uncertainty' ranges (in square brackets) are: ~~22~~23 [~~17, 32~~16, 33] (~~IMF12~~IMF11), and ~~34~~39 [~~25, 52.5~~29, 53] (~~IMF13~~IMF12) months. These ranges are defined based on the 6.7^{th} and 93.3^{rd} percentiles of ~~IMF12,IMF13~~IMF11,IMF12's instantaneous period/frequency values, which roughly correspond to $\pm 1.5\sigma$ (instantaneous periods). Fig. S1 (supplement) shows power spectra of ~~IMF12 and IMF13~~IMF11 and IMF12, which visualise the range of periods the two modes have. ~~Note~~ Recall that instantaneous period/frequency can be obtained via Hilbert transform (see Eq. (B4) in Appendix B1). This range was chosen as it ~~yielded the best results, but other (~~captures most of the variability in a given mode. Additionally, this range yields a good agreement with other methods (e.g., band-pass filtering), although other (reasonable) percentile ranges give qualitatively similar results. We then use these period/frequency ranges to perform a ~~band pass~~band-pass filter (via 5th order Butterworth filter) of the Niño3 index that is consistent with the individual modes (e.g., ~~25-52.5~~29-53 months band of ~~IMF13~~IMF12 is used to construct band-pass filtered Niño3 index composites in Fig. 6b; more below)~~, as well as~~. We also use a band-pass filter of Niño3 index consistent with the sum of the two modes where the band-pass range encompasses periods of both significant modes, i.e., ~~17-52.5~~16-53 months (e.g., used to construct black solid line in Fig. 3).

Timeseries of the sum of IMFs (black dashed line) and band passed index (black solid line) ~~in Fig. 3 largely agree ,~~largely agree (i.e., their correlation is ~~0.96. However, also~~0.95; Fig. 3). Also, the individual modes show very good agreement with the band-passed index with correlations of ~~0.69 (IMF12~~0.68 (IMF11) and 0.83 (~~IMF13~~IMF12), which can be increased further if we consider only the specific IMF's timescale range (specified above) when band-passing the Niño3 index. This merely confirms that MEMD results are consistent with other filtering methods.
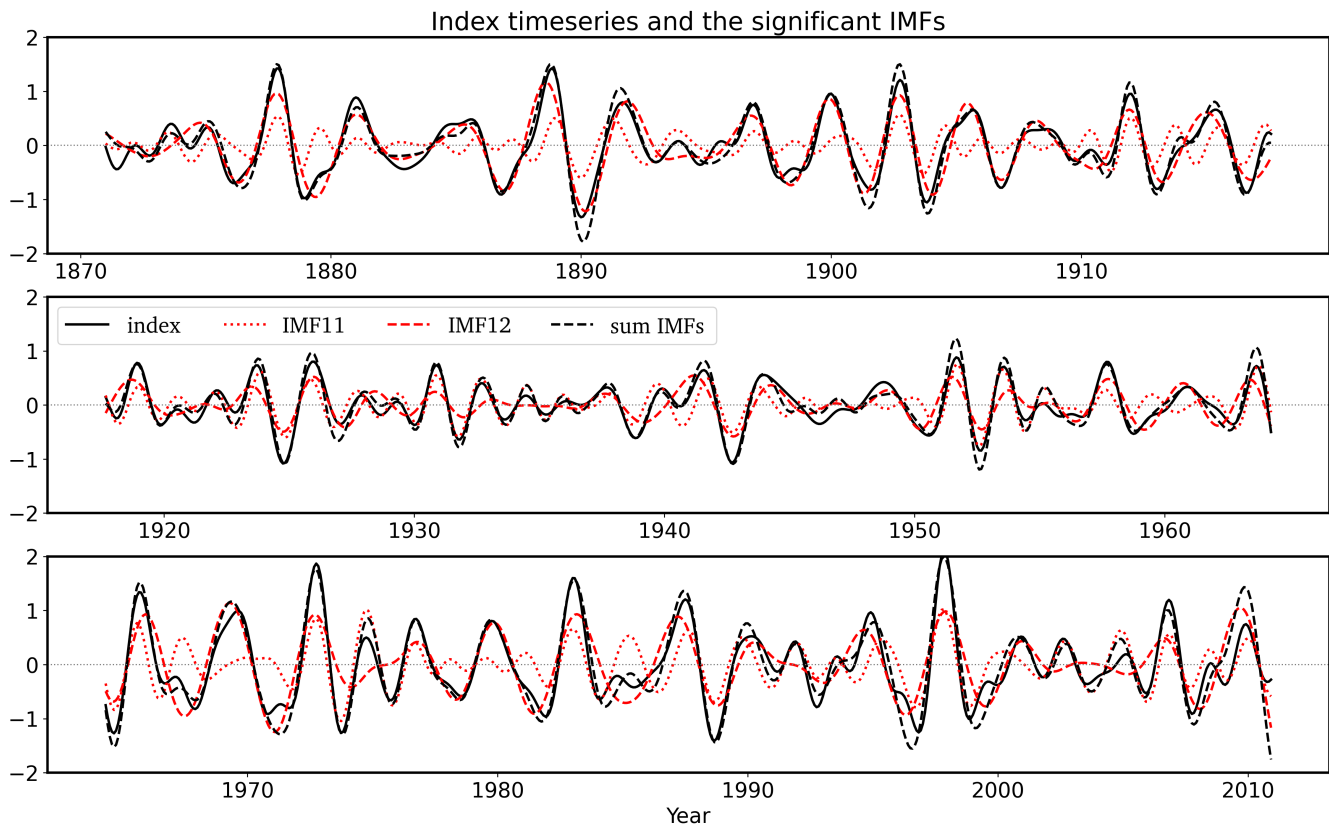
Figure 3. Timeseries (1871-2010) of Niño3 index from IMF11 (red dotted line) and IMF12 (red dashed line) obtained via MEMD. The sum of the two IMF indices (black dashed line) and band-pass filtered (16-53 months) Niño3 index (black solid line) are also shown.

470   Additionally, timeseries of Niño3 index extracted from ~~IMF12,13~~ IMF11,12 (Fig. 3) are largely consistent with the modes identified in Jiang et al. (1995) (their Fig. 9a), who used MSSA analysis, and with Wang and Ren (2020) (their Fig. 3), who used EEMD on Niño3.4 index. ~~IMF13~~ IMF12 is also similar to Froyland et al. (2021) 4-year mode (their Fig. 10), who used an operator-theoretic approach. ~~Also, the~~ The average periods of the 'significant' modes of variability (~~IMF12,13~~IMF11,12) in this study are typically lower than in other studies, however please recall that we have used much longer timeseries and that

475 the timeseries of IMFs are nonstationary. Thus, within the overlapping time periods (e.g., recent decades), the timescale (and corresponding timeseries) is generally consistent across different studies. These similarities provide further confidence in the results ~~presented below~~from the MEMD. Note that these studies have focused on different ENSO timescales and associated different spatial patterns during QB and LF/QQ ENSO events~~, but have not considered the~~ . However, IMFs are nonstationary and can thus capture changing patterns over time (section 5.2). Also, previous studies have not necessarily considered relationships

480 between different variables that are relevant for the ENSO dynamics (section 6) and related conceptual oscillator models ~~that we analyse below~~on specific timescales.

Overall, MEMD is consistent with other filtering methods (i.e., it acts like an effective band-pass filter) and can thus be used for further analysis of ENSO dynamics and its nonstationarity to determine if IMFs hold any physical significance. If IMFs have physical meaning, the information inferred from them can be valuable for enhancing climate models, long-term predictions, understanding teleconnections, and exploring the underlying physics and variability of specific fields of interest, such as ENSO. In the following sections we address some of these aspects.

5.2 Nonstationarity of ENSO

The period-ranges provided above (see also Fig. S1) suggest that there is some overlap between the period/frequency bands of IMF11, IMF12, which is a result of nonlinear and nonstationary evolution of the modes (period is not constant as seen in Fig. 3). Indeed, there is a low (yet statistically significant) correlation between the two modes (~~~0.27~~0.28). Thus, in some time-windows the two IMFs can describe the variability of a similar timescale (e.g., similar ~~time~~ temporal evolution in Fig. 3 around years ~~1902-3,~~ 1923-7, ~~1983~~1951-1955, 1997), but in other time-windows they describe variability on different timescales. Despite changing periods, the two IMFs ~~overall~~ pass the red noise threshold in the Niño3 region and are therefore still considered significant and thus *quasi*-periodic (~~Appendix B2~~sections 4, 5.1).

~~Fig. 3 also shows that in~~ In some decades the band-pass-filtered Niño3 index (black solid line) is more consistent with the lower frequency ~~IMF13~~ IMF12 (red dashed line; approx. 1870-1917, ~~1938-1950,~~ 1968-2000) and in other periods with the higher frequency ~~IMF12~~ IMF11 (red dotted line; approx. ~~1917-1938, 1950-1968,~~ 1917-1968, 2000-2010) ~~. This is likely~~ (Fig. 3). This is consistent with the interdecadal shifts in the frequency of ENSO (Hu et al., 2017, 2020) that have occurred around years 1970 (from higher frequency to lower frequency) and 2000 (from lower frequency to higher frequency)~~, and~~. Similar behaviour can also be seen earlier in the record~~in Fig. 3.~~. We can also see interdecadal changes in ~~amplitude of~~the amplitude of the ENSO modes, i.e., middle panel (~~periods~~ period 1920-1965) in Fig. 3 compared with top and bottom panels (periods 1870-1920, 1965-2010). This is somewhat consistent with Crespo et al. (2022), who found reduced amplitude of ENSO during 1901-1931 and 1935-1965 periods relative to post-1970 period.

Similarly, IMFs can capture different propagation directions of SST anomalies (Fig. 4). Previous studies have highlighted that SST anomaly propagation changed from westward to stationary or eastward around 1970 (Fedorov and Philander, 2000, 2001; Wang and A . Indeed, the two IMFs show westward propagation of SST anomalies prior to 1970. However, after 1970 it is slightly more complex. SST anomaly propagation becomes stationary or eastward in IMF11 (except for a period just before 1990; Fig. 4a), but IMF12 (Fig. 4b) still indicates either stationary or westward propagation with some rare (eastward) exceptions. This is then reflected in the sum of the two modes (Fig. 4c), which generally shows variability in propagation of SST anomalies post-1970. Furthermore, some of the characteristics of SST anomaly propagation from the sum of IMF11 and IMF12 carry over to 1-year low-passed SST anomalies (e.g., in the 1960s, 1980s; Fig. 4d).

ENSO events have also been characterised as east Pacific (EP) or central Pacific (CP) depending on the longitude where SST anomalies maximise (e.g., Kao and Yu, 2009; Singh and Delcroix, 2013; Zhang et al., 2019). The two IMFs are less able to reproduce this feature of ENSO diversity (Fig. 4). Most EP events are captured by IMF11,12. However, some CP events (e.g., early 1990s; Fig. 4d) are not captured by quasi-periodic modes of variability (Fig. 4a,b,c). This suggests that other processes are

likely causing them, such as sub-ENSO variability (Keenlyside et al., 2007). Similarly, some persistent La Niña events (e.g., mid 1970s, mid-1980s and around year 2000; Fig. 4d) are not necessarily captured by the quasi-periodic modes of variability alone (potentially related to ENSO asymmetry; e.g., Choi et al. 2013; An et al. 2020). These results could also reflect that some aspects of ENSO nonlinearity are not represented by these two IMFs (e.g., Dommenget et al. 2013).

520   Note that the magnitude of ENSO ultimately depends on all underlying modes of variability in the tropical Pacific (not just on the IMFs discussed here) ~~.~~ — as also seen in Fig. 4. In fact, we find ∼5 modes (i.e., IMFs ~~11-15~~ 10-14 here) with timescales ranging from ∼1 to ∼~~12~~ 11 years (assessed via 6.7^{th} and 93.3^{rd} percentiles as above) that can reproduce the majority of ENSO variability (not shown), but the rest of these modes (i.e., ~~IMF11,14,15~~ IMF10,13,14 here) are consistent with red noise (Fig 2).

~~Fig. 3also shows~~ Figs. 3, 4 also show that weak Niño3 events have either small amplitudes (e.g., ~~1933-7~~ 1930-40) of both

525   IMFs or opposite amplitudes (e.g., ~~1908-11), whereas~~ 1985-87). However, strong Niño3 events generally show a constructive interference or mode-combination (e.g., 1997, a super-El Niño event), which is consistent with, e.g., Slawinska and Giannakis (2017); Jajcay et al. (2018); Wang and Ren (2020); Froyland et al. (2021).

~~Note that the band-pass range here was chosen using an 'objective' data-driven approach, which shows that MEMD can be used as an effective band-pass filter method~~ This section has clearly demonstrated that MEMD together with a red noise test

530   is suitable for identifying nonstationary quasi-periodic multivariate signals. This is ~~supported by a good agreement between the band-passed data and the IMFs (similarly, with spectral analysis; Appendix B2.2; Fig. B3) , which suggests consistency between modes obtained via MEMD and other statistical methods. This provides grounds for using MEMD for further analysis of ENSO dynamics or, more specifically, for revisiting ENSO conceptual oscillator models, which we turn to next~~ a clear advantage of MEMD over other multivariate signal processing methods (e.g., PCA, MSSA). Below we now turn to ENSO

535   dynamics to show that IMFs are also physical.

6   ENSO dynamics

The dynamics of ENSO typically involves positive (e.g., Bjerknes) and negative feedbacks between the atmosphere and ocean. The Bjerknes feedback (Bjerknes, 1969) refers to any decrease (increase) of trade winds that leads to reduced (enhanced) ocean upwelling (downwelling) and thus warming (cooling) in the eastern tropical Pacific leading to reduced (enhanced) zonal SST-

540   and pressure-gradients, which in turn reinforce the initial increase (decrease) of the trade winds. The most prominent and also the simplest negative feedback in the tropical Pacific involves recharge/discharge of ocean heat content via Sverdrup transport (e.g., Jin, 1997a, b; Burgers et al., 2005). Other negative feedbacks involve propagation and reflection of ocean Rossby and Kelvin waves (e.g., Suarez and Schopf, 1988; Battisti and Hirst, 1989; Picaut et al., 1997; Weisberg and Wang, 1997; Wang et al., 1999; W , where the latter can also be wind-forced. These processes generally involve changes in the thermocline depth, surface wind

545   stress, and SST anomalies.

~~Timeseries (1871-2010) of band-pass filtered (17-52.5 months) Niño3 index (black solid line), IMF12 obtained via MEMD (red dotted line), IMF13 obtained via MEMD (red dashed line), and the sum of the two IMFs (black dashed line).~~

6.1 Quasi-oscillatory timescale

6.2 ~~Observational evidence~~

550 ~~ENSO evolution~~ As ENSO dynamics is primarily related to the evolution of the ocean surface zonal wind stress (τ_x), thermocline depth and the ~~SSTs~~ SST in the tropical Pacific (e.g., Wang, 2018)~~. Their relationship is demonstrated with map phase composites of IMF13~~, we demonstrate their relationship with phase composite maps of IMF12 (Fig. 5). Shading in Fig. 5 ~~, where shading~~ represents SST anomalies, contours represent the thermocline depth anomalies (solid contours represent positive values and deeper thermocline), yellow arrows represent τ_x anomalies, and grey shading represents grid points of the ~~SSTs~~
555 SST that do not pass the red noise threshold. All values are standardised (i.e., divided by standard deviation).

The phase composites are computed using instantaneous phase of the IMF12's eastern Pacific SST (Niño3) timeseries (~~from IMF13~~i.e., IMF12 (SST (Niño3))) that we can obtain via Hilbert Transform (Appendix B1, Eq. (B4)). This 'assigns' every point in the eastern Pacific SST (Niño3) timeseries a phase between 0 and 360 degrees, which can then be split into 12 phases (marked phase 0 through 11; e.g., phase 0 is 0-30°, ..., phase 11 is 330-360°) and all points in timeseries (of 1-D or 3-D fields)
560 belonging to a specific phase are then averaged to form map phase composites shown in Fig. 5 (see also Fig. ~~S2 for IMF12~~S3 for IMF11). Line phase composites ~~shown below~~ in Fig. 6 (see also Fig. S4 for IMF11) are constructed similarly. Eastern Pacific SST (Niño3) is chosen as ~~index here as conceptual models discussed below (section ??) involve SST (Niño3).~~ an index here to focus on East-Pacific (EP) ENSO events and related dynamics.

Fig. 5 shows a typical cycle of EP ENSO in the tropical Pacific (on a ∼3-year timescale) ~~, which can be summarised also~~
565 as composited over the IMF12's Niño3 index. This cycle can also be summarised with line phase composites (Fig. 6a,c) ~~of timeseries averaged~~ by averaging over specific regions ~~(as labeled~~of the tropical Pacific (as labelled; see also Table 1)~~of the tropical Pacific (e.g., Wang, 2018)~~. Here, we analyse ~~SSTs~~ eastern Pacific SST (Niño3), central Pacific τ_x (Niño4), western Pacific τ_x (Niño5), off-equatorial western Pacific thermocline depth (Niño6), and ~~thermocline depth (Pacific mean)~~Pacific mean thermocline depth as they have historically been considered important for ENSO dynamics and have been used in ENSO
570 conceptual oscillator models (e.g., Jin, 1997a; Burgers et al., 2005; Wang, 2018).

Together the two figures (Figs. 5, 6a,c) suggest the following sequence of events: ~~(i)~~

(i) during La Niña (phases 5-7) we have negative SST anomalies and shallower thermocline in Niño3 region, stronger easterly wind stress in Niño4 region, and deeper thermocline in the western Pacific (including Niño6 region); ~~(ii)~~

(ii) as La Niña weakens (phases 8-10), the westerly wind stress in Niño5 region and thermocline depth averaged across the
575 tropical Pacific peak, starting the El Niño cycle; ~~(iii) SSTs warm,~~

(iii) SST warm, eastern Pacific thermocline (Niño3) becomes deeper, central Pacific wind stress (Niño4) becomes westerly, and thermocline in the western Pacific (including Niño6 region) becomes shallower (phases 11,0,1); ~~(iv)~~

(iv) El Niño weakens (phases 2-4) and western Pacific τ_x (Niño5) becomes easterly as well as the thermocline averaged across the Pacific becomes shallower, starting a La Niña event (phases 5-7); ~~(v)~~

580   (v) the cycle repeats. ~~Note that there are also wind stress changes in the far eastern tropical Pacific that generally oppose the ones in the Niño4 region (Fig. 5), but occur roughly at the same time and are thus not explicitly considered further.~~

~~As in Fig. 6b, but for (a) higher frequency (band-pass filter over 11-19 months) data and (b) lower frequency (band-pass filter over 50-112 months) data. Note that higher and lower frequency timescale bands were chosen based on timescale ranges of 11th and 14th IMFs, i.e., the IMFs with slightly smaller or larger (respectively) timescales than IMF12, IMF13.~~

585   The evolution described above is also seen in the band-pass (~~25-52.5 months; 2-4.5~~ 29-53 months; 2.5-4.5 years) filtered data (Fig. 6b). Note that the values in Fig. 6b are slightly larger than in Fig. 6a, because slightly different frequency ranges are ultimately represented in the two panels, but they remain qualitatively similar. ~~Similar results can also be obtained for the 17-32 months band-passed data and IMF12 (e.g., Fig. S2 in supplement). However, at lower and higher frequencies the evolution is different with τ_x (Niño5) closely following thermocline (Niño6) (Fig. 8a,b), whereas other variables remain similar~~
590 ~~across timescales (to $\pm$ 1 phase). This suggests a different behaviour of the high- and low-frequency tropical Pacific variability compared with the 1.5-4.5-year variability (IMF12,13) discussed here.~~

~~Since we observe clear relationships between the relevant variables (e.g., Fig. 5) that strongly resemble a unified oscillator of Wang (2001a) (see also Fig. 5 in Wang 2018), recharge-discharge oscillator (e.g., Burgers et al., 2005), and others, we now revisit the theory of ENSO dynamics using the relevant conceptual oscillator models.~~

595 **6.2**   ~~**Conceptual oscillator models of ENSO**~~

~~**The Unified Oscillator**~~

~~As mentioned in the introduction, there are several different conceptual oscillator models (e.g., Wang, 2018): (i) the delayed; (ii) the recharge-discharge; (iii) the advective-reflective; (iv) the Western Pacific; and (v) the unified oscillator. The latter encompasses all the previous oscillators, and can be described by four relatively simple equations (Wang, 2001a)~~

600 $$\frac{\mathrm{d}T}{\mathrm{d}t} = a\tau_1 - b_1\tau_1(t-\eta) + b_2\tau_2(t-\delta) - b_3\tau_1(t-\mu) - \varepsilon T^3$$

$$\frac{\mathrm{d}h}{\mathrm{d}t} = -c\tau_1(t-\lambda) - R_h h$$

$$\frac{\mathrm{d}\tau_1}{\mathrm{d}t} = dT - R_{\tau_1}\tau_1$$

$$\frac{\mathrm{d}\tau_2}{\mathrm{d}t} = eh - R_{\tau_2}\tau_2$$

~~where T is SST in Niño3 region, h is thermocline depth in Niño6 region, τ_1 is τ_x in Niño4 region, τ_2 is τ_x in Niño5 region,~~
605 ~~the constants are: $a = 1.5 \times 10^{-2}$ K m^2 N^{-1} yr^{-1}, $b_1 = b_3 = 2.5 \times 10^2$ K m^2 N^{-1} yr^{-1}, $b_2 = 7.5 \times 10^2$ K m^2 N^{-1} yr^{-1},~~
~~$c = 1.5 \times 10^3$ m^3 N^{-1} yr^{-1}, $d = 3.6 \times 10^{-2}$ N m^{-2} K^{-1} yr^{-1}, $e = 3 \times 10^{-3}$ N m^{-3} yr^{-1}, the damping coefficients are:~~
~~$\varepsilon = 1.2$ K^{-2} yr^{-1}, $R_h = 5$ yr^{-1}, $R_{\tau_1} = R_{\tau_2} = 2$ yr^{-1}, and the delay times (lags) are: $\eta = 150$ days, $\delta = 30$ days, $\lambda = 180$~~

days, $\mu = 90$ days. The unified oscillator and all its special cases listed above have a timescale range between 2 and 5 years with the provided parameters.

610   Eqs. (??)-(??) represent different processes in the tropical Pacific (Wang, 2001a). Eq. (??) represents changes to the SSTs in the Niño3 region via: (i) positive Bjerknes feedback (first term on the right-hand-side (RHS)); (ii) negative feedback due to Kelvin wave reflection at the western ocean boundary (second term on the RHS); (iii) negative feedback due to wind-forced Kelvin wave contribution in the equatorial western Pacific (third term on the RHS); (iv) negative feedback due to Rossby wave reflection at the eastern ocean boundary (fourth term on the RHS); and (v) cubic damping term that limits anomaly growth

615   (last term on the RHS). Eq. (??) represents changes to the off-equatorial thermocline depth (Niño6 region) via the wind stress in the equatorial central Pacific (Niño4 region) (first term on the RHS) and linear damping (second term on the RHS). Eq. (??) represents changes to the zonal wind stress (τ_x) in the equatorial central Pacific (Niño4 region) via SSTs in equatorial eastern Pacific (Niño3 region) (first term on the RHS) and linear damping (second term on the RHS). Finally, Eq. (??) represents changes to the zonal wind stress in the equatorial western Pacific (Niño5 region) via off-equatorial western Pacific thermocline

620   depth (Niño6 region; first term on the RHS) and linear damping (second term on the RHS).

However, the results from section 6 suggest that the average evolution on 2-3 year (average) timescale is different from the unified model and many other oscillator models discussed in Wang (2001a). In section 6 we have established that τ_x (Niño4) closely follows SST (Niño3) (While Fig. 6 a,b,c), i.e., they co-vary. Similarly, Fig. 6a,b,c suggests that thermocline depth (Niño6) is 'anticorrelated' with SST (Niño3) and τ_x (Niño4). Therefore, we now assume: (i) $\tau_1 \propto T$ by setting $\mathrm{d}\tau_1/\mathrm{d}t = 0$,

625   yielding $\tau_1 = dT/R_{\tau_1}$, i.e., we replace τ_1 by dT/R_{τ_1} everywhere; (ii) $h \propto -\tau_1, -T$ by setting $\mathrm{d}h/\mathrm{d}t = 0$, yielding $h = -c\tau_1/R_h = -cdT/$ i.e., we replace h by $-cdT/R_hR_{\tau_1}$ everywhere; (iii) assume that the wave reflections at both eastern and western ocean boundaries are not necessary for this oscillator (typically assumed for the Western Pacific oscillator) by setting $b_1 = b_3 = 0$; (iv) for simplicity set $\lambda = \eta = \mu = 0$, and keep the other parameters the same. This yields a modified unified oscillator

$$\frac{\mathrm{d}T}{\mathrm{d}t} = a\frac{d}{R_{\tau_1}}T + b_2\tau_2(t-\delta) - \varepsilon T^3$$

630    $$h = -\frac{cd}{R_hR_{\tau_1}}T$$

$$\tau_1 = \frac{d}{R_{\tau_1}}T$$

$$\frac{\mathrm{d}\tau_2}{\mathrm{d}t} = -e\frac{cd}{R_hR_{\tau_1}}T - R_{\tau_2}\tau_2.$$

Fig. 6d shows a phase composite over the synthetic timeseries generated by solving equations (??)-(??) numerically with a first order Euler method (other more complex methods yield the same results). The phase composite (lines are the same as

635   in Fig. 6c, except for exclusion of the blue dotted line) clearly demonstrates that the lags between τ_x (Niño5), thermocline (Niño6), and SST (Niño3) or τ_x (Niño4) are consistent with observations (Fig. 6c). The timescale of this oscillator (using the same parameters as above) is ~3 years, i.e., similar to IMF13 timescale. The timescale was assessed via the timelapse between different peaks in the timeseries that the oscillator model of Eqs. (??)-(??) yields.

The parameters in Eqs. (??)-(??) can theoretically be adjusted (e.g., to change timescale) until the oscillatory behaviour breaks (or the 'system' becomes unphysical), which is true also for the unified oscillator in Eqs. (??)-(??). Note that if $\delta = 0$ this oscillator and its timescale remain similar (not shown). This is consistent with Weisberg and Wang (1997), who noted that Western Pacific oscillator remains qualitatively similar when all lags are set to 0. We will refer to Eqs. (??)-(??) as a *simplified West-Pacific (SWP) Oscillator* from hereon, since this model is somewhat similar to the Western Pacific oscillator developed by Weisberg and Wang (1997), except for the conditions $\tau_1 = dT/R_{\tau_1}$ and $h = -cdT/R_h R_{\tau_1}$. We view the SWP Oscillator model as another special case of the unified oscillator.

As mentioned above, the Western Pacific oscillator can be captured by Eqs. (??)-(??) by setting $b_1 = b_3 = 0$, typically describing the following series of events (e.g., Wang, 2001b, their Fig. 2). During the warm phase of ENSO there is (*atmospheric*) condensation heating in the equatorial central Pacific (Niño4) (e.g., Deser and Wallace, 1990) that induces a pair of low pressure anomalies in the off-equatorial central Pacific region. This drives the westerly wind anomalies in the central Pacific (Gill 1980 ; see also yellow arrows in phases 10,11,1,2 in Fig. 5). The wind stress (Niño4) then leads to deeper thermocline and warmer SSTs in Niño3 region (solid contours and red colours in phases 10,11,1,2 in Fig. 5), i.e., a positive (Bjerknes) feedback further amplifies the signal (reflected in Eqs. (??) first term, (??)).

However, these off-equatorial low pressure anomalies act to raise the off-equatorial thermocline via Ekman pumping (evident through dashed contours in phases 10,11,1,2 in Fig. 5) (e.g., Wang, 2001b). This brings colder waters to the off-equatorial ocean surface, cooling the SSTs in Niño6 region (captured in Eq. (??) via thermocline impacts on SSTs), and introduces a pair of off-equatorial high pressure anomalies in the western Pacific. These then induce easterlies in equatorial western Pacific (Niño5) (yellow arrows in phase 3 in Fig. 5; also reflected in Eq. (??)), which can ultimately cause upwelling of the cold waters. This upwelling then extends further eastward (dashed contours and blue colours in phases 3-8 in Fig. 5; also reflected in Eq. (??) – third term on the RHS) with Kelvin wave propagation (reflected in delay time δ), leading to a negative phase of ENSO, La Niña, and the cycle repeats.

What is different in the SWP Oscillator compared with, e.g., the West-Pacific oscillator? The relationship between h, T and τ_1 in Eqs. (??), (??) suggests that on 17-52.5 months timescale the atmosphere over the central-eastern Pacific is co-varying (is in a 'steady state') with the underlying ocean in SWP Oscillator (cf., Wang, 2001a), and that eastern (co-varying with SSTs (Niño3)) and western Pacific thermocline depth anomalies also co-vary, but are anti-correlated, unlike in the West-Pacific oscillator. However, the western Pacific wind stress (Niño5) keeps an out-of-phase relationship with the SSTs (Niño3), suggesting an important role of the western Pacific in ENSO variability on 17-52.5 months timescale. Note that this analysis does not necessarily imply causality, and that other processes are likely needed to ultimately cause an ENSO event (e.g., related to a delayed oscillator or recharge-discharge oscillator discussed below). This is because the western Pacific winds are weak compared to the central Pacific winds. However, the western Pacific winds are important for forcing eastward-propagating Kelvin waves, which can represent additional feedback for ENSO growth/decay (e.g., Wang, 2018). The western Pacific wind anomalies could also reflect 'state-dependent westerly wind bursts' in the western Pacific (Lopez et al., 2013; Lopez a that are important for the onset of ENSO events.

Fig. 6 and Eqs. (??)-(??) describe an ~~describes an~~ average evolution of the ~~parameters (τ_1, τ_2, T, h), but~~ variables considered, not all individual events have this exact behaviour~~—~~. This is partly reflected in the reduced amplitudes of τ_x and thermocline depth in Fig. 6a compared with the SST amplitude. The full timeseries of these parameters from IMF12 (Fig. 7; line types/colours are the same as in Fig. 6) show that the ~~relationship~~ relationships from Fig. 6 ~~a,b,c occurs~~ occur often in the analysed period, especially for strong events. However, for weak events (middle panel in Fig. 7) the ~~relationship is~~ relationships are harder to establish ~~——~~ every event seems to be slightly different. This is somewhat consistent with Crespo et al. (2022), who noted that the dynamics of ENSO was different prior to 1970 relative to after 1970 with a dominant recharge-discharge oscillator (Jin, 1997a, b; Burgers et al., 2005) in the latter period. ~~Note that below we show that there is likely a relationship between the SWP and the recharge-discharge oscillators.~~

~~IMF13's standardised timeseries of SST (Niño3) (black solid line), thermocline depth (Niño6) (black dashed line), thermocline depth (across tropical Pacific) (blue dotted line), τ_x (Niño4) (grey dashed line), τ_x (Niño5) (grey solid line).~~

~~As seen in section 6, the dynamics on timescales of about 2-3 years (with a range of 1.5-4.5 years; Fig. 6) is different from the higher and lower frequencies (Fig. 8). Thus, the variables that are important for the SWP Oscillator are not necessarily important for the variability on shorter~~Similar results (phase composites) can also be obtained for the 16-33 months band-passed data and IMF11 (Figs. S3, S4 in the supplement). This suggests that on average the QB and LF~~/longer timescales. On shorter/longer timescales, one can make the assumption $\tau_2 \propto h$, since τ_x (Niño5)~~closely follows thermocline (Niño6). Additionally, Fig. ~~8 shows that $\tau_1, T \propto -\tau_2, -h$, i.e., these quantities are anticorrelated. This correlation suggests that a unified model on shorter/longer timescales may be simplified to a delayed oscillator (involving T alone). Note that we discuss below the recharge-discharge oscillator of Burgers et al. (2005), which is also present on those timescales (blue dotted lines~~ QQ ENSO events have similar evolution and associated dynamics. However, the frequency of events (see full timeseries in Fig. 8)~~.~~

~~Some of the results here are consistent with Graham et al. (2015), who considered 12-month low-pass filtered data to assess the conceptual unified model of Wang (2001a). They provided several suggestions to improve the unified model of Wang (2001a), e.g., to remove the tendency terms for both wind stress terms (in Eqs. (??) , (??)). As mentioned above, Fig. 8 suggests that such approximation would likely yield better results than the original unified oscillator model. However, this would only work in the high- and low-frequency ranges discussed here~~ S5 in the supplement) that follow the dynamics identified in the phase composites (Fig. 8) ~~, but not on timescales in between, i.e., timescales of about 1.5-4.5 years (Fig. 6). On those intermediate timescales, the SWP Oscillator presented here should be used instead. This means that on these intermediate timescales the tendency of the wind stress in the Niño5 region should be kept, but the one in the Niño4 region may be omitted~~S4) is lower in IMF11 case than in IMF12 case. This may be indicative of other processes that could be relevant for the QB ENSO.

~~Graham et al. (2015) also provided other suggestions, like adding a stochastic forcing term to wind stress equations (not attempted here for simplicity), and they also suggested using the thermocline depth in the Western Pacific averaged over a region that lies on the equator, rather than off-Equator. From Fig. 5 we can see that this would likely yield similar results as thermocline depth averaged over the Niño6 region (to ± 1 phase). They also suggested that a delayed oscillator is generally~~

sufficient for describing ENSO variability, which may be true but only for shorter/longer timescales The results presented here and summarised in the line phase composites (Fig. 8) that are not necessarily quasi-periodic.

The Recharge-Discharge Oscillator

The SWP Oscillator is based on Wang (2001a), which uses West-Pacific wind stress and off-equatorial thermocline depth as part of the overall ENSO oscillator. However, Burgers et al. (2005) suggested using only the 6, S4) have some interesting implications for the dynamics of ENSO and conceptual oscillator models. The out-of-phase relationship between Pacific mean thermocline depth and SSTs (Niño3), which led to the (simplest) recharge-discharge oscillator

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} T \\ h_{\mathrm{Pac}} \end{pmatrix} = \begin{pmatrix} -2\gamma & \omega_0 \\ -\omega_0 & 0 \end{pmatrix}\begin{pmatrix} T \\ h_{\mathrm{Pac}} \end{pmatrix}$$

where T is SST (Niño3), h_{Pac} is the thermocline depth averaged across the tropical Pacific, $\gamma^{-1} = 24^{+22}_{-11}$ months is the damping timescale, and the period of this oscillator is $2\pi\omega^{-1} = 37^{+8}_{-4}$ months ($\approx 2\pi\omega_0^{-1}$), with $\omega^2 = \omega_0^2 - \gamma^2$. The sub- (blue dotted line) and super-scripts in the period and damping timescale provide their 95% confidence levels. Note that the Burgers et al. (2005) model is based on the model developed by Jin (1997a, b). The recharge-discharge oscillator (e.g., Wang, 2001b) is related to Sverdrup transport (Sverdrup, 1947), which is associated with both zonal wind stress in the central Pacific (i.e., off-equatorial wind-stress curl) and SSTs in the eastern Pacific. This is driving the recharge-discharge of the equatorial heat content that ultimately gives rise to oscillations.

Fig. 6a,b,c clearly demonstrates that eastern Pacific SST (Niño3; black solid line) and thermocline depth averaged across the tropical Pacific (blue dotted line) are ~ 90° out of phase, consistent with is a typical feature of the recharge-discharge oscillator of Jin (1997a, b); Burgers et al. (2005). We can also see (Fig. 6a, b,c) that τ_x ((Jin, 1997a, b; Burgers et al., 2005). This suggests that IMFs that emerge from MEMD analysis can capture physical processes in the tropical Pacific. Furthermore, co-variability of the Pacific mean thermocline depth and the western Pacific wind stress (Niño5) (grey solid line) shows variability consistent with (proportional to) the thermocline depth averaged across the tropical Pacific (blue dotted line), i.e., one may assume $\tau_2 \propto h_{\mathrm{Pac}}$ (in fact, their standardised values are essentially equal for IMF13). As with the SWP Oscillator, the relationship between the relevant variables in the recharge-discharge oscillator on the 2-3 year timescale largely remains the same over time (Fig. 7).

Wang (2001a) suggested that a recharge-discharge oscillator can be derived from the unified model by assuming $\tau_2 \propto h$. However, Fig. 6 suggests that on 2-3 year timescale the Pacific mean thermocline (h_{Pac}) and wind stress ; grey solid line) suggests that wind forcing in the western Pacific (τ_2) co-vary, thus the approximation should be $\tau_2 \propto h_{\mathrm{Pac}}$ to retrieve the Burgers et al. (2005) model. The recharge-discharge and SWP oscillators are thus likely related on 2-3-year timescale, which could suggest a role of western Pacific wind-forced Kelvin waves in recharge-discharge process (and vice versa). Thus, another way of looking at the two oscillators is to combine them, e.g.: (i) by using the recharge-discharge oscillator (Eq. **??**) or SWP Oscillator (Eqs. **??-??**) and setting $\tau_2 = \alpha_1 h_{\mathrm{Pac}}$ or $h_{\mathrm{Pac}} = \alpha_2 \tau_2$ (with $\alpha_{1,2}$ a constant); (ii) by replacing h in the unified model with h_{Pac} and re-tuning the parameters of the model; or (iii) setting τ during the recharge/discharge process to a non-zero value

τ_2 (see, e.g., Fig. 1b,d in Jin 1997a), i.e., potentially re-tuning may play a role in the recharge-discharge oscillator. It may also be necessary to reassess the relative importance of different variables in the oscillator models. process or in the onset of ENSO in general as it precedes ENSO events (see also Roulston and Neelin, 2000; Capotondi et al., 2018; Lopez et al., 2013; Lopez and Kirtman, 20 .

The relationship between the SSTs (Niño3) and the Pacific mean thermocline depth remains the same at lower and higher frequencies (see black solid and blue dotted lines in Fig. 8), suggesting that the recharge-discharge oscillator (Jin, 1997a, b; Burgers et al., 2 is present on all timescales considered here. This makes the relationship between the SWP and recharge-discharge oscillators less clear on these longer/shorter timescales (Fig. 8). This On the other hand, the co-variability between central Pacific wind stress (Niño4; grey dashed line), off-equatorial western Pacific thermocline depth (Niño6; black dashed line) and eastern Pacific SST (Niño3; black solid line) suggests that the relationship between the two models is non-trivial, and can depend on the timescale (Figs. 6, 8). However, this also suggests that the unified oscillator captures different dynamics on different timescales, and that this is likely related to different behaviour in the western Pacific on different timescales, since τ_x (Niño5)'s relationship with other variables changes substantially. These relationships should be explored further in the future.

Overall, in this section we have identified a SWP Oscillator in the frequency range 1.5-4.5 years in observations/reanalyses (Fig. 6c) and through a conceptual oscillator model (equations (??)-(??); Fig. 6d) by modifying unified oscillator proposed by Wang (2001a) may need to be revised (see also Graham et al., 2015). This is because: (i) in the unified oscillator model (equations (??)-(??); Wang 2001a). At lower and higher frequencies we have found a different behaviour, likely related to different dynamical processes in the western Pacific on those timescales. Note, however, that a quasi-periodic behaviour in observations on shorter/longer timescales has not been detected. We have also found a recharge-discharge oscillator (Jin, 1997a; Burgers et a on *all* timescales, suggesting similar dynamical processes in the eastern Pacific on all timescales considered, and a complex relationship with the unified model (Wang, 2001a) and SWP Oscillator. This suggests that a reassessment of the conceptual oscillator models and their links across timescales may be necessary (left for future work). Additionally, this section has confirmed that MEMD can extract physical modes of variability from the data.

7 Further implications of understanding ENSO's quasi-periodic variability

The previous section established that the unified model (e.g., Wang, 2001a) can be simplified to SWP Oscillator model in observations on quasi-periodic timescales and that its relationship with recharge-discharge oscillator (e.g., Burgers et al., 2005) is complex. Here, we first discuss the representation of the SWP and recharge-discharge oscillators in a climate model (section ??). Then, we test ENSO predictability on quasi-periodic timescales using a simple statistical model (section ??, Appendix C).

6.1 Model evaluation

Since the modes extracted via MEMD are physical and the significant modes are quasi-periodic (via a red noise test), we can now repeat the analysis in a climate model. Here, we use the NorCPM1 historical simulation (first ensemble member; the other

ensemble members show qualitatively similar results) to test the ENSO dynamics in different frequency bands (similar to Fig. 6). The MEMD analysis yields two significant (quasi-periodic) modes also in the model: IMF14 with a timescale of 27 months (range: 22 - 35 months), and IMF15 with a timescale of 38 months (range: 27 - 54 months) . As in the observations, the two modes show similar behaviour (see Figs. S3-S5 in supplement) and have a similar range (22-54 months) of quasi-periodic behaviour as observations (17-52.5 months).

As in Fig. 6a, but for the corresponding IMF15 fields in the NorCPM1 (ensemble member 1).

Fig. **??** shows the model's IMF15 phase composites, a figure equivalent to observations' Fig. 6a (note that as in observations, also here the same band pass filter yields similar results to an IMF). For full temporal evolution (model equivalent of Fig. 7) see Fig. S6 in the supplement. While the recharge-discharge oscillator (e.g., Burgers et al., 2005) is well represented in the model (black solid line for SST (Niño3) and blue dotted line for the tropical these three quantities should all be somewhat out-of-phase, but here we show that only western Pacific wind stress and Pacific mean thermocline depth exhibit out-of-phase relationship with eastern Pacific SST (on average); and (ii) Pacific mean thermocline depth), the SWP Oscillator (other lines and black solid line) shows slightly different behaviour compared with observations. This is especially evident in the τ_x (Niño5; grey solid line), which peaks just before the SST (Niño3; black solid line), leading it by ∼1 phase (Fig. **??**), whereas in observations (Fig. 6a) it peaks well before the SST (Niño3), leading it by ∼3 phases. At the same time, τ_x (Niño5; grey solid line) is anticorrelated with does not co-vary with off-equatorial western Pacific thermocline depth (Niño6; black dashed line), which is different from observations where thermocline depth (Niño6) was anticorrelated with SST (Niño3) and τ_x (Niño4).

This suggests an issue in the), rendering the unified model's representation of the ENSO dynamics on 2-3 year timescale, which is especially pronounced in the western Pacific (SWP Oscillator), where both τ_x (Niño5) and recharge-discharge oscillator simplification, which uses off-equatorial western Pacific thermocline depth (Niño6)tend to have phase-lag relationship with SST (Niño3) slightly incorrectly represented relative to observations (Figs. **??**, 6a). On the other hand, the recharge-discharge mechanism, which is stronger in the eastern Pacific , is more correctly represented, i.e., phase-lag relationship between thermocline depth (Pacific mean) and SSTs (Niño3) is similar to observations. The τ_x (Niño4) shows similar evolution as in the observations, but its strength is stronger in the model (compare grey dashed lines in Figs. 6a, **??**), suggesting a stronger feedback to the SST (Niño3) or less varied dynamical processes. Consistently, there is also a stronger relation of the SST (Niño3) to the tropical Pacific mean thermocline depth (larger amplitude relative to the SSTs compared with observations; compare blue dotted lines in Figs. 6a, **??**; see also large values in spatial composites of Fig. S4), likely suggesting a stronger (more dominant) recharge-discharge oscillator in the model compared with observations.

The strong East-Pacific thermocline-SST feedback in the model was also identified as a general model bias in Chen et al. (2021) , who also identified model biases in the western Pacific (e.g., SST anomalies during ENSO events extending too far west; see also Figs. S3, S4) . These are likely responsible for different model behaviour in the western Pacific compared with observations. We also found differences at low and high frequencies between the model and the observations in the unified oscillator, further suggesting potential issues with the dynamics of ENSO on different timescales in the model. questionable.

This analysis suggests that timeseries analysis methods (e.g., MEMD) may be used to evaluate models. Here, such analysis confirms that the analysed model has correct periodic behaviour (perhaps too periodic), however it may be struggling with the

exact physics that lead to the onset of ENSO (or ENSO diversity), which should be considered for future model improvements (see also Guilyardi et al. 2020; Planton et al. 2021). Nonetheless, since (i) the recharge-discharge process is well represented in the model; (ii) the model exhibits quasi-periodic behaviour; and (iii) SST-variability alone could be sufficient for describing/predicting ENSO (see section **??**; Graham et al. 2015) – the NorCPM1 (prediction system) should be able to predict ENSO, especially on timescales ~2-4 years (left for future work)That the central Pacific wind stress may be omitted in the *unified model* (due to co-variability with the eastern Pacific SST) was also mentioned in Graham et al. (2015). However, they suggested that western Pacific wind stress (Niño5) could also be omitted from the *unified model*, but Fig. 6 suggests that on 1.5-4.5 year timescale this variable should be kept. This may be because Graham et al. (2015) used 1-year low-pass filtered data, which could have obscured the signal on 1.5-4.5 year timescale (see section 6.1). Additionally, Graham et al. (2015) suggested using the thermocline depth in the western Pacific averaged over a region that lies on the equator, rather than off-equator. From Fig. 5 we can see that this would likely yield similar results as off-equatorial western Pacific thermocline depth (Niño6) (to ± 1 phase).

6.1   ~~Prediction~~Other timescales

While a better understanding of the quasi-periodic variability (e.g., in ENSO) can help climate model improvements and thus ultimately their prediction skill, it can also help identification of timescale-windows in which predictions (e.g., of ENSO) work well. The latter means that quasi-periodic signal of ENSO (as identified above) may be predictable far in advance, even though full ENSO prediction is more challenging. This is what we test in this section.

To do this, we use a relatively simple multi-linear regression model (for details see Appendix **??**; cf., Omrani et al. 2022) with three (input) timeseries: ENSO is a phenomenon that occurs on timescales of 2-8 years and previous work has often used 1-year low pass filter to obtain ENSO. Thus, we now test the relationships between eastern Pacific SST (Niño3), central Pacific τ_x (Niño5), and thermocline depth (Niño6)to predict the SST (Niño3). We included Niño4), western Pacific τ_x (Niño5), western Pacific off-equatorial thermocline depth (Niño6)as a predictor as this improved the results compared with using just SST (Niño3) and τ_x (Niño5) as predictors, which are the two main quantities in the SWP Oscillator (Eqs. **(??)** - **(??)**). On the other hand, we do not use τ_x (Niño4) as it is clear from Figs. 6, 8 that it is somewhat redundant (has the same evolution as the SST (Niño3)), and including it actually weakens predictability (not shown). Similarly, predictions from tropical , and Pacific mean thermocline depth are not considered here as they did not improve predictions (not shown). on slightly shorter and slightly longer timescales. We do this to test how relationships between different variables change across timescales that are still somewhat within the ENSO range.

Before inputting timeseries into the prediction model (for details see Appendix **??**) we lag the data so that all timeseries maximise at 'lag 0' and we additionally use up to 19 past timesteps of each timeseries to improve predictions. The prediction model is constructed separately for each lead time and is repeated for all months within the 1871-2010 period (except for the timesteps used for lagging the data). We use 60% of data for training and 40% of the data for testing.

The skill of the prediction is assessed using the correlation between the true SST (Fig. 8 shows phase composites similar to those in Fig. 6b, but for band-passed data over 12-19 months range (shorter periods; Fig. 8a) and over 42-135 months range

(longer periods; Fig. 8b). These timescale bands are consistent with IMF10 (shorter periods) and IMFs 13,14 (longer periods), which are the modes that together with IMF11,12 explain the majority of variance in, e.g., eastern Pacific SST (Niño3) and its prediction in the testing dataset. This is done on raw data (i.e., unsmoothed) and on smoothed data using 25-52.5 months band-pass, where the SST (Niño3)quasi-periodic behaviour is strongly significant in observations (i.e., IMF13 timescale band). Note that the prediction model is always applied on raw data, and the prediction itself is later smoothed by using the band-pass filter, before computing the correlations. This is done because any smoothing of the data before inputting it into the prediction model can yield misleading (over-confident) results. An extension to 17-52.5 months band (used above) reduces the skill significantly (perhaps because IMF12 is marginally significant; e.g., Fig. S2) and is hence not pursued further (not shown).

Correlation skill score of a multi-linear regression model for predicting SST (Niño3) from (a) 7 past timesteps (in months), and (b) 19 past timesteps (in months). The dashed lines are for raw predictions (without smoothing), and solid lines are for a (smoothed) band-pass filtered prediction (25-52.5 months). Note that we smooth the raw prediction to get the smooth-prediction skill score. The grey lines are for a prediction from SST (Niño3) alone, whereas black lines are for prediction from SST (Niño3), τ_x (Niño5) and thermocline depth (Niño6). For further details see the text.

Fig. **??** shows SST (Niño3) prediction skill for 30 months lead time using (a) 7 months of past timesteps within predictors, and (b)19 months of past timesteps within predictors for predictions from SST (Niño3) alone (grey lines), and predictions from all predictors (black lines). Dashed lines represent the raw prediction skill,and solid lines represent prediction skill of smoothed data (25-52.5 months band-pass). This reveals that there is extended prediction skill in the smoothed data to ~18 months lead time (skill score over 0.6) compared with raw prediction skill (~5 months) in the simple statistical model. This is consistent with the oscillatory behaviour inthat timescale-band (identified in section 5). Note that smoothing predictions over a narrower timescale band (e.g., 26.5-51 months) before computing correlations can yield slightly better results, i.e., longer predictability range (not shown).

Fig. **??**a reveals that we can get a good prediction of SST Interestingly, on these shorter and longer timescales the evolution is different than on quasi-oscillatory timescales (1.5-4.5-year periods of IMF11,12). Namely, western Pacific wind stress (Niño5) closely follows western Pacific off-equatorial thermocline depth (Niño6) (Niño3) from SST (Niño3) alone by using 7 past timesteps (months), which can be slightly extended by a month or two when all predictors are considered (black solid line crosses the 0.6-skill-score line at a slightly longer lead time beyond 18 months). By using 19 past timesteps (months) (Fig. **??**b), we can improve raw predictions (dashed lines) for lead times 10-15 months, and the skill-score of smoothed predictions using all predictors (black solid line)improves relative to Fig. **??**a. However, the latter yields slightly worse skill-score than prediction of SST (Niño3) from SST (Niño3) alone with 7 past timesteps (Fig. **??**a,grey solid line), i.e., it lasts ~8a,b). This suggests a very different role of the western Pacific wind stress on different timescales, which may be relevant for ENSO diversity/asymmetry. However, other variables remain similar across timescales (to $\pm$ 1 month lead-time less in Fig. **??**b(black solid line)– until lead time ~17 months. Note that root-mean-square error was not sensitive to the choice of predictors, hence it is not shown.

These results suggest that (linear)statistical predictions of ENSO can be well captured by SST (Niño3)alone, which is perhaps consistent with Graham et al. (2015), who suggested that a delayed oscillator is sufficient for describing ENSO. The

~~importance of~~ phase). Thus, recharge-discharge processes operate throughout the ENSO timescale range (2-8 years), since Pacific mean thermocline depth and eastern Pacific SST (Niño3) ~~alone may also be due to other processes that can affect variables like thermocline depth (affected by, e.g., deep ocean dynamics) and~~ remain out of phase also on these shorter/longer timescales. Also, the results from these shorter/longer timescales indicate that the western Pacific wind stress ~~(affected by, e.g., high-frequency atmospheric variability)~~ . ~~These processes can then interfere with a (linear) statistical prediction of SSTs~~ Niño5) may be omitted from conceptual oscillator models (on these timescales) as it is ~~computed from raw (unfiltered)data. Note that, here, we use a very simple linear statistical model and it is therefore hard to compare it directly to dynamical models or more sophisticated statistical models. Either way, this shows that ENSO's quasi-periodic signal can be well predicted months in advance (even with a simple linear statistical model).~~

~~While this section suggests that there is indeed potential for very good predictability (for ~18 months lead time) of ENSO on timescales 25-52.5 months, such predictability can only provide the sign of ENSO in this frequency band, and it can suggest a higher chance of positive or negative ENSO event, but it cannot provide the actual magnitude and prediction of the full ENSO event. A lot of work has been done on dynamical and statistical model predictions of ENSO, which can predict an ENSO event reasonably well up to 6 months ahead (sometimes more) from raw data(e.g., L'Heureux et al., 2020; Dijkstra et al., 2019) . However, we show that such models might have a predictable (up to 1.5 years or more) component on timescales 25-52.5 months. This is also because models can simulate ENSO variability on this timescale well (as established in section ??). Perhaps ENSO events that can be predicted up to 2 years ahead (e.g., Chen et al., 2004; Park et al., 2018) have a stronger component of this ~2-4.5 year mode of variability (relative to red-noise variability of other modes~~ anti-correlated with eastern Pacific SST (Niño3). This is consistent with Graham et al. (2015), who suggested this revision of the *unified oscillator*, but using 1-year low-pass filtered data, which may have obscured the different behaviour on quasi-periodic timescales (section 6.1; Fig. 6).

~~Note that here we computed 'prediction' skill without any preconditioning or a selection of a season, therefore better skill scores may be obtained if the model is initialised in relevant seasons or based on certain precursors (left for future work)~~

The above analysis shows that it is important to filter the data to "correct" frequency bands as there may be different behaviour present on different timescales, even within the ENSO range of 2-8 years.

7 Conclusions

In this study we have used ~~observational and reanalysis products to study the variability in the tropical Pacific on interannual timescales (i.e., ENSO) and to revisit conceptual oscillator models of ENSO. To do this, we have used~~ a recently developed nonlinear and non-stationary method for identifying intrinsic variability of multivariate systems, the multivariate empirical mode decomposition (MEMD; Rehman and Mandic 2010). The method can objectively identify modes of variability on different timescales within a nonlinear and non-stationary dataset describing a complex system such as the climate system. The timescale identification is objective as it is done without any pre-selection of a timescale window in which we expect the quasi-periodic behaviour. It finds a signal that is synchronised across input timeseries (here PC timeseries of combined fields, i.e., ~~it works as an objective band-pass filter. Then,~~ SST, wind stress, thermocline depth over the tropical Pacific) for

every timescale within the given system. Also, the multivariate modes of variability that emerge from MEMD analysis have non-stationary (i.e., evolving in time) patterns of variability (e.g., Fig. 4)—a clear advantage over some other multivariate timeseries analysis tools whose patterns are stationary.

Additionally, a red noise significance test has been ~~used to extract~~ developed to robustly identify quasi-periodic modes of variability in the given data~~.~~, which had not been used before in the framework of MEMD. This means that MEMD can now be used as an alternative method for objective identification of the timescale of quasi-periodic motions in the climate system. Since the red noise test can be applied on every grid point separately, MEMD together with the red noise test can also be used for identifying potential new regions of quasi-periodic variability (similar to Fig. 5).

~~Here, we~~ We demonstrate that MEMD can identify physical quasi-periodic modes of variability within the climate system by analysing tropical Pacific SST variability. We have identified a clear quasi-periodic behaviour on a timescale of about ~~3 years (25-52.5 months; arguably also on extended timescale, 17-52.5 months, as discussed in previous sections)~~ 2-3 years (16-53 months) in the tropical Pacific. This timescale falls within the typical timescale range of ENSO, i.e., 2-8 years ~~. While the~~ and the dynamics of this quasi-periodic variability is consistent with ENSO dynamics. While ENSO quasi-periodic variability is ~~a~~ well-known~~feature~~, an identification (via MEMD) of a ~~range of timescales of the~~ frequency range linked to the two dominant quasi-periodic modes ~~(e.g., 25-52.5~~ of variability (i.e., 16-53 months) has still led to a few interesting results.

By analysing composites (e.g., Figs. 5, 6) of the thermocline depth, ~~surface~~ wind stress and ~~the sea surface temperature (SST)~~ SST, we have shown that the ~~∼3-year ENSO variability generally follows an oscillator model that is a subset of the unified oscillator (Wang, 2001a). We refer to this oscillator as a simplified West-Pacific (SWP) Oscillator, as it is slightly different from the other established oscillator models (i.e., the Western Pacific, delayed, advective-reflective,~~ 2-3-year (Eastern Pacific) ENSO variability is generally consistent with the recharge-discharge ~~oscillators) . We have also identified the recharge-discharge oscillator of Burgers et al. (2005), which is described by a slightly different set of equations than the unified oscillator (i.e., uses Pacific mean thermocline depth instead of average over the Niño6 region). However, the recharge-discharge oscillator is present also on other timescales (not just on ∼3-year timescale), unlike the SWP Oscillator. This suggests similar behaviour in the eastern Pacific on all timescales considered, but different behaviour in the western Pacific. This means that the relative role, phasing of and relationship between relevant variables in the tropical Pacific may change depending on the timescale. This may ultimately be important for interactions across scales that give rise to, e.g., an ENSO event, thus these relationships should be explored further in the future.~~

~~The SWP Oscillator (Eqs. (??)-(??)) resembles the Western Pacific Oscillator of Weisberg and Wang (1997), in that it keeps the evolution (tendency) of wind stress in the western Pacific (Niño5). However, based on the results from the reanalyses/observations (Fig. 6a,c) we assume that wind stress in the Niño4 region and thermocline depth in Niño6 region can be modelled through the SSTs in the Niño3 region (i.e., their tendencies can be omitted). This is because wind stress (Niño4) closely follows SST (Niño3) and thermocline depth (Niño6) is largely anticorrelated with~~ conceptual oscillator model of ENSO (e.g., Jin, 1997a; Burgers et al., ?). This oscillator describes an interplay between Pacific mean thermocline depth and eastern Pacific SST (Niño3)~~. This suggests co-variability of the atmosphere and ocean on interannual timescales in the central-eastern Pacific, but an out-of-phase relationship with the western Pacific coupled atmosphere-ocean processes, which are important for the generation of Kelvin waves at~~

the western boundary. The latter provides additional feedbacks (e.g., Wang, 2018) for the variability in the eastern Pacific, suggesting that the western Pacific processes are important for ENSO variability. There might be an additional link to the

945 'state-dependent westerly wind bursts' that are important for ENSO onset (e.g., Lopez et al. 2013; Lopez and Kirtman 2014; Timmermann). However, since the amplitude of the western Pacific wind stress anomalies is generally small, it is likely that other processes are important as well – e.g., recharge-discharge of the heat content (see below).

The evolution of the SWP Oscillator on timescales of 17-52.5 months can be described in the following way (cf., Wang 2001b ; Figs. 5, 6a ,c): warm phase of ENSO is associated with westerly winds in Niño4, warmer SSTs and deeper thermocline in

950 Niño3 (phases 10, 11,0 in Fig. 5); this is associated with off-equatorial pair of low pressure anomalies in central Pacific, which can cause upwelling (via Ekman pumping) of the off-equatorial thermocline (becoming shallower), extending further west (incl. into the Niño6 region). This largely occurs together with the warmer SSTs in Niño3 (see evolution of dashed contours in phases 9-11, 0-2 in Fig. 5); the colder waters caused by upwelling of the thermocline in Niño6 lead to off-equatorial pair of high pressure anomalies in the western Pacific , which ultimately gives rise to easterlies in the western Pacific (Niño5)(e.g.,

955 yellow arrows in phase which are related to recharge and discharge of ocean heat content on a timescale of about 3 in Fig. 5); these easterlies can cause further upwelling of cold waters (shallower thermocline) in the equatorial western Pacific , which can then propagate as Kelvin waves further east (dashed contours in phases 0-4 in Fig. 5), bringing colder waters to the eastern Pacific and initiating the cold phase of ENSO (phases 4-6 in Fig. 5) ; the cycle reverses and ultimately repeats.

The recharge-discharge oscillator of Burgers et al. (2005) is also present on the 17-52.5 months timescale throughout the

960 analysed period (Fig. 7) and should be considered as part of (or alternative to) the SWP Oscillator (or vice versa). This is because the years. However, the unified model (Wang, 2001a) may need to be revised (see also Graham et al., 2015) as most of the variables relevant for the unified oscillator model co-vary. In particular, western Pacific off-equatorial thermocline depth (Niño6), central Pacific wind stress (Niño4), and eastern Pacific SST (Niño3) exhibit the same phase relation and thus describe the same dynamics. As also western Pacific wind stress (Niño5) and the ; relevant for the unified oscillator) and Pacific mean

965 thermocline depth co-vary and are likely related to each other – their relationship should be explored further in the future. The evolution of the , there seems to be a close relationship between the recharge-discharge oscillator (alone) can be summarised as follows (e.g., Burgers et al. 2005; Wang 2001b; Figs. 5, 6a,c): (i) during the warm phase of ENSO (phases 10,11,0 in Fig. 5) there is westerly wind stress in the central Pacific, warmer SSTs and deeper thermocline in the eastern Pacific; (ii) this is associated with divergent Sverdrup transport that ultimately drives discharge of the equatorial heat content, which leads to

970 (climatological) upwelling of cold waters throughout the equatorial Pacific (phase 3 in Fig. 5); (iii) this initiates a cold phase of ENSO (phases 4-6 in Fig. 5) , the cycle reverses and ultimately repeats. and unified oscillator models, but it may be different than previously thought.

Exploring the representation of ENSO's intrinsic variability in a climate model (NorCPM1) showed that the model has a similar quasi-periodic behaviour as the observations/reanalyses, with largely accurate (though too strong) recharge-discharge

975 oscillator. However, SWP Oscillator is not really present, i.e., the dynamics is different, since On shorter and longer timescales (12-19 months, 42-135 months; Fig. 8) the relationships between variables are different, specifically the western Pacific wind stress (Niño5) peaks too late and thermocline depth (Niño6)peaks too early (Fig. ??). Similarly, the co-variability of the

recarge-discharge and SWP oscillators is absent in the model. Therefore, we speculate that climate models (likely not just NorCPM1) may struggle with the relationship between the western and eastern Pacific coupled on these timescales co-varies

980 with eastern Pacific SST (Niño3). This suggests that the role of the western Pacific atmosphere-ocean processes .

By constructing a statistical prediction model, we have shown that SST variability on timescales of 25-52.5 months in the tropical Pacific may be predictable for up to 18 months in advance (Fig. **??**) . Also, given that NorCPM1 can reproduce the (via wind stress) in tropical Pacific variability (and ENSO) can be very different on different timescales, whereas other variables (and in other regions) largely keep their relationships across timescales. This means that recharge-discharge oscillator well

985 and that it exhibits periodic variability on oscillator model operates on all timescales 2-4 years, we believe that it should be able to predict the variability on the 2-4 year timescale reasonably well – this is likely true for other models as well (left for future work) . However, such prediction can only tell us that there is a higher chance of a certain ENSO event and it cannot yield a prediction of full ENSO amplitude, i.e., this remains challenging. This is because an ENSO event ultimately depends on variability on all timescales (and their interactions), thus its peak magnitude and timing will likely differ from the one identified

990 on a specific (band-passed) timescale.

Therefore, a better understanding of ENSO variability considered (12-135 months), even if it has a characteristic timescale. However, the relations underlying the unified oscillator model exhibit different behaviour on different timescalesis important for a better representation of ENSO dynamics in the climate models. Additionally, it is important to understand ENSO impacts on different timescales both locally and remotely (teleconnections; e.g., Brönnimann, 2007; Fereday et al., 2008; Jiménez-Esteve and Domeis

995 . The latter can be achieved through similar analysis as in this study, but including other fields and other (remote) regions in the analysis (e.g., mean-sea-level pressure in the Euro-Atlantic region) , implications of which should be explored further in the future.

MEMD analysis could be extended in several ways. For example: (i) to assess ENSO dynamics in models as they typically struggle with the representation of the western Pacific processes (i.e., relevant for Niño5 wind stress; Guilyardi et al., 2020; Planton et al., 2

1000 ; and (ii) to study ENSO teleconnections on different timescales (e.g., Brönnimann, 2007; Fereday et al., 2008; Jiménez-Esteve and Domeis . Also, future studies should involve an examination of sensitivity and causal links (not established here) between different variables (and their links across scales) within the Tropics and beyond (e.g., Runge et al., 2015; Jajcay et al., 2018; Jenney et al., 2019; Kretschmer et al., 2021), as well as dedicated model-experiments.

Overall, this study has analysed the variability in the tropical Pacific , identified a (using MEMD with a red noise test),

1005 identified two quasi-periodic mode modes of variability (on ∼3-year timescale)and related its 2-3-year timescale), related their physics to the SWP and recharge-discharge oscillators, which are likely related to each other on this timescale. Variability oscillator, and suggested a revision to the unified oscillator model (somewhat consistent with Graham et al., 2015). As the variability on this timescale is quasi-periodic, it may be predictable far in advance, which calls for further investigations of the tropical Pacific variability and related teleconnections, their prediction, and for further model improvements (see also Chen

1010 et al., 2021; Lee et al., 2021).

Code availability. EMD and MEMD Python codes are available on Github (https://github.com/laszukdawid/PyEMD, https://github.com/mariogrune/MEMD-Python-). Other scripts are available upon request.

Data availability. SODA2 data can be downloaded from http://apdrc.soest.hawaii.edu/dods/public_data/SODA/soda_pop2.2.4; HadISST data can be downloaded from https://www.metoffice.gov.uk/hadobs/hadisst.

Appendix A: MEMD for 3-D fields

1015

To find the intrinsic variability of our 3-D field, i.e., ~~$\mathbf{A}'(t,y,x)$~~ $\mathbf{A}'(L_t, L_y, L_x)$ mentioned in section ~~3~~3.3, we first reduce dimensionality of our data by decomposing it using the singular value decomposition (SVD), which yields spatial patterns of our data (empirical orthogonal functions, EOFs) and corresponding timeseries (principal components, PCs). First, we multiply $\mathbf{A}'$ by $\sqrt{\cos\phi}$ (area weighting; with ϕ latitude), divide by standard deviation (σ) at each grid point (and for each variable separately, if relevant), and reshape $\mathbf{A}'$ from ~~(t,y,x) to (t,xy)~~3-D array (L_t, L_y, L_x) to 2-D array $(L_t, L_x \cdot L_y)$. Then $\mathbf{A}'$ can be expressed with a singular value decomposition as

1020

$$\sigma^{-1}(\underset{xy}{L_x \cdot L_y})\mathbf{A}'(\underset{t}{L_t}, \underset{xy}{L_x \cdot L_y}) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{A1}$$

where $\mathbf{U}$ and $\mathbf{V}$ represent left and right singular vectors ~~, i.e., the normalised~~ related to PCs and EOFs, ~~$\mathbf{\Sigma} = \sqrt{(N-1)\mathbf{\Lambda}}$~~ $\mathbf{\Sigma} = \sqrt{(L_t-1)\mathbf{\Lambda}}$ is a diagonal matrix with square roots of variance explained of each mode (denoted $\mathbf{\Lambda}$, i.e., eigenvalues) along the diagonal, ~~N is the number of spatial points~~L_t is the length of timeseries, and superscript T denotes transpose. ~~In order to keep the information of the variance explained of each mode within the data we define PCs as $\mathbf{U}\mathbf{\Sigma}/\sqrt{N-1}$~~PCs are defined as $\mathbf{U}\sqrt{L_t-1}$ and EOFs as $\mathbf{\Sigma}\mathbf{V}^T/\sqrt{L_t-1}$, such that ~~$\sigma^{-1}(xy)\mathbf{A}'$~~ $\sigma^{-1}(L_x \cdot L_y)\mathbf{A}'$ can be represented as a function of PCs and EOFs ~~(recall EOFs are in $\mathbf{V}$)~~

1025

$$\sigma^{-1}(\underset{xy}{L_x \cdot L_y})\mathbf{A}'(\underset{t}{L_t}, \underset{xy}{L_x \cdot L_y}) \underset{=\sqrt{N-1}}{\approx} \sum_{m=1}^{m=20} \text{EOF}(m, \underset{xy}{L_x \cdot L_y})\text{PC}(m, \underset{t}{L_t}) \tag{A2}$$

where m corresponds to PC-number and is ordered according to the eigenvalues ($m=1$ for the largest eigenvalue). We retain only the leading 20 PCs for the analysis (they generally describe the majority of the variance in $\mathbf{A}'$).

1030

Now we can use the 20 PCs (PC~~(m,t)~~(m, L_t)) as input to MEMD algorithm (for details on algorithm itself see Rehman and Mandic 2010). This algorithm finds common timescales (i.e., Intrinsic Mode Functions, IMFs) within the 20 PCs and splits each PC into several IMFs (the number of IMFs is not predetermined). Thus, each PC can be represented as a sum of IMFs

1035

$$\text{PC}(m, \underset{t}{L_t}) = \sum_{s=1}^{s=s_{max}} \text{IMF}(s, m, \underset{t}{L_t}) \tag{A3}$$

where s corresponds to IMF-number and is ordered according to the timescale ($s=1$ for the shortest timescale, s_{max} for the longest timescale, which is usually a trend or a residual). Eq. (A3) shows that each PC is a superposition of different IMFs (see also Table S2) and with it also a superposition of modes of variability in the selected field(s) with different timescales.

Each PC is associated with a spatial pattern (EOF~~(m,xy)~~$(m, L_x \cdot L_y)$), which allows a reconstruction of the time-space ~~(t,xy)~~ $(L_t, L_x \cdot L_y)$ structure/evolution for each IMF, yielding IMFs of initial dataset $\mathbf{A}'$ (**IMFA**). We do this by multiplying PCs for each IMF with corresponding EOFs and summing over all 20 PCs/EOFs (similar to Eq. A2)

$$\sigma^{-1}(\underset{xy}{L_x \cdot L_y})\mathbf{IMFA}(s, \underset{t}{L_t}, \underset{yx}{L_x \cdot L_y}) \simeq \sqrt{N-1} \sum_{m=1}^{m=20} \text{IMF}(s, m, \underset{t}{L_t})\text{EOF}(m, \underset{xy}{L_x \cdot L_y}). \tag{A4}$$

Here note that to get **IMFA** in the units of the input field we must multiply it by the field's standard deviation as the input data for the SVD algorithm ~~was~~ were standardised (Eq. A1). Again, **IMFA**s are ordered by timescale, i.e., **IMFA**1 with the shortest timescale and **IMFA**s_{max} with the longest timescale (trend). From here we can reconstruct ~~$\mathbf{A}$~~ $\mathbf{A}'$ by summing over all **IMFA**

$$\sigma^{-1}(\underset{xy}{L_x \cdot L_y})\mathbf{A}'(\underset{t}{L_t}, \underset{xy}{L_x \cdot L_y}) \simeq \sum_{s=1}^{s=s_{max}} \mathbf{IMFA}(s, \underset{t}{L_t}, \underset{yx}{L_x \cdot L_y}) \tag{A5}$$

and ultimately one can also reshape $\mathbf{A}'$ from ~~(t,xy) to (t,y,x)~~ 2-D array $(L_t, L_x \cdot L_y)$ to 3-D array (L_t, L_x, L_y). The importance of each **IMFA** for $\mathbf{A}'$ can be assessed by computing variance explained of each **IMFA** or other significance methods. To find **IMFA** modes (and grid-points) that correspond to potentially oscillatory behaviour we must perform a red noise test (see Appendix B).

Note that from here on (and in the main text) **IMFA**s are referred to as IMFs for simplicity.

Appendix B: Significance tests

Typically we can test if the modes (IMFs) are different from white or red noise, ~~if we expect such distribution in~~ depending on the distribution of our data. In the climate system, ~~more often than not we expect the variables to behave as~~ variables often exhibit behaviour that resembles white or red noise. The IMFs that are significant (i.e., different from both red and white noise) ~~are likely representing oscillations instead, which can suggest~~ likely represent quasi-oscillations, indicating higher potential for predictability of processes that correspond to the timescale of the significant IMF. Thus, this distinction is very important in climate ~~system~~ science. Therefore, we discuss the white and red noise tests (for 1-D data, i.e., timeseries) below, whereas the robustness of IMFs ~~and the~~ from MEMD analysis and the relevant significance tests are ~~briefly mentioned where relevant~~ discussed in the main text (sections 4, 5).

Note that the white and red noise tests are performed on 1-D timeseries, hence EMD (univariate decomposition; see main text) is first used to test ~~their performance~~ the performance of IMFs that arise from the EMD analysis. The multivariate data (via MEMD) ~~is later~~ in the main text is analysed with the simplest and most relevant test (i.e., theoretical red noise test described below).

B1 White noise test

The white noise significance test has been derived by Wu and Huang (2004), who showed that the energy density function of sth IMF (E_s; i.e., average squared amplitude of the sth IMF) is

$$E_s = \frac{1}{L_t} \sum_{j=1}^{L_t} [\text{IMF}_s(j)]^2, \tag{B1}$$

where $\text{IMF}_s(j)$ is the sth IMF at time-step j ($j = 1, ..., L_t$), and L_t is the length of the timeseries. Wu and Huang (2004) further showed that the total energy density of the sth IMF can then be expressed as

$$L_t E_s = \int S(\nu)_s \, \mathrm{d}\nu \tag{B2}$$

where ν is frequency, and $S(\nu)_s$ is the power spectrum of the sth IMF. From this they showed that for white noise

$$\ln E_s \approx -\ln T_s \tag{B3}$$

where T_s is the average period of the sth IMF, and ln denotes natural logarithm.

Note that frequency (and thus also period) of each IMF is computed using Hilbert transform by first generating an analytical signal (e.g., Huang et al., 1998)

$$Z(t) = X(t) + \mathrm{i}Y(t) = |Z(t)| \mathrm{e}^{\mathrm{i}\theta(t)} \tag{B4}$$

where t is time dimension, $X(t)$ is our IMF timeseries, $Y(t)$ is its Hilbert transform, $Z(t)$ is the analytical signal, and $\theta(t) = \arctan(Y(t)/X(t))$ is instantaneous phase. Instantaneous frequency can then be computed by taking a time-derivative of the phase, i.e., $\nu = \mathrm{d}\theta(t)/\mathrm{d}t/2\pi$, and the average frequency of each IMF is computed by averaging instantaneous frequency in time (note that period = $1/\nu$, i.e., T_s).

The relationship between the logarithms of energy density and average period of the IMFs (Eq. B3) is then used in Fig. B1a (black solid line) to test whether an IMF (using EMD decomposition of Niño3 index; blue dots in Fig. B1a; see also section 3) is different from white noise or not. The mode is significant with respect to white noise if it exceeds one-tailed 95th percentile threshold (denoted by black dotted line). The percentile range serves as a significance test, i.e., if IMFs from our data are above, e.g., 95th percentile they are significant at 95th percentile level. The percentile range can be expressed analytically as (Wu and Huang, 2004)

$$\ln E_s = -\ln T_s \pm p \sqrt{\frac{2}{L}} \sqrt{\frac{2}{L_t}} \exp\left(\ln T_s / 2\right) \tag{B5}$$

where p denotes a threshold ($p = 1.645$ for one-tailed 95th percentile of Gaussian distribution). Note that typically the number of degrees of freedom (DoF) for white noise data is expected to be equal to the total energy density of the sth IMF (i.e., $\text{DoF}_s = L_t E_s$; Wu and Huang 2004).

~~Blue dots in Fig. 2 are (in general) constructed the following way. First, we obtain IMFs via either EMD (Fig. 2a,b) or MEMD (Fig. 2c,d) analysis (for details see section 3 and Appendix A). Then, we compute average energy density (Eq. B1) from timeseries of each IMF, and average period of each IMF (for periods of IMFs from MEMD analysis see Tables S1, S2)~~

1095 ~~using Hilbert transform (see text under Eq. B4). Then we obtain a natural logarithm of both average amplitude and average period of each IMF and plot them in Fig. 2 as $\log_e$(period) vs. $\log_e$(amplitude) of each IMF as a scatter plot. Thus, in Fig. 2 each blue dot represents amplitude and frequency of each IMF that we have identified within our timeseries/data. IMFs are ordered by frequency along the x-axis, i.e. left-most blue dot is IMF1 and right-most blue dot is trend (last IMF). Note that all dots and lines in Fig. 2 represent logarithmic values of period/energy-density.~~

1100 Alternatively, we can test whether the input data is different from white noise by constructing multiple (I) realisations of synthetic white noise timeseries w_i (i^{th} random normally distributed timeseries with standard deviation σ of 1). Then we can compute its IMFs via EMD (section 3) and we can repeat the process I-times. Employing Eq. B1 on these IMFs yields scattered grey dots in Fig. ~~2~~B1a (constructed in the same way as blue dots; see also section 4, Fig. 2), where their mean and 5th-95th percentile are shown as grey solid and dotted lines, respectively.

1105 A comparison with the IMFs from the input data (Niño3 index; blue dots in Fig. ~~2~~B1a) reveals that many IMFs lie outside the white noise range and that overall the data (blue dots) distribution does not resemble the white noise (grey dots) distribution (not noted in Wu and Huang 2004). This suggests that a white noise test for such data is not a good test. Indeed, atmosphere-ocean coupled systems, such as ENSO, can often be represented as a red noise process (e.g., Hasselmann 1976; Frankignoul and Hasselmann 1977), thus we now turn to a similar test, but for data distributed as red noise.

1110 ~~Significance tests for EMD and MEMD: (a) white noise significance test and (b) red noise significance test for EMD-IMFs of Niño3 index (blue dots); (c) (theoretical) red noise test for IMFs of PC1 of the combined field (via MEMD; blue dots) (d) (theoretical) red noise test for Niño3 obtained via MEMD (blue dots). (a) black solid line represents the theoretical linear relationship between the logarithms of period (T_s) and logarithm of energy density (E_s) (Eq. (B3)), black dotted line represents 5th-95th percentile (Eq. B5), respectively; grey dots represent $I = L$ realisations of IMFs of white noise timeseries (length is~~

1115 ~~the same as for Niño3 index), whereas grey solid and dotted lines represent their mean and the 5th-95th percentile, respectively. (b) red solid line represents the theoretical red spectrum energy density (Eqs. (B7-B9)), red dotted line represents the 95th percentile (via χ^2-test); light pink dots represent $I = L$ realisations of IMFs of red noise timeseries (Eq. B6; length L is the same as for Niño3 index), whereas pink solid and dotted lines represent their mean and the 95th percentile, respectively. In (c) and (d) the red solid and dotted lines are as in (b) but for the respective data from MEMD. Note that x-axis shows the~~

1120 ~~logarithms of period (T_s) ordered from shortest period (highest frequency) to longest period (trend). Also, in (d) IMF22 was too large to fit the graph, but was identified (see Fig. S7, Table S1). For further description of the figure see text.~~

B2 Red noise test

B2.1 Synthetic red noise data

To test if our data (e.g., Niño3 index) is purely red noise or it has inherent oscillations that we can identify through the (M)EMD analysis, we generate I-realisations of synthetic red noise timeseries x (AR(1) process) as (e.g., Gilman et al. 1963)

$$x_{i,j+1} = rx_{i,j} + \sqrt{1-r^2}\, w_{i,j+1} \tag{B6}$$

where r is lag-1 auto-correlation from our data (e.g., Niño3 index), w is white noise (as in Appendix B1), i runs over I realisations of synthetic red noise data, and j ($j > 1$ and ~~$j \leq L$; with L~~ $j \leq L_t$; with L_t length of our data, e.g., the length of Niño3 record) is an index that runs over the time-steps (one time step is one time unit, e.g., 1 month). For $j = 1$ (the first time-step) we set $x_{i,1} = w_{i,1}$.

Once we obtain the red noise timeseries x_i we can compute its IMFs via EMD (section 3) and we can repeat the process I-times (as for the white noise; Appendix B1). This yields the pink scattered dots in Fig. ~~2~~B1b. The mean over I cases for each IMF (frequency band) is shown by pink solid line and the (one-tailed) 95th percentile across the I cases are shown by pink dotted line. Note that we plot logarithmic values in Fig. ~~2~~B1, as mentioned above.

Note that Franzke (2009) used a similar approach for indices such as the North Atlantic Oscillation, and found a simple relationship between the power spectrum and frequency, consistent with Kolotkov, D. Y. et al. (2016). However, we follow Gilman et al. (1963) to define a relationship between the power spectrum and frequency. This incorporates the lag-1 auto-correlation of the timeseries into the theoretical red noise power spectrum (see below).

B2.2 Theoretical red noise test

Alternatively, one can compute a theoretical power spectrum of the red noise (cf., Gilman et al. 1963)

$$S(\nu) = \frac{1-r^2}{1-2r\cos 2\pi\nu + r^2} \tag{B7}$$

where S is the power spectrum of red noise, $\nu = 1/t$ is frequency, and r is again lag-1 auto-correlation from our data. For each frequency estimate we must multiply $S(\nu)$ by frequency range ($\Delta\nu$) (cf. Eq. (B2)) to obtain a theoretical estimate of the (mean) energy of the red noise (E^{red}) (cf. Kolotkov, D. Y. et al. 2016)

$$E^{red}(\nu) = S(\nu)(\nu\alpha - \nu/\beta) \tag{B8}$$

where $\beta = \sqrt{\nu_s/\nu_{s+1}}$ and $\alpha = \sqrt{\nu_{s-1}/\nu_s}$ with s running over frequencies (from higher to lower frequency). Note that since EMD is a dyadic filter (each lower frequency is a half of the previous one; e.g., Flandrin et al. 2004; Rehman and Mandic 2011) both α and β typically take a value of $\sqrt{2}$ (consistent with, e.g., Kolotkov, D. Y. et al. 2016). ~~Note that~~ However, when mode-mixing is present (e.g., ~~in this study~~ here it is generally present at higher ~~frequencies~~ and lower frequencies of IMFs from MEMD analysis) this is not necessarily true, hence the use of α and β in Eq. (B8). Finally, ~~E_{red}~~ E^{red} for every s^{th} IMF

(i.e., E_s^{red}) must be scaled such that its total energy is the same as the total energy of our data (e.g., Madden and Julian 1971; Bretherton et al. 1999)

$$E_s^{red} = E_s^{red} \frac{\sum_{s=2}^{s_{max}-1} E_s}{\sum_{s=2}^{s_{max}-1} E_s^{red}} \tag{B9}$$

where s_{max} is the number of IMFs (as above), and s represents the sth IMF of frequency ν_s, and E_s was defined above (Eq. (B1)). Note that we scale E_s^{red} from total energies of IMFs between $s = 2$ and $s = s_{max} - 1$ as the last IMF is typically a trend/residual and the first IMF does not necessarily follow the distribution correctly (but including the two usually does not significantly alter the results). E_s^{red} is shown in Fig. ~~2~~B1b as red solid line.

This red noise test is typically used in climate science to determine the significance of power spectra peaks in our data (using $S(\nu)$ from Eq. (B7)), and it differs from the red noise test of Kolotkov, D. Y. et al. (2016) as it takes into consideration the lag-1 auto-correlation of the data. If the cosine function in the $S(\nu)$ (Eq. (B7)) is expanded into Taylor series ($\cos 2\pi\nu \approx 1 - (2\pi\nu)^2/2 + ..$) one can realise that for large ν (high frequencies) $S(\nu)$ indeed reduces to the spectrum $\gamma\nu^{-2}$ (with γ a constant) suggested by Kolotkov, D. Y. et al. (2016); Franzke (2009). However, for low frequencies (small ν) they do not agree well and ultimately $S(\nu)$ also becomes a constant (see Fig. B2 for comparison). Furthermore, as $S(\nu)$ depends on lag-1 auto-correlation (r) we can see from Eq. (B7) that for $r = 0$, $S(\nu) = 1$, i.e., it reduces to the power spectrum of the white noise. This means that this theoretical test can potentially be used for testing the significance of the data that corresponds to either white or red noise.

The significance of the IMFs from the input data is tested using χ^2-test, where s^{th} IMF's χ_s^2 value for the (one-tailed) 95th percentile is computed from $\mathrm{DoF}_s = L_{eff} E_s$ degrees of freedom (instead of ~~LE_s~~ $L_t E_s$ as was the case for white noise, due to strong correlations between neighbouring data-points; Bretherton et al. 1999; Wu and Huang 2004; Kolotkov, D. Y. et al. 2016), where ~~(Bretherton et al., 1999)~~

$$L_{eff} = \frac{1-r^2}{1+r^2} L_t. \tag{B10}$$

Then we multiply the expected red noise curve E_s^{red} by χ_s^2/DoF_s (e.g., Madden and Julian 1971; Bretherton et al. 1999) to ultimately obtain a threshold for 95th percentile (red dotted line in Fig. ~~2~~B1b). Note that for $\mathrm{DoF}_s < 1$ we set $\mathrm{DoF}_s = 1$ (to avoid numerical issues). The IMFs derived from the data (blue dots in Fig. ~~2~~B1b) that exceed the red noise threshold (i.e., lie above the red/pink dotted line in Fig. ~~2~~B1b) are considered significant at 95th percentile (one-tailed).

Fig. ~~2~~B1b shows that the two (synthetic and theoretical) red noise tests (for Niño3 index via EMD) are somewhat comparable and that the majority of the input (e.g., Niño3 index) data (blue dots) lies within the red noise range (i.e., within the pink-dots, and below the pink/red dotted line). However, we can identify one IMF (period ~31 months or ~2.5 years) that is above the red noise threshold and well within the typical ENSO timescale (2-8 years), suggesting quasi-periodic behaviour (oscillations).

Similarly, we can ~~identify one significant IMF (via MEMD) in the PC1 (Fig. 2c) of the 3-variable field (discussed in section ??, Appendix A) with a period of ~37 months (~3 years; i.e., IMF13). Note that we use PC1 of the combined field as an example of the overall tropical Pacific IMFs (via MEMD) in Fig. 2c, since PC1 strongly dominates the EOF decomposition of the SST field there. One could potentially also compute significant modes by computing red noise test at each grid-point and~~

~~then average the results over all grid-points, but we have not used this here. Instead, we use an additional test on map-plots in~~
1185 ~~section 5 (Fig. 5; Figs. S2-S4 in supplement), where we identify *potentially* "oscillatory" grid points and use grey shading on areas that are well represented with red noise alone (i.e., not significant).~~

~~Upon a reconstruction of IMFs for the x-y-t SSTs (Eq. A4), we can average SSTs over the Niño3 region (using the same IMFs~~ use the theoretical red noise significance test on IMFs obtained via MEMD~~; see also section 3). This yields two significant modes (i.e., IMF12,~~ IMF13; Fig. 2d) ~~both within the quasi-biennial QB and LF/QQ range of the ENSO~~
1190 ~~(Allan, 2000; Jajcay et al., 2018), i.e., with periods of ~22 and ~34 months (~2-3 years), respectively. This means that Niño3 index has~~ which yields 2 significant modes of variability ~~(identified via MEMD), but the region of significance is narrow (small) for the ~22 month-mode (IMF12; see Fig. S2), thus it is not apparent in the PC1 test that encompasses the whole tropical Pacific (Fig. 2c). The latter serves as an additional (more subjective) test for quasi-periodic modes of variability across the tropical Pacific, which suggests that IMF12 is only marginally significant and only in the~~ in the eastern Pacific SST (Niño3~~region (Fig.~~
1195 ~~S2) . Similarly, another (more subjective) test was used to make sure that the two quasi-periodic modes identified are well separated (in terms of timescale)from the other modes (and from each other) to avoid identifying mixed modes as significant quasi-periodic modes of variability (see below).~~

~~Note~~) index (see the main text). Note, however, that we do not necessarily expect exactly the same results from EMD (Fig. ~~2B~~1b) and MEMD (Fig. 2~~d~~) methods since the MEMD finds a "synchronised" signal within the tropical Pacific and
1200 across different variables, whereas EMD only analyses the 1-D Niño3 timeseries. This is also true for the number of IMFs obtained via the two different methods. MEMD yields significantly more IMFs as EMD (~~22~~ 21 versus 10), which is likely a result of inputting several different timeseries with different timescales, especially in the ~~high-frequency range here~~ high- and low-frequency range (i.e., periods shorter than about 8 months ~~)~~and longer than about 700 months).

Alternatively, one can also compute significant modes by computing red noise test at each grid-point and then average the
1205 results over all grid-points, but we have not used this here. Instead, we use an additional test on map-plots in section 5 (Fig. ~~For example, wind stress can be much noisier than SSTs (or different parts of the tropical Pacific have different variabilities) and can thus lead to identification of several high frequency modes of variability. This can then result in mode-mixing in that frequency range (see somewhat overlapping and close blue dots in Fig. 2c,d)if timescale separation is not clear across different input timeseries, and can ultimately yield more modes (here 22)than when only one timeseries is considered (here 10)~~5; Fig. S3
1210 in supplement), where we identify *potentially* "oscillatory" grid points and use grey shading on areas that are well represented with red noise alone (i.e., not significant).

~~Nonetheless~~Despite some differences between the MEMD and EMD modes of variability (primarily due to different input timeseries), the two methods agree on the quasi-periodic timescale of 2-3 years in the Niño3 region. This is also consistent with the significant periods inferred from the usual 1-D wavelet transform (not shown) and the power spectrum analysis of the
1215 Niño3 index (1-D) obtained via Fourier Transform (Fig. B3~~;~~). The latter confirms that 2-3 year timescale is quasi-periodic as there is a significant peak in the power spectrum of the Niño3 index on those timescales, i.e., black solid line ~~(power spectrum of~~ in Fig. B3 (Niño3 index) is above the red dashed line (red noise one-tailed 95% threshold))~~, as well as via a 1-D wavelet transform (not shown). ~~.
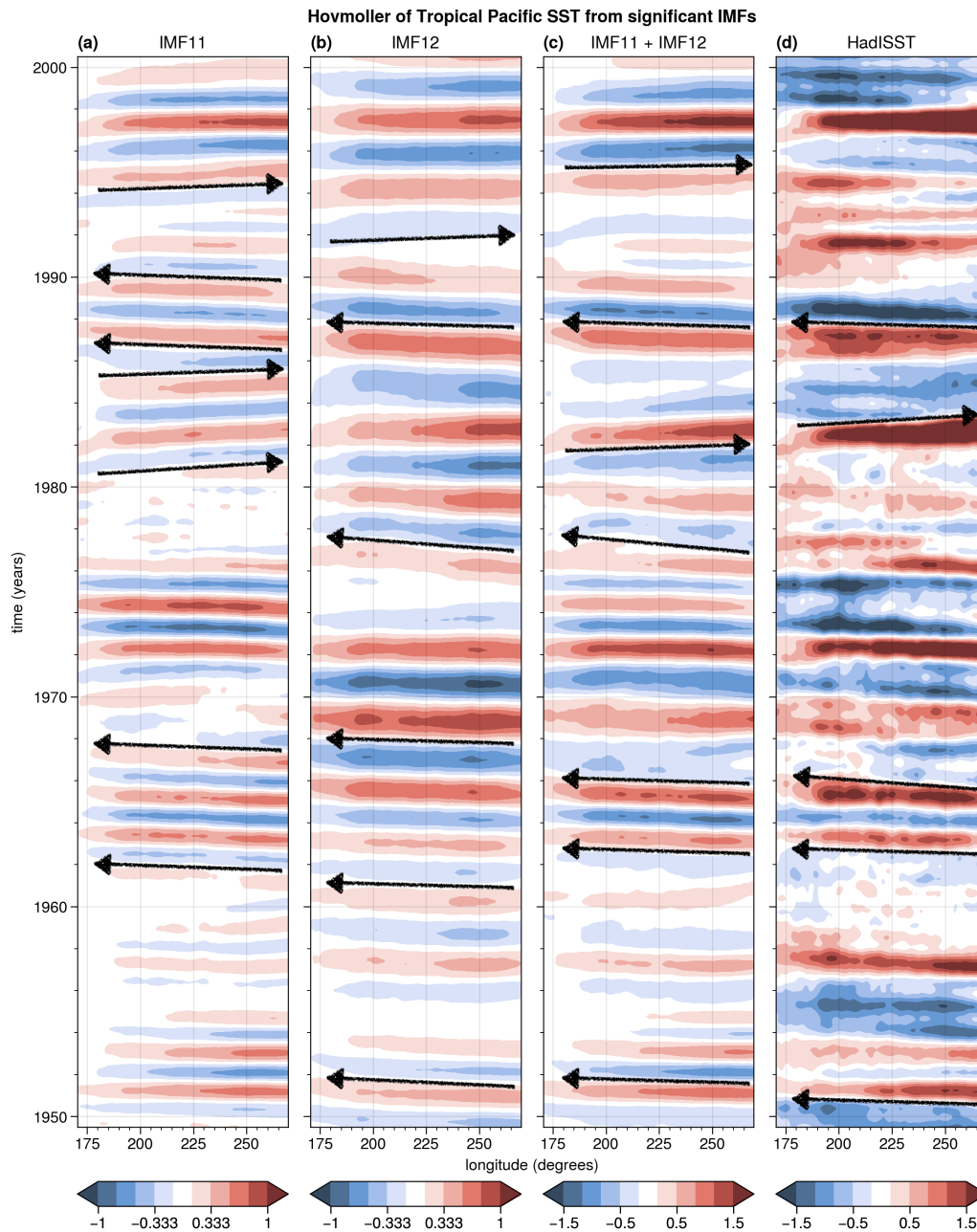
Figure 4. Time-longitude Hovmöller diagram of tropical Pacific SST anomalies (in K) averaged between 5°S and 5°N from (a) IMF11, (b) IMF12, (c) IMF11 + IMF12, and (d) 1-year low-passed (via 5th order Butterworth filter) SST data. Black arrows help visualise the direction of propagation of SST anomalies in some periods (other periods are more stationary). We show the time period around year 1970 (1950-2000) where a change in propagation direction has been identified in previous work. Note that colourscale in (a)-(b) is smaller than in (c)-(d).
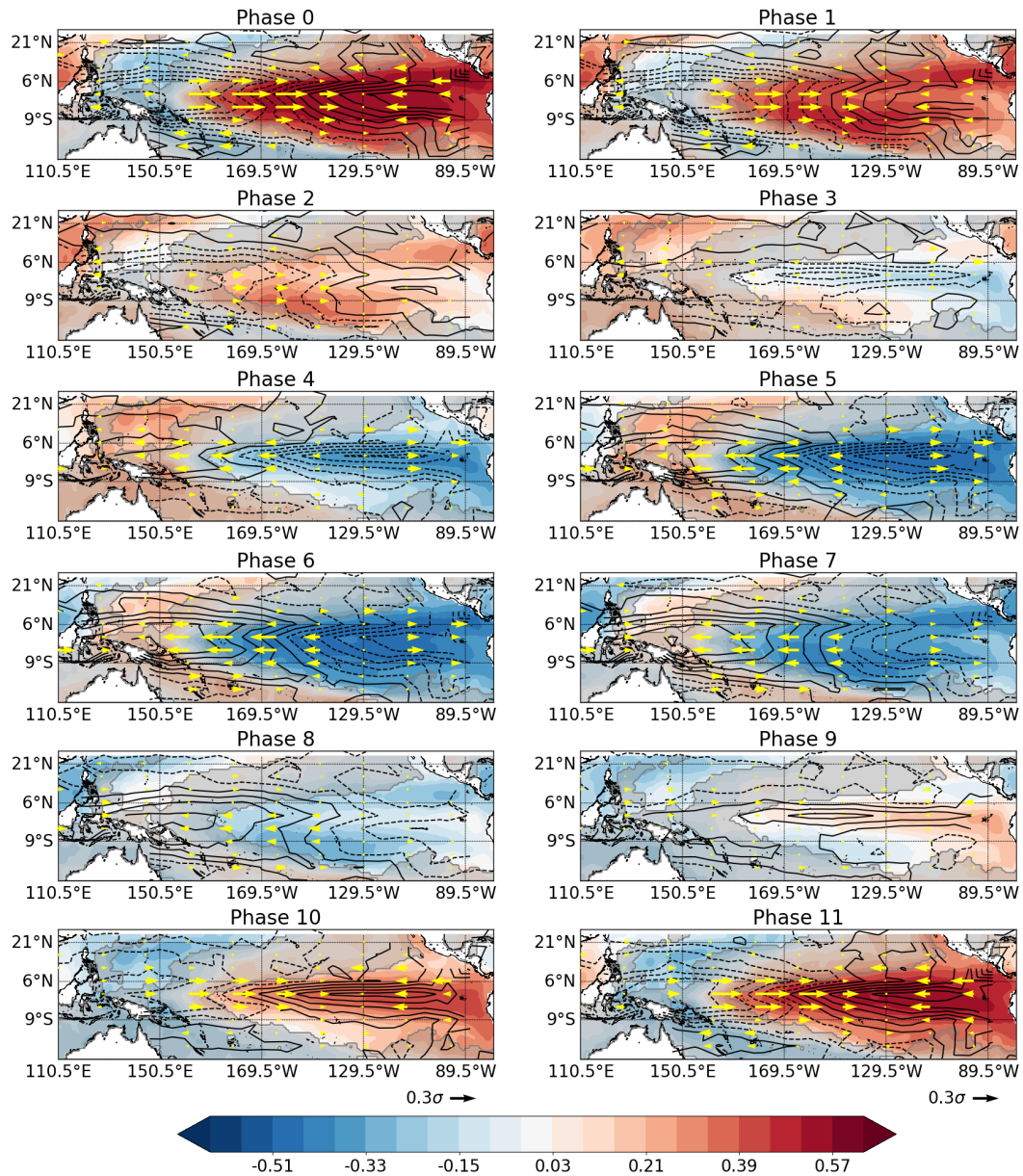
Figure 5. Latitude-longitude phase composite (phases 0 to 11 as labelled) of ~~IMF13~~IMF12: shading for ~~SSTs~~SST, contours for thermocline depth (contour interval is the same as in the colourbar with solid contours representing positive values, and dashed contours represent negative values), and arrows for τ_x (the scale is shown in the bottom left corner of panels for phases 10,11). All data is standardised and all fields were composited based on the phase of the eastern Pacific SST (Niño3) index.
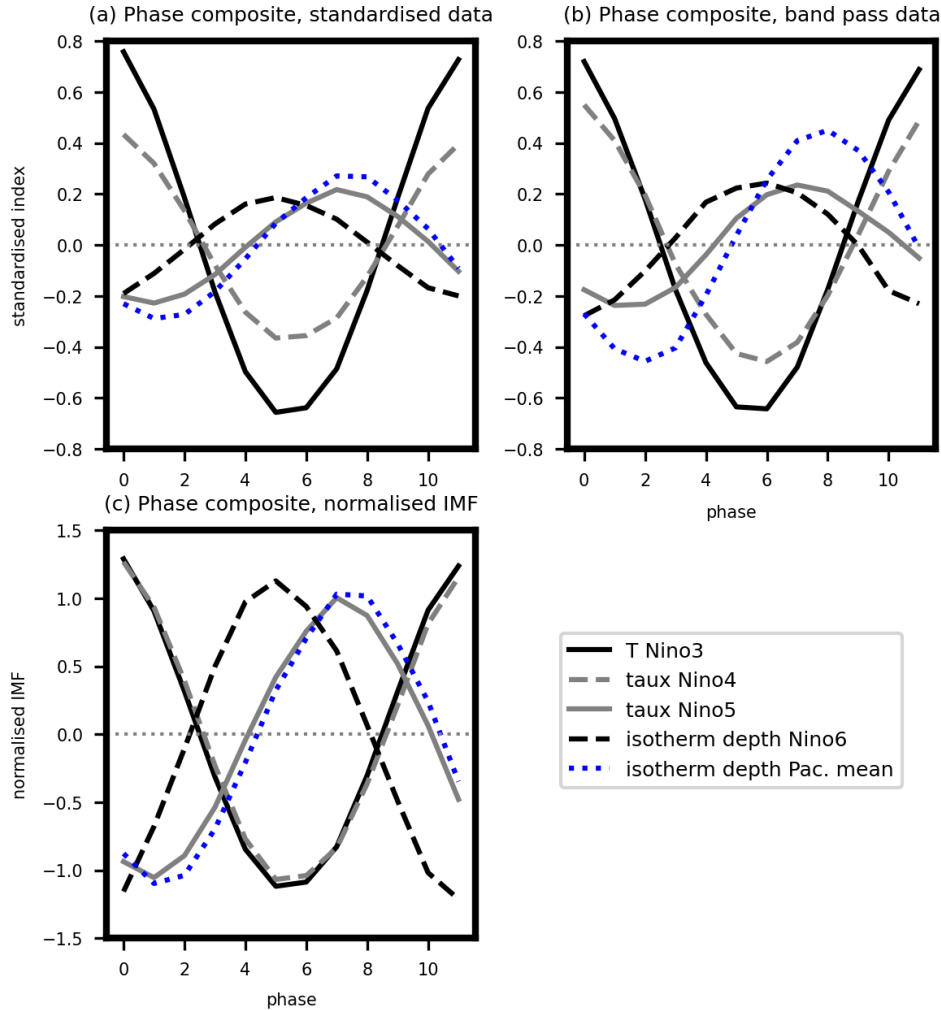
Figure 6. Phase composites of eastern Pacific SST (Niño3) (black solid line), off-equatorial western Pacific isotherm/thermocline depth (Niño6) (black dashed line), Pacific mean isotherm/thermocline depth (~~across tropical Pacific)~~ (blue dotted line), central Pacific τ_x (Niño4) (grey dashed line), western Pacific τ_x (Niño5) (grey solid line). All fields are composited over the phases of eastern Pacific SST (Niño3), such that they fit the phases in Fig. 5. (a) composites of ~~IMF13~~ IMF12 for data divided by the standard deviation of corresponding timeseries (e.g., ~~IMF13~~ IMF12 (thermocline)/σ (thermocline)); (b) composites of band-pass filtered (~~25-52.5~~ 29-53 months) standardised timeseries; (c) as in (a) but IMF-timeseries are divided by IMF's standard deviation (e.g., ~~IMF13~~ IMF12 (thermocline)/σ (~~IMF13~~ IMF12 of thermocline))~~; (d) composites of the fields from the conceptual oscillator timeseries (Eqs.(??)-(??)), divided by standard deviations of corresponding timeseries (similar to (c)).~~
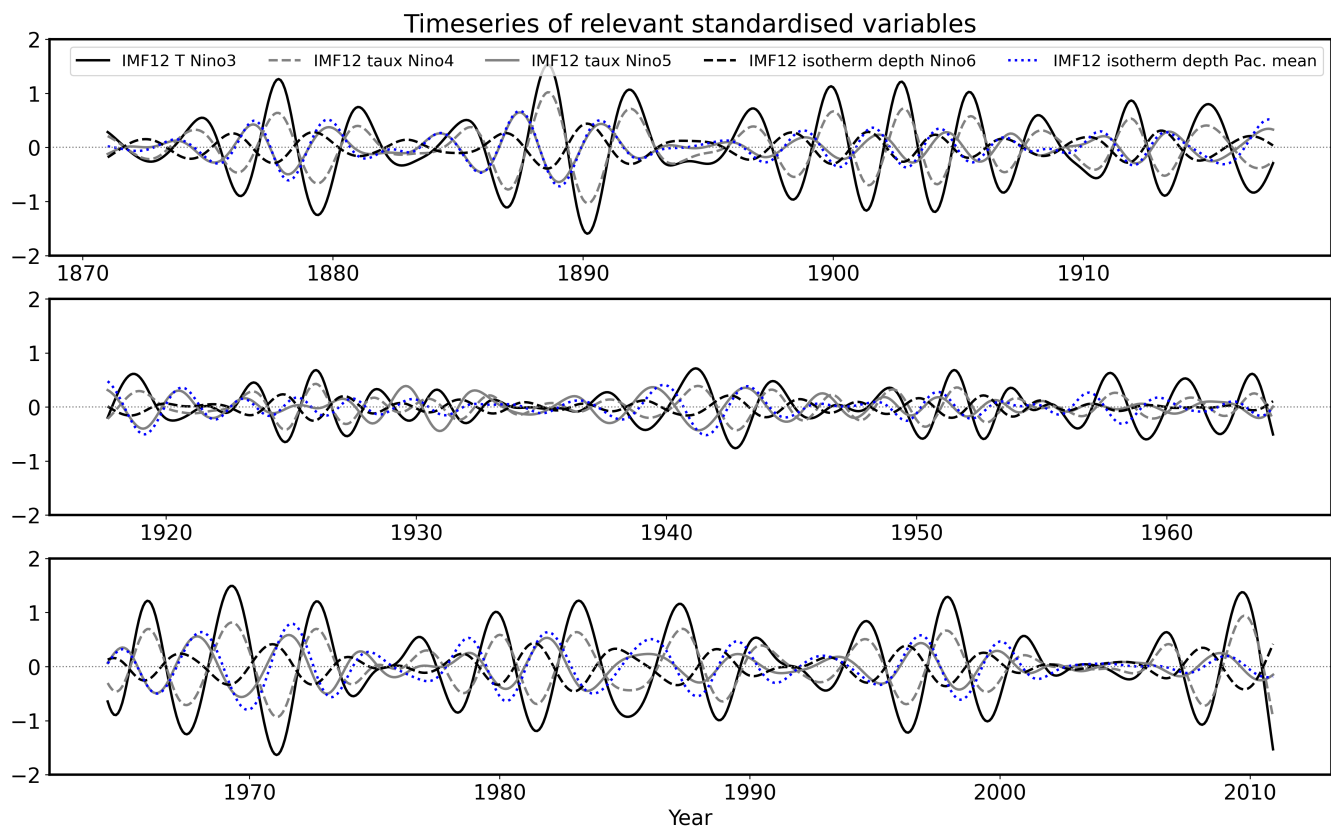
Figure 7. IMF12's standardised timeseries of eastern Pacific SST (Niño3) (black solid line), off-equatorial western Pacific thermocline depth (Niño6) (black dashed line), Pacific mean thermocline depth (blue dotted line), central Pacific τ_x (Niño4) (grey dashed line), western Pacific τ_x (Niño5) (grey solid line).
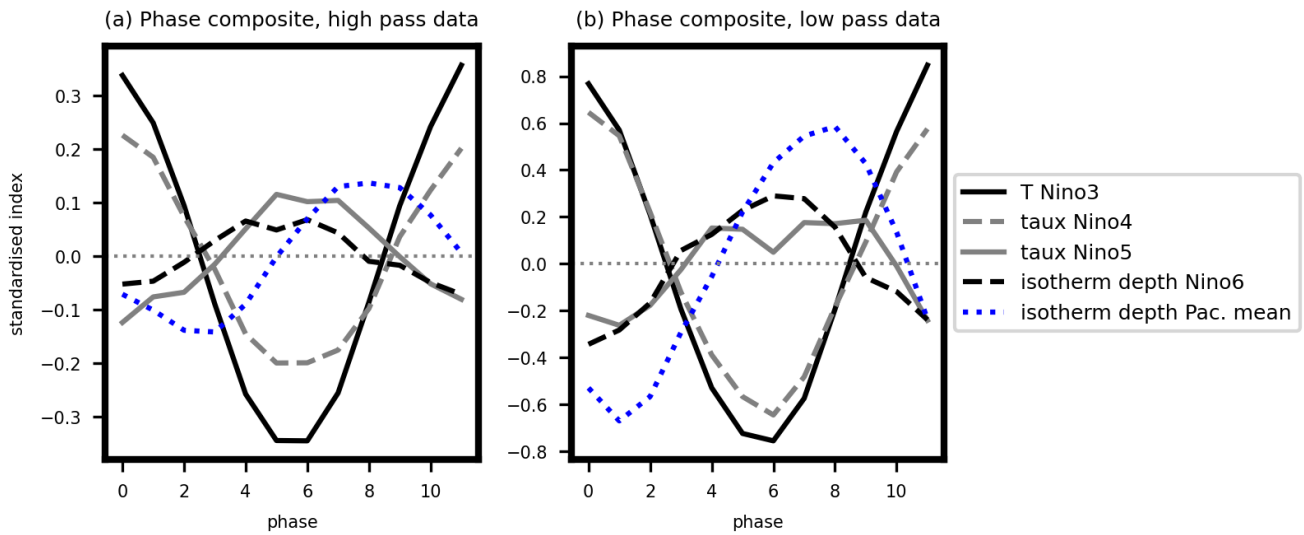
Figure 8. As in Fig. 6b, but for (a) higher frequency (band-pass filter over 12-19 months) data and (b) lower frequency (band-pass filter over 42-135 months) data. Note that higher and lower frequency timescale bands were chosen based on timescale ranges of 10th and 13th-14th IMFs, i.e., the IMFs with slightly smaller or larger (respectively) timescales than IMF11, IMF12.
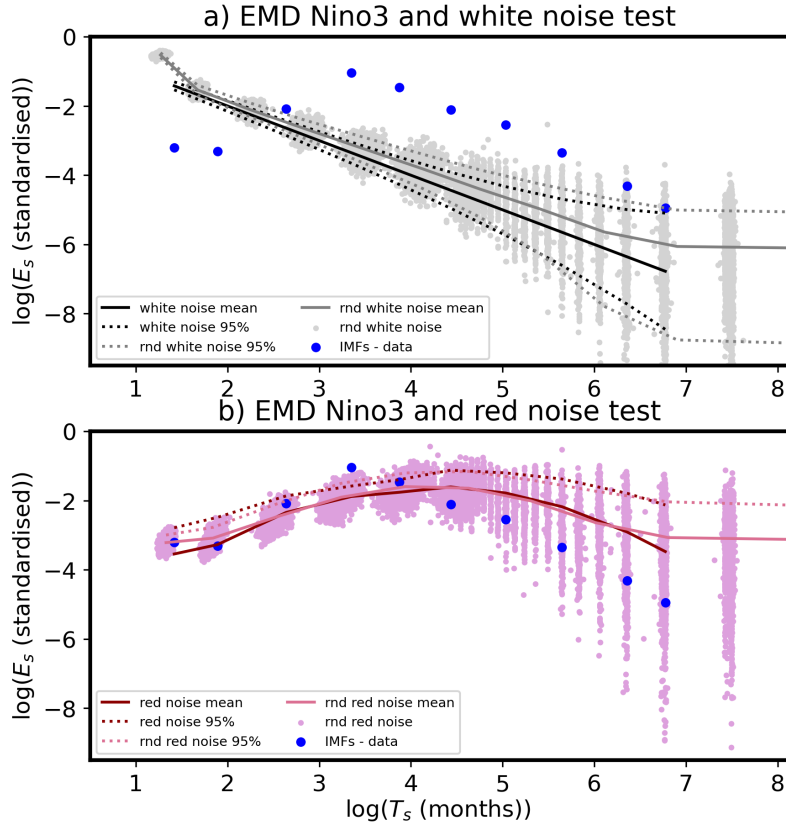
Figure B1. Significance tests for EMD modes: (a) white noise significance test and (b) red noise significance test for EMD-IMFs of Niño3 index (blue dots). Blue dots are computed as in Fig. 2. (a) Black solid line represents the theoretical linear relationship between the logarithms of period (T_s) and logarithm of energy density (E_s; i.e., average squared amplitude; Eq. (B3)), black dotted line represents 5th-95th percentile (Eq. B5), respectively; grey dots represent $I = L_t$ realisations of IMFs of white noise timeseries (length is the same as for Niño3 index), whereas grey solid and dotted lines represent their mean and the 5th-95th percentile, respectively. (b) Red solid line represents the theoretical red spectrum energy density (Eqs. (B7-B9)), red dotted line represents the 95th percentile (via χ^2-test); light pink dots represent $I = L_t$ realisations of IMFs of red noise timeseries (Eq. B6; length L_t is the same as for Niño3 index), whereas pink solid and dotted lines represent their mean and the 95th percentile, respectively. Note that x-axis shows the logarithms of period ($\log_e(T_s)$) ordered from shortest period (highest frequency) to longest period (trend). For further descriptions of the figure see text.
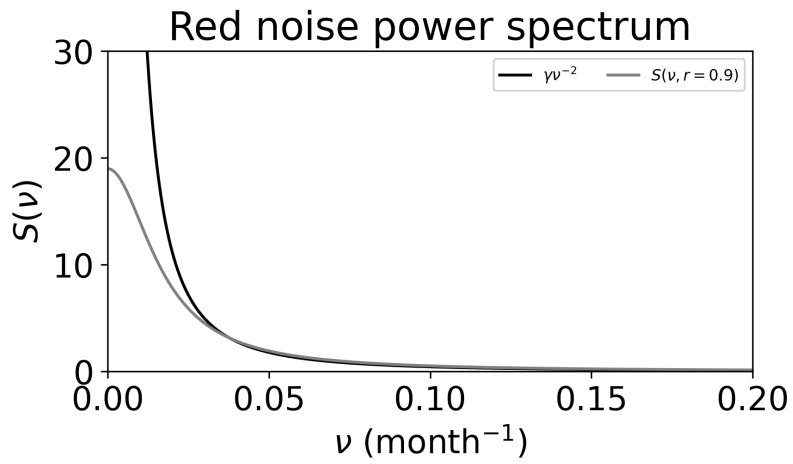
Figure B2. Red noise power spectrum ($S(\nu)$) for (black line) $S(\nu) = \gamma \nu^{-2}$, and (grey line) $S(\nu, r)$ from Eq. (B7) for $r = 0.9$. γ was estimated as a ratio between the integrated power spectra of the two spectra for frequencies higher than 0.02/month ($\gamma = \sum_\nu \nu^{-2} / \sum_\nu S(\nu, r)$) where the two power spectra generally agree well.
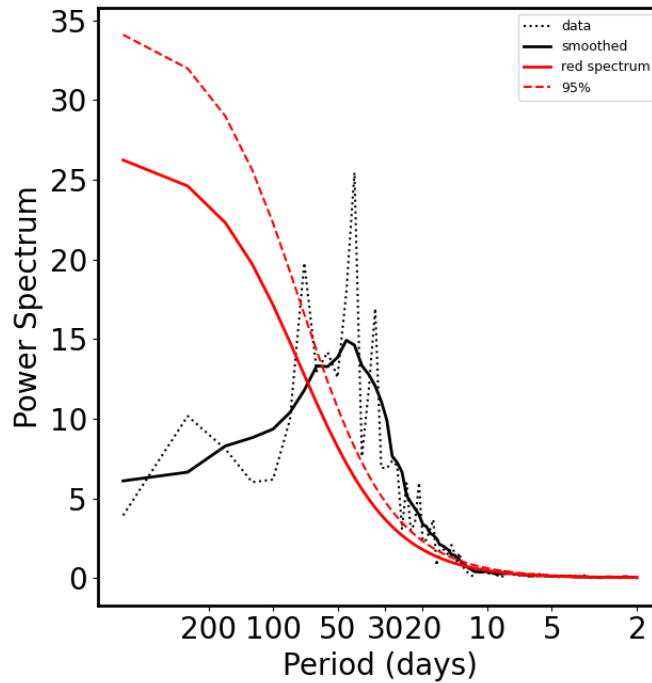
Figure B3. Power spectrum of Niño3 index. The power spectra are first computed for 500-months long chunks (overlapped by 250 months) and then averaged over all cases (grey solid line). The black solid line represents a 10-point running mean of the black dotted line (to increase the number of degrees of freedom, which is $f_\omega L_t/0.5L_{chunk} = 10 \times 1680/250 \approx 67$; see also Boljka et al. 2021). The red solid and dashed lines represent the theoretical red noise test and its (one-tailed) 95th percentile, respectively.

Appendix C: Simple statistical prediction model

A schematic of the procedure for lagging the data for predictors (relative to a predictand, SST (Niño3)), before inputting them into the simple (multi-linear) statistical prediction model. Black solid lines represent idealised SST (Niño3), black dotted lines represent idealised τ_x(Niño5), and black dashed lines represent idealised thermocline depth (Niño6). Vertical lines in (a) represent different lags (as specified). Similarly, stars represent locations of maximum (dark red) and minimum (light green) correlations of different variables with SST (Niño3). These peaks/minima are then lagged according to the identified lag of the maximum/minimum correlation as corresponding arrows show, i.e., they are "moved" towards lag 0. This yields two timeseries from each variable (see transition between (a) and (b)), where the minimum correlations are multiplied by (-1). (b) shows the predictors data as they were input into the statistical prediction model – when using only one past timestep. Note that top black solid line here is lagged by a specific lead time relative to the predictand (if lead time is 0 then predictand equals the top line). (c) shows how each of the lines from (b) (top line in (c)) can be further lagged with its past timesteps (light green squares and subsequent lines in (c) – see arrows). This yields up to 19 more timeseries from each timeseries in (b). For more details see the text.

The simple (multi-linear regression) statistical prediction model is constructed as follows. We first prepare timeseries that we can input into the statistical prediction model based on multi-linear regression the following way (Fig. **??**; see also Omrani et al. 2022). First, we lag the input timeseries relative to the SST (Niño3) based on the lag of the maximum correlation of data band-pass-filtered (smoothed) in 25-52.5 months band. We capture both the first maximum (i.e., the strongest positive correlation) and minimum (i.e., the strongest negative correlation) of the correlation between the different timeseries, which includes a full wave/oscillation (Fig. **??**a). This generates 6 predictors for the SST (Niño3) (the predictand): SST (Niño3) max at lag 0 months, SST (Niño3) min at lag -14 months, τ_x (Niño5) max at lag -9 months, τ_x (Niño5) min at lag -34 months, thermocline depth (Niño6) min at lag -13 months, and thermocline depth (Niño6) max at lag -37 months. Note that this yields better results than if only half of the lags (i.e., half of the oscillation) were used (not shown). The results from the prediction based on these 6 predictors are compared with the prediction of SST (Niño3) from SST (Niño3) at lag 0 alone.

By lagging the data we generate timeseries of 6 predictors and the predictand that have maximum correlations at lag zero (Fig. **??**b). We can then input the timeseries into the multi-linear regression model and predict SST (Niño3) at different lead times (1, 2, 3, ... , 30 months). For each lead time we construct a new prediction model.

To improve predictions we use the predictors' data from several past timesteps (i.e., use each of the predictors at months ..., -19, -18, ..., -1 relative to the predictand; Fig. **??**c). This can increase the 'number' of predictors significantly, but must be used carefully (i.e., sensitivity to the number of the past timesteps used should be tested).

The above yields the predictions of SST (Niño3) for different lead times and can then be used to assess the skill of the prediction – see section **??**.

1250   *Author contributions.*   LB performed the analysis, prepared the figures, and wrote the first draft of the manuscript. NEO and NSK provided additional insight and helped improve the manuscript for the final version.

Competing interests.   The authors declare no competing interests.

Acknowledgements.   This work was supported by the Trond Mohn ~~foundation~~ Foundation (project BCPU, grant number ~~BFS2018TMT01~~BFS3018TMT01

1255   and was performed on NIRD/Sigma2 (project NS9039K). We thank three anonymous reviewers and William Roberts for their constructive comments that helped improve the original manuscript. We also thank Lander Crespo for helpful discussions, and Ingo Bethke for the help with the data.

References

Alberti, T., Donner, R. V., and Vannitsem, S.: Multiscale fractal dimension analysis of a reduced order model of coupled ocean–atmosphere dynamics, Earth System Dynamics, 12, 837–855, https://doi.org/10.5194/esd-12-837-2021, 2021.

Allan, R. J.: ENSO and Climatic Variability in the Past 150 Years, p. 3–56, Cambridge University Press, https://doi.org/10.1017/CBO9780511573125.002, 2000.

An, S.-I., Tziperman, E., Okumura, Y. M., and Li, T.: ENSO Irregularity and Asymmetry, chap. 7, pp. 153–172, American Geophysical Union (AGU), https://doi.org/https://doi.org/10.1002/9781119548164.ch7, 2020.

Ashok, K., Behera, S. K., Rao, S. A., Weng, H., and Yamagata, T.: El Niño Modoki and its possible teleconnection, Journal of Geophysical Research: Oceans, 112, https://doi.org/https://doi.org/10.1029/2006JC003798, 2007.

Baede, A., Ahlonsou, E., Ding, Y., and Schimel, D.: The Climate System: an Overview, pp. 1–881, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://www.ipcc.ch/report/ar3/wg1/chapter-1-the-climate-system-an-overview/, 2001.

Battisti, D. S. and Hirst, A. C.: Interannual Variability in a Tropical Atmosphere–Ocean Model: Influence of the Basic State, Ocean Geometry and Nonlinearity, Journal of Atmospheric Sciences, 46, 1687 – 1712, https://doi.org/10.1175/1520-0469(1989)046<1687:IVIATA>2.0.CO;2, 1989.

Bejarano, L. and Jin, F.-F.: Coexistence of Equatorial Coupled Modes of ENSO, Journal of Climate, 21, 3051 – 3067, https://doi.org/10.1175/2007JCLI1679.1, 2008.

Bethke, I., Wang, Y., Counillon, F., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H. R., Chiu, P.-G., Bentsen, M., Guo, C., Tjiputra, J., Kirkevåg, A., Oliviè, D. J. L., Seland, y., Fan, Y., Lawrence, P., Eldevik, T., and Keenlyside, N.: NCC NorCPM1 model output prepared for CMIP6 CMIP historical, https://doi.org/10.22033/ESGF/CMIP6.10894, 2019.

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H., Svendsen, L., Chiu, P.-G., Passos, L., Bentsen, M., Guo, C., Gupta, A., Tjiputra, J., Kirkevåg, A., Olivié, D., Seland, Ø., Solsvik Vågane, J., Fan, Y., and Eldevik, T.: NorCPM1 and its contribution to CMIP6 DCPP, Geoscientific Model Development, 14, 7073–7116, https://doi.org/10.5194/gmd-14-7073-2021, 2021.

Bjerknes, J.: ATMOSPHERIC TELECONNECTIONS FROM THE EQUATORIAL PACIFIC, Monthly Weather Review, 97, 163 – 172, https://doi.org/10.1175/1520-0493(1969)097<0163:ATFTEP>2.3.CO;2, 1969.

Boljka, L., Thompson, D. W. J., and Li, Y.: Downstream Suppression of Baroclinic Waves, Journal of Climate, 34, 919 – 930, https://doi.org/10.1175/JCLI-D-20-0483.1, 2021.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., and Bladé, I.: The Effective Number of Spatial Degrees of Freedom of a Time-Varying Field, Journal of Climate, 12, 1990 – 2009, https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2, 1999.

Broomhead, D. S., Jones, R., and King, G. P.: Topological dimension and local coordinates from time series data, Journal of Physics A: Mathematical and General, 20, L563–L569, https://doi.org/10.1088/0305-4470/20/9/003, 1987.

Brönnimann, S.: Impact of El Niño–Southern Oscillation on European climate, Reviews of Geophysics, 45, https://doi.org/https://doi.org/10.1029/2006RG000199, 2007.

Burgers, G., Jin, F.-F., and van Oldenborgh, G. J.: The simplest ENSO recharge oscillator, Geophysical Research Letters, 32, L13 706, https://doi.org/https://doi.org/10.1029/2005GL022951, 2005.

Cane, M., Zebiak, S., and Dolan, S.: Experimental forecasts of El Niño, Nature, 321, 827–832, https://doi.org/10.1038/321827a0, 1986.

Capotondi, A., Sardeshmukh, P. D., and Ricciardulli, L.: The Nature of the Stochastic Wind Forcing of ENSO, Journal of Climate, 31, 8081 – 8099, https://doi.org/10.1175/JCLI-D-17-0842.1, 2018.

1295 Carton, J. A. and Giese, B. S.: A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA), Monthly Weather Review, 136, 2999 – 3017, https://doi.org/10.1175/2007MWR1978.1, 2008.

Cashin, P., Mohaddes, K., and Raissi, M.: Fair weather or foul? The macroeconomic effects of El Niño, Journal of International Economics, 106, 37–54, https://doi.org/https://doi.org/10.1016/j.jinteco.2017.01.010, 2017.

Chen, D., Cane, M., Kaplan, A., Zebiak, S. E., and Huang, D.: Predictability of El Niño over the past 148 years, Nature, 428, 733–736,
1300 https://doi.org/10.1038/nature02439, 2004.

Chen, H.-C., Jin, F.-F., Zhao, S., Wittenberg, A. T., and Xie, S.: ENSO Dynamics in the E3SM-1-0, CESM2, and GFDL-CM4 Climate Models, Journal of Climate, pp. 1 – 59, https://doi.org/10.1175/JCLI-D-21-0355.1, 2021.

Choi, K.-Y., Vecchi, G. A., and Wittenberg, A. T.: ENSO Transition, Duration, and Amplitude Asymmetries: Role of the Nonlinear Wind Stress Coupling in a Conceptual Model, Journal of Climate, 26, 9462 – 9476, https://doi.org/https://doi.org/10.1175/JCLI-D-13-00045.1,
1305 2013.

Clement, A., DiNezio, P., and Deser, C.: Rethinking the Ocean's Role in the Southern Oscillation, Journal of Climate, 24, 4056 – 4072, https://doi.org/10.1175/2011JCLI3973.1, 2011.

Crespo, L. R., Rodríguez-Fonseca, M. B., Polo, I., Keenlyside, N., and Dommenget, D.: Multidecadal variability of ENSO in a recharge oscillator framework, Environmental Research Letters, 17, 074 008, https://doi.org/10.1088/1748-9326/ac72a3, 2022.

1310 Dai, A. and Wigley, T. M. L.: Global patterns of ENSO-induced precipitation, Geophysical Research Letters, 27, 1283–1286, https://doi.org/https://doi.org/10.1029/1999GL011140, 2000.

de la Cámara, A., Birner, T., and Albers, J. R.: Are Sudden Stratospheric Warmings Preceded by Anomalous Tropospheric Wave Activity?, Journal of Climate, 32, 7173–7189, https://doi.org/10.1175/JCLI-D-19-0269.1, 2019.

Deser, C. and Wallace, J. M.: Large-Scale Atmospheric Circulation Features of Warm and Cold Episodes in the Tropical Pacific, Journal of
1315 Climate, 3, 1254 – 1281, https://doi.org/10.1175/1520-0442(1990)003<1254:LSACFO>2.0.CO;2, 1990.

Dijkstra, H. A., Petersik, P., Hernández-García, E., and López, C.: The Application of Machine Learning Techniques to Improve El Niño Prediction Skill, Frontiers in Physics, 7, 153, https://doi.org/10.3389/fphy.2019.00153, 2019.

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

1320 Duffy, D. G.: The Application of Hilbert–Huang Transforms to Meteorological Datasets, Journal of Atmospheric and Oceanic Technology, 21, 599 – 611, https://doi.org/10.1175/1520-0426(2004)021<0599:TAOHTT>2.0.CO;2, 2004.

Enfield, D. B. and Mestas-Nuñez, A. M.: Multiscale Variabilities in Global Sea Surface Temperatures and Their Relationships with Tropospheric Climate Patterns, Journal of Climate, 12, 2719 – 2733, https://doi.org/10.1175/1520-0442(1999)012<2719:MVIGSS>2.0.CO;2, 1999.

1325 Ezer, T. and Corlett, W. B.: Is sea level rise accelerating in the Chesapeake Bay? A demonstration of a novel new approach for analyzing sea level data, Geophysical Research Letters, 39, https://doi.org/https://doi.org/10.1029/2012GL053435, 2012.

Ezer, T., Atkinson, L. P., Corlett, W. B., and Blanco, J. L.: Gulf Stream's induced sea level rise and variability along the U.S. mid-Atlantic coast, Journal of Geophysical Research: Oceans, 118, 685–697, https://doi.org/https://doi.org/10.1002/jgrc.20091, 2013.

Fedorov, A. V. and Philander, S. G.: Is El Niño Changing?, Science, 288, 1997–2002, https://doi.org/10.1126/science.288.5473.1997, 2000.

1330   Fedorov, A. V. and Philander, S. G.: A Stability Analysis of Tropical Ocean–Atmosphere Interactions: Bridging Measurements and Theory for El Niño, Journal of Climate, 14, 3086 – 3101, https://doi.org/https://doi.org/10.1175/1520-0442(2001)014<3086:ASAOTO>2.0.CO;2, 2001.

Fedorov, A. V., Harper, S. L., Philander, S. G., Winter, B., and Wittenberg, A.: How Predictable is El Niño?, Bulletin of the American Meteorological Society, 84, 911 – 920, https://doi.org/10.1175/BAMS-84-7-911, 2003.

1335   Fereday, D. R., Knight, J. R., Scaife, A. A., Folland, C. K., and Philipp, A.: Cluster Analysis of North Atlantic–European Circulation Types and Links with Tropical Pacific Sea Surface Temperatures, Journal of Climate, 21, 3687 – 3703, https://doi.org/10.1175/2007JCLI1875.1, 2008.

Flandrin, P., Rilling, G., and Goncalves, P.: Empirical mode decomposition as a filter bank, IEEE Signal Processing Letters, 11, 112–114, https://doi.org/10.1109/LSP.2003.821662, 2004.

1340   Frankignoul, C. and Hasselmann, K.: Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability, Tellus, 29, 289–305, https://doi.org/10.3402/tellusa.v29i4.11362, 1977.

Franzke, C.: Multi-scale analysis of teleconnection indices: climate noise and nonlinear trend analysis, Nonlinear Processes in Geophysics, 16, 65–76, https://doi.org/10.5194/npg-16-65-2009, 2009.

Franzke, C.: Nonlinear Trends, Long-Range Dependence, and Climate Noise Properties of Surface Temperature, Journal of Climate, 25,
1345   4172 – 4183, https://doi.org/10.1175/JCLI-D-11-00293.1, 2012.

Franzke, C. and Woollings, T.: On the Persistence and Predictability Properties of North Atlantic Climate Variability, Journal of Climate, 24, 466 – 472, https://doi.org/10.1175/2010JCLI3739.1, 2011.

Froyland, G., Giannakis, D., Lintner, B., Pike, M., and Slawinska, J.: Spectral analysis of climate dynamics with operator-theoretic approaches, https://doi.org/10.1038/s41467-021-26357-x, 2021.

1350   Ghil, M. and Jiang, N.: Recent forecast skill for the El Niño/Southern Oscillation, Geophysical Research Letters, 25, 171–174, https://doi.org/https://doi.org/10.1029/97GL03635, 1998.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F., and Yiou, P.: ADVANCED SPECTRAL METHODS FOR CLIMATIC TIME SERIES, Reviews of Geophysics, 40, 3–1–3–41, https://doi.org/https://doi.org/10.1029/2000RG000092, 2002.

1355   Gill, A. E.: Some simple solutions for heat-induced tropical circulation, Quarterly Journal of the Royal Meteorological Society, 106, 447–462, https://doi.org/https://doi.org/10.1002/qj.49710644905, 1980.

Gilman, D. L., Fuglister, F. J., and Mitchell, J. M.: On the power spectrum of "red noise.", J. Atmos. Sci., 20, 182–184, 1963.

Graham, F. S., Brown, J. N., Wittenberg, A. T., and Holbrook, N. J.: Reassessing Conceptual Models of ENSO, Journal of Climate, 28, 9121 – 9142, https://doi.org/10.1175/JCLI-D-14-00812.1, 2015.

1360   Guilyardi, E., Capotondi, A., Lengaigne, M., Thual, S., and Wittenberg, A. T.: ENSO Modeling, chap. 9, pp. 199–226, American Geophysical Union (AGU), https://doi.org/https://doi.org/10.1002/9781119548164.ch9, 2020.

Guimarães Nobre, G., Muis, S., Veldkamp, T. I., and Ward, P. J.: Achieving the reduction of disaster risk by better predicting impacts of El Niño and La Niña, Progress in Disaster Science, 2, 100 022, https://doi.org/https://doi.org/10.1016/j.pdisas.2019.100022, 2019.

Ham, Y., Kim, J., and Luo, J.: Deep learning for multi-year ENSO forecasts, Nature, 573, 568–572, https://doi.org/10.1038/s41586-019-
1365   1559-7, 2019.

Ham, Y.-G., Kim, J.-H., Kim, E.-S., and On, K.-W.: Unified deep learning model for El Niño/Southern Oscillation forecasts by incorporating seasonality in climate data, Science Bulletin, 66, 1358–1366, https://doi.org/https://doi.org/10.1016/j.scib.2021.03.009, 2021.

Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Ineson, S., Lim, J., and Fereday, D.: The Impact of Strong El Niño and La Niña Events on the North Atlantic, Geophysical Research Letters, 46, 2874–2883, https://doi.org/https://doi.org/10.1029/2018GL081776, 2019.

Hasselmann, K.: Stochastic climate models Part I. Theory, Tellus, 28, 473–485, https://doi.org/https://doi.org/10.1111/j.2153-3490.1976.tb00696.x, 1976.

Hasselmann, K.: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns, Journal of Geophysical Research: Atmospheres, 93, 11 015–11 021, https://doi.org/https://doi.org/10.1029/JD093iD09p11015, 1988.

Hu, Z.-Z., Kumar, A., Zhu, J., Huang, B., Tseng, Y.-h., and Wang, X.: On the Shortening of the Lead Time of Ocean Warm Water Volume to ENSO SST Since 2000, Scientific Reports, 7, 4294, https://doi.org/10.1038/s41598-017-04566-z, 2017.

Hu, Z.-Z., Kumar, A., Huang, B., Zhu, J., L'Heureux, M., McPhaden, M. J., and Yu, J.-Y.: The Interdecadal Shift of ENSO Properties in 1999/2000: A Review, Journal of Climate, 33, 4441 – 4462, https://doi.org/10.1175/JCLI-D-19-0316.1, 2020.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 454, 903–995, https://doi.org/10.1098/rspa.1998.0193, 1998.

Huang, N. E., Shen, Z., and Long, S. R.: A NEW VIEW OF NONLINEAR WATER WAVES: The Hilbert Spectrum, Annual Review of Fluid Mechanics, 31, 417–457, https://doi.org/10.1146/annurev.fluid.31.1.417, 1999.

Huang, N. E., Wu, M.-L. C., Long, S. R., Shen, S. S., Qu, W., Gloersen, P., and Fan, K. L.: A confidence limit for the empirical mode decomposition and Hilbert spectral analysis, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 459, 2317–2345, https://doi.org/10.1098/rspa.2003.1123, 2003.

Jajcay, N., Kravtsov, S., Sugihara, G., Tsonis, A. A., and Paluš, M.: Synchronization and causality across time scales in El Nño Southern Oscillation, npj Climate and Atmospheric Science, 1, 33, https://doi.org/10.1038/s41612-018-0043-7, 2018.

Jenney, A. M., Randall, D. A., and Barnes, E. A.: Quantifying Regional Sensitivities to Periodic Events: Application to the MJO, Journal of Geophysical Research: Atmospheres, 124, 3671–3683, https://doi.org/https://doi.org/10.1029/2018JD029457, 2019.

Jiang, N., Neelin, J. D., and Ghil, M.: Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific, Climate Dynamics, 12, 101–112, https://doi.org/10.1007/BF00223723, 1995.

Jiménez-Esteve, B. and Domeisen, D. I. V.: Nonlinearity in the tropospheric pathway of ENSO to the North Atlantic, Weather and Climate Dynamics, 1, 225–245, https://doi.org/10.5194/wcd-1-225-2020, 2020.

Jiménez-Esteve, B. and Domeisen, D. I. V.: The Tropospheric Pathway of the ENSO–North Atlantic Teleconnection, Journal of Climate, 31, 4563 – 4584, https://doi.org/10.1175/JCLI-D-17-0716.1, 2018.

Jin, F.-F.: An Equatorial Ocean Recharge Paradigm for ENSO. Part I: Conceptual Model, Journal of the Atmospheric Sciences, 54, 811 – 829, https://doi.org/10.1175/1520-0469(1997)054<0811:AEORPF>2.0.CO;2, 1997a.

Jin, F.-F.: An Equatorial Ocean Recharge Paradigm for ENSO. Part II: A Stripped-Down Coupled Model, Journal of the Atmospheric Sciences, 54, 830 – 847, https://doi.org/10.1175/1520-0469(1997)054<0830:AEORPF>2.0.CO;2, 1997b.

Kao, H.-Y. and Yu, J.-Y.: Contrasting Eastern-Pacific and Central-Pacific Types of ENSO, Journal of Climate, 22, 615 – 632, https://doi.org/10.1175/2008JCLI2309.1, 2009.

Keenlyside, N. S., Latif, M., and Dürkop, A.: On Sub-ENSO Variability, Journal of Climate, 20, 3452 – 3469, https://doi.org/10.1175/JCLI4199.1, 2007.

Kim, K.-Y., O'Brien, J. J., and Barcilon, A. I.: The Principal Physical Modes of Variability over the Tropical Pacific, Earth Interactions, 7, 1 – 32, https://doi.org/10.1175/1087-3562(2003)007<0001:TPPMOV>2.0.CO;2, 2003.

Kolotkov, D. Y., Anfinogentov, S. A., and Nakariakov, V. M.: Empirical mode decomposition analysis of random processes in the solar atmosphere, Astronomy & Astrophysics, 592, A153, https://doi.org/10.1051/0004-6361/201628306, 2016.

Kretschmer, M., Adams, S. V., Arribas, A., Prudden, R., Robinson, N., Saggioro, E., and Shepherd, T. G.: Quantifying causal pathways of teleconnections, Bulletin of the American Meteorological Society, pp. 1 – 34, https://doi.org/10.1175/BAMS-D-20-0117.1, 2021.

Lam, H. C. Y., Haines, A., McGregor, G., Chan, E. Y. Y., and Hajat, S.: Time-Series Study of Associations between Rates of People Affected by Disasters and the El Niño Southern Oscillation (ENSO) Cycle, International Journal of Environmental Research and Public Health, 16, https://doi.org/10.3390/ijerph16173146, 2019.

Lee, J., Planton, Y. Y., Gleckler, P. J., Sperber, K. R., Guilyardi, E., Wittenberg, A. T., McPhaden, M. J., and Pallotta, G.: Robust Evaluation of ENSO in Climate Models: How Many Ensemble Members Are Needed?, Geophysical Research Letters, 48, e2021GL095 041, https://doi.org/https://doi.org/10.1029/2021GL095041, 2021.

Lee, T. and Ouarda, T. B. M. J.: Prediction of climate nonstationary oscillation processes with empirical mode decomposition, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/https://doi.org/10.1029/2010JD015142, 2011.

L'Heureux, M. L., Levine, A. F. Z., Newman, M., Ganter, C., Luo, J.-J., Tippett, M. K., and Stockdale, T. N.: ENSO Prediction, chap. 10, pp. 227–246, American Geophysical Union (AGU), https://doi.org/https://doi.org/10.1002/9781119548164.ch10, 2020.

Lopez, H. and Kirtman, B. P.: WWBs, ENSO predictability, the spring barrier and extreme events, Journal of Geophysical Research: Atmospheres, 119, 10,114–10,138, https://doi.org/https://doi.org/10.1002/2014JD021908, 2014.

Lopez, H., Kirtman, B. P., Tziperman, E., and Gebbie, G.: Impact of interactive westerly wind bursts on CCSM3, Dynamics of Atmospheres and Oceans, 59, 24–51, https://doi.org/https://doi.org/10.1016/j.dynatmoce.2012.11.001, 2013.

Madden, R. A. and Julian, P. R.: Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific, Journal of Atmospheric Sciences, 28, 702 – 708, https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2, 1971.

Maloney, E. D., Chelton, D. B., and Esbensen, S. K.: Subseasonal SST Variability in the Tropical Eastern North Pacific during Boreal Summer, Journal of Climate, 21, 4149 – 4167, https://doi.org/10.1175/2007JCLI1856.1, 2008.

Mariotti, A., Ruti, P. M., and Rixen, M.: Progress in subseasonal to seasonal prediction through a joint weather and climate community effort, npj Climate and Atmospheric Science, 1, https://doi.org/10.1038/s41612-018-0014-z, 2018.

Mariotti, A., Baggett, C., Barnes, E. A., Becker, E., Butler, A., Collins, D. C., Dirmeyer, P. A., Ferranti, L., Johnson, N. C., Jones, J., Kirtman, B. P., Lang, A. L., Molod, A., Newman, M., Robertson, A. W., Schubert, S., Waliser, D. E., and Albers, J.: Windows of Opportunity for Skillful Forecasts Subseasonal to Seasonal and Beyond, Bulletin of the American Meteorological Society, 101, E608 – E625, https://doi.org/https://doi.org/10.1175/BAMS-D-18-0326.1, 2020.

McPhaden, M. J.: Playing hide and seek with El Niño, Nature Climate Change, 5, 791–795, https://doi.org/10.1038/nclimate2775, 2015.

Mestas-Nuñez, A. M. and Enfield, D. B.: Eastern Equatorial Pacific SST Variability: ENSO and Non-ENSO Components and Their Climatic Associations, Journal of Climate, 14, 391 – 402, https://doi.org/10.1175/1520-0442(2001)014<0391:EEPSVE>2.0.CO;2, 2001.

Omrani, N.-E., Keenlyside, N. S., Matthes, K., Boljka, L., Zanchettin, D., Jungclaus, J. H., and Lubis, S. W.: Coupled stratosphere-troposphere-Atlantic multidecadal oscillation and its importance for near-future climate projection, npj Climate and Atmospheric Science, 5, 59, https://doi.org/https://doi.org/10.1038/s41612-022-00275-1, 2022.

Park, J., Kug, J., Li, T., and Behera, S. K.: Predicting El Niño Beyond 1-year Lead: Effect of the Western Hemisphere Warm Pool, Scientific Reports, 8, 14 957, https://doi.org/10.1038/s41598-018-33191-7, 2018.

Penland, C. and Sardeshmukh, P. D.: The Optimal Growth of Tropical Sea Surface Temperature Anomalies, Journal of Climate, 8, 1999 – 2024, https://doi.org/10.1175/1520-0442(1995)008<1999:TOGOTS>2.0.CO;2, 1995.

Philander, S. G. H.: El Niño, La Niña, and the southern oscillation, Academic Press, Inc., San Diego, California, 1990.

Picaut, J., Masia, F., and du Penhoat, Y.: An Advective-Reflective Conceptual Model for the Oscillatory Nature of the ENSO, Science, 277, 663–666, https://doi.org/10.1126/science.277.5326.663, 1997.

Planton, Y. Y., Guilyardi, E., Wittenberg, A. T., Lee, J., Gleckler, P. J., Bayr, T., McGregor, S., McPhaden, M. J., Power, S., Roehrig, R., Vialard, J., and Voldoire, A.: Evaluating Climate Models with the CLIVAR 2020 ENSO Metrics Package, Bulletin of the American Meteorological Society, 102, E193 – E217, https://doi.org/10.1175/BAMS-D-19-0337.1, 2021.

Qian, C., Wu, Z., Fu, C., and Wang, D.: On Changing El Niño: A View from Time-Varying Annual Cycle, Interannual Variability, and Mean State, Journal of Climate, 24, 6486 – 6500, https://doi.org/10.1175/JCLI-D-10-05012.1, 2011.

Quinn, W. H.: Monitoring and Predicting El Niño Invasions, Journal of Applied Meteorology and Climatology, 13, 825 – 830, https://doi.org/10.1175/1520-0450(1974)013<0825:MAPENI>2.0.CO;2, 1974a.

Quinn, W. H.: Outlook for El Niño-like conditions in 1975, NORPAX Highlights, 2, 2–3, 1974b.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research, 108, D14, 4407, https://doi.org/10.1029/2002JD002670, 2003.

Rehman, N. and Mandic, D. P.: Multivariate empirical mode decomposition, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 466, 1291–1302, https://doi.org/10.1098/rspa.2009.0502, 2010.

Rehman, N. and Mandic, D. P.: Filter Bank Property of Multivariate Empirical Mode Decomposition, IEEE Transactions on Signal Processing, 59, 2421–2426, https://doi.org/10.1109/TSP.2011.2106779, 2011.

Rilling, G., Flandrin, P., and Gonçalves, P.: On empirical mode decomposition and its algorithms, in: Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado, Italy, https://hal.inria.fr/inria-00570628, 2003.

Rilling, G., Flandrin, P., Goncalves, P., and Lilly, J. M.: Bivariate Empirical Mode Decomposition, IEEE Signal Processing Letters, 14, 936–939, https://doi.org/10.1109/LSP.2007.904710, 2007.

Roulston, M. S. and Neelin, J. D.: The response of an ENSO Model to climate noise, weather noise and intraseasonal forcing, Geophysical Research Letters, 27, 3723–3726, https://doi.org/https://doi.org/10.1029/2000GL011941, 2000.

Runge, J., Petoukhov, V., Donges, J. F., Hlinka, J., Jajcay, N., Vejmelka, M., Hartman, D., Marwan, N., Paluš, M., and Kurths, J.: Identifying causal gateways and mediators in complex spatio-temporal systems, Nature Communications, 6, 8502, https://doi.org/10.1038/ncomms9502, 2015.

Santos, J. L.: The Impact of El Niño - Southern Oscillation Events on South America, https://doi.org/10.5194/adgeo-6-221-2006, 2006.

Singh, A. and Delcroix, T.: Eastern and Central Pacific ENSO and their relationships to the recharge/discharge oscillator paradigm, Deep Sea Research Part I: Oceanographic Research Papers, 82, 32–43, https://doi.org/https://doi.org/10.1016/j.dsr.2013.08.002, 2013.

Slawinska, J. and Giannakis, D.: Indo-Pacific Variability on Seasonal to Multidecadal Time Scales. Part I: Intrinsic SST Modes in Models and Observations, Journal of Climate, 30, 5265 – 5294, https://doi.org/10.1175/JCLI-D-16-0176.1, 2017.

Stallone, A., Cicone, A., and Materassi, M.: New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms, Scientific Reports, 10, 15 161, https://doi.org/10.1038/s41598-020-72193-2, 2020.

Stein, K., Schneider, N., Timmermann, A., and Jin, F.-F.: Seasonal Synchronization of ENSO Events in a Linear Stochastic Model, Journal of Climate, 23, 5629 – 5643, https://doi.org/10.1175/2010JCLI3292.1, 2010.

Suarez, M. J. and Schopf, P. S.: A Delayed Action Oscillator for ENSO, Journal of Atmospheric Sciences, 45, 3283 – 3287, https://doi.org/10.1175/1520-0469(1988)045<3283:ADAOFE>2.0.CO;2, 1988.

Sverdrup, H. U.: Wind-Driven Currents in a Baroclinic Ocean; with Application to the Equatorial Currents of the Eastern Pacific*, Proceedings of the National Academy of Sciences, 33, 318–326, https://doi.org/10.1073/pnas.33.11.318, 1947.

1485 Timmermann, A., An, S. I., Kug, J. S., Jin, F. F., Cai, W., Capotondi, A., Cobb, K. M., Lengaigne, M., McPhaden, M. J., Stuecker, M. F., Stein, K., Wittenberg, A. T., Yun, K. S., Bayr, T., Chen, H. C., Chikamoto, Y., Dewitte, B., Dommenget, D., Grothe, P., Guilyardi, E., Ham, Y. G., Hayashi, M., Ineson, S., Kang, D., Kim, S., Kim, W., Lee, J. Y., Li, T., Luo, J. J., McGregor, S., Planton, Y., Power, S., Rashid, H., Ren, H. L., Santoso, A., Takahashi, K., Todd, A., Wang, G., Wang, G., Xie, R., Yang, W. H., Yeh, S. W., Yoon, J., Zeller, E., and Zhang, X.: El Nino-Southern Oscillation complexity, Nature, 559, 535–545, https://doi.org/10.1038/s41586-018-0252-6, 2018.

1490 Wang, B. and An, S.-I.: Why the properties of El Niño changed during the late 1970s, Geophysical Research Letters, 28, 3709–3712, https://doi.org/https://doi.org/10.1029/2001GL012862, 2001.

Wang, C.: A Unified Oscillator Model for the El Niño-Southern Oscillation, Journal of Climate, 14, 98 – 115, https://doi.org/10.1175/1520-0442(2001)014<0098:AUOMFT>2.0.CO;2, 2001a.

Wang, C.: On the ENSO Mechanisms, ADVANCES IN ATMOSPHERIC SCIENCES, 18, 2001b.

1495 Wang, C.: A review of ENSO theories, National Science Review, 5, 813–825, https://doi.org/10.1093/nsr/nwy104, 2018.

Wang, C. and Fiedler, P. C.: ENSO variability and the eastern tropical Pacific: A review, Progress in Oceanography, 69, 239–266, https://doi.org/https://doi.org/10.1016/j.pocean.2006.03.004, a Review of Eastern Tropical Pacific Oceanography, 2006.

Wang, C., Weisberg, R. H., and Virmani, J. I.: Western Pacific interannual variability associated with the El Niño-Southern Oscillation, Journal of Geophysical Research: Oceans, 104, 5131–5149, https://doi.org/https://doi.org/10.1029/1998JC900090, 1999.

1500 Wang, C., Deser, C., Yu, J.-Y., DiNezio, P., and Clement, A.: El Niño Southern Oscillation (ENSO): A Review, in: Coral Reefs of the Eastern Tropical Pacific, edited by Glymm, P., Manzello, D., and Enochs, I., chap. 4, pp. 85–106, Springer Science+Business Media, Heidelberg, Germany, 2017.

Wang, R. and Ren, H.-L.: Understanding Key Roles of Two ENSO Modes in Spatiotemporal Diversity of ENSO, Journal of Climate, 33, 6453 – 6469, https://doi.org/10.1175/JCLI-D-19-0770.1, 2020.

1505 Weisberg, R. H. and Wang, C.: A Western Pacific Oscillator Paradigm for the El Niño-Southern Oscillation, Geophysical Research Letters, 24, 779–782, https://doi.org/https://doi.org/10.1029/97GL00689, 1997.

Wengel, C., Latif, M., Park, W., Harlaß, J., and Bayr, T.: Seasonal ENSO phase locking in the Kiel Climate Model: The importance of the equatorial cold sea surface temperature bias, Climate Dynamics, 50, 901–919, https://doi.org/10.1007/s00382-017-3648-3, 2018.

Wittenberg, A. T.: Extended Wind Stress Analyses for ENSO, Journal of Climate, 17, 2526 – 2540, https://doi.org/10.1175/1520-1510 0442(2004)017<2526:EWSAFE>2.0.CO;2, 2004.

Wu, Z. and Huang, N. E.: A study of the characteristics of white noise using the empirical mode decomposition method, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 460, 1597–1611, https://doi.org/10.1098/rspa.2003.1221, 2004.

Wu, Z. and Huang, N. E.: ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOISE-ASSISTED DATA ANALYSIS METHOD, 1515 Advances in Adaptive Data Analysis, 01, 1–41, https://doi.org/10.1142/S1793536909000047, 2009.

Wu, Z., Huang, N. E., Long, S. R., and Peng, C.-K.: On the trend, detrending, and variability of nonlinear and nonstationary time series, Proceedings of the National Academy of Sciences, 104, 14 889–14 894, https://doi.org/10.1073/pnas.0701020104, 2007.

Zebiak, S. E. and Cane, M. A.: A Model El Niño–Southern Oscillation, Monthly Weather Review, 115, 2262 – 2278, https://doi.org/10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2, 1987.

1520   Zhang, Z., Ren, B., and Zheng, J.: A unified complex index to characterize two types of ENSO simultaneously, Scientific Reports, 9, 8373, https://doi.org/10.1038/s41598-019-44617-1, 2019.

Supplementary Information for 'Identifying Quasi-periodic Variability Using Multivariate Empirical Mode Decomposition: a Case of the Tropical Pacific'

Lina Boljka[1], Nour-Eddine Omrani[1], and Noel S. Keenlyside[1]

[1]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway

Correspondence: Lina Boljka (lina.boljka@uib.no), Nour-Eddine Omrani (noureddine.omrani@uib.no)

S.1 Supplementary Figures
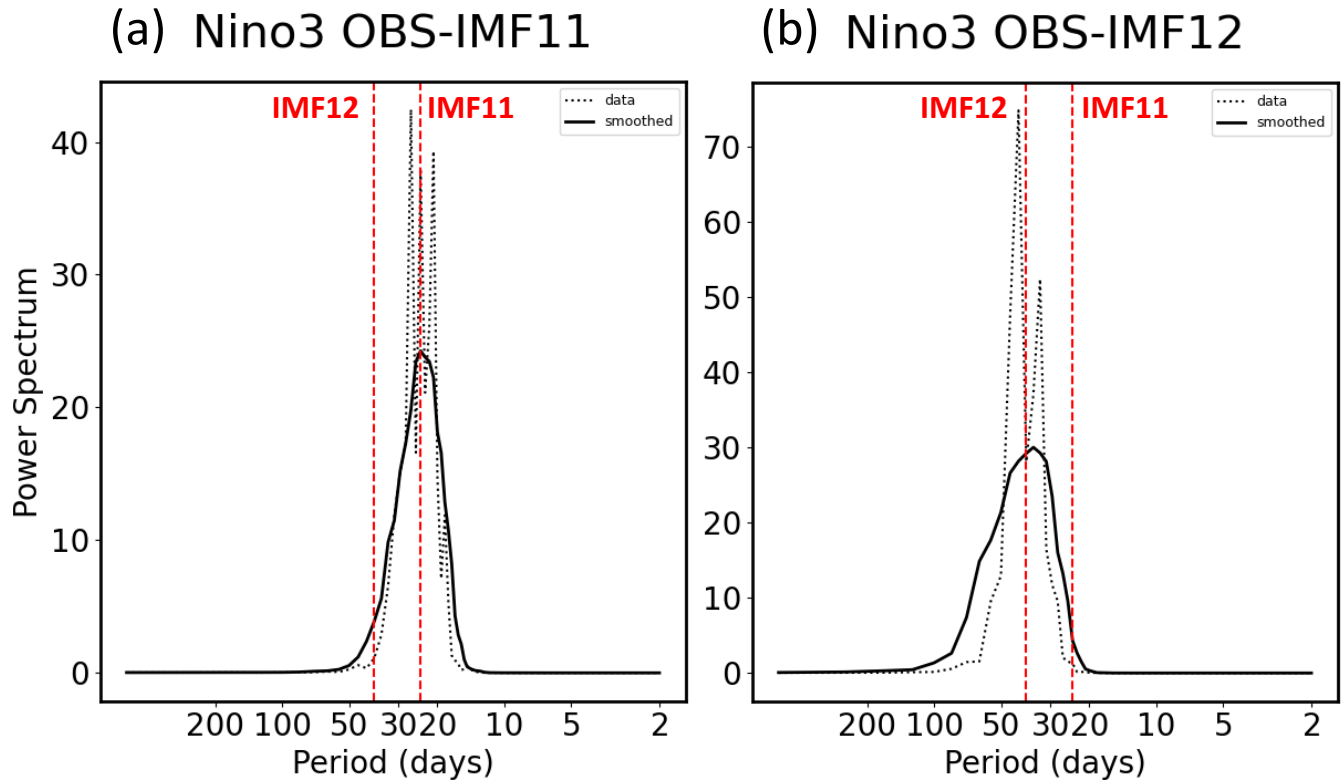
Figure S1. Power spectra of (a) ~~IMF12~~ IMF11 and (b) ~~IMF13 from reanalysis/observational data~~IMF12, their eastern Pacific SST (Niño3) index. Black dotted lines represent raw power spectra of IMFs, black solid line is 10-point smoothing of the raw power spectra, and red dashed lines represent ~~averaged~~ average frequencies of ~~IMF12~~ IMF11 and ~~IMF13~~ IMF12 (as labelled) — for values see the main text or Table S1 (second column).
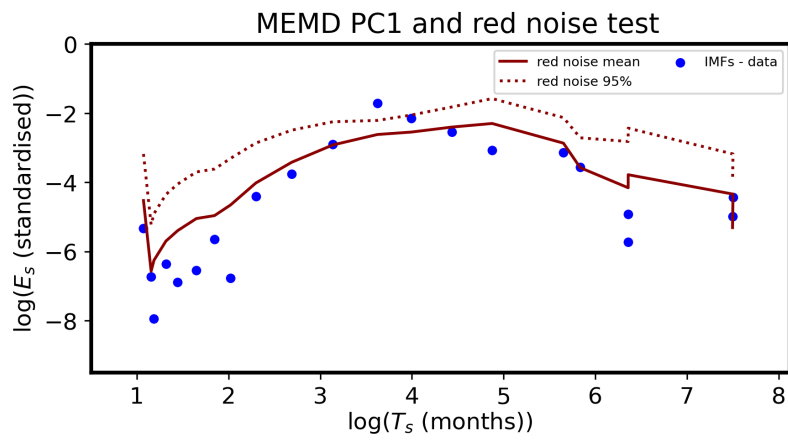
Figure S2. As Fig. 2 in the main text but for ~~IMF12~~IMFs of PC1 of the combined field (via MEMD; blue dots) instead of eastern Pacific SST (Niño3) index.
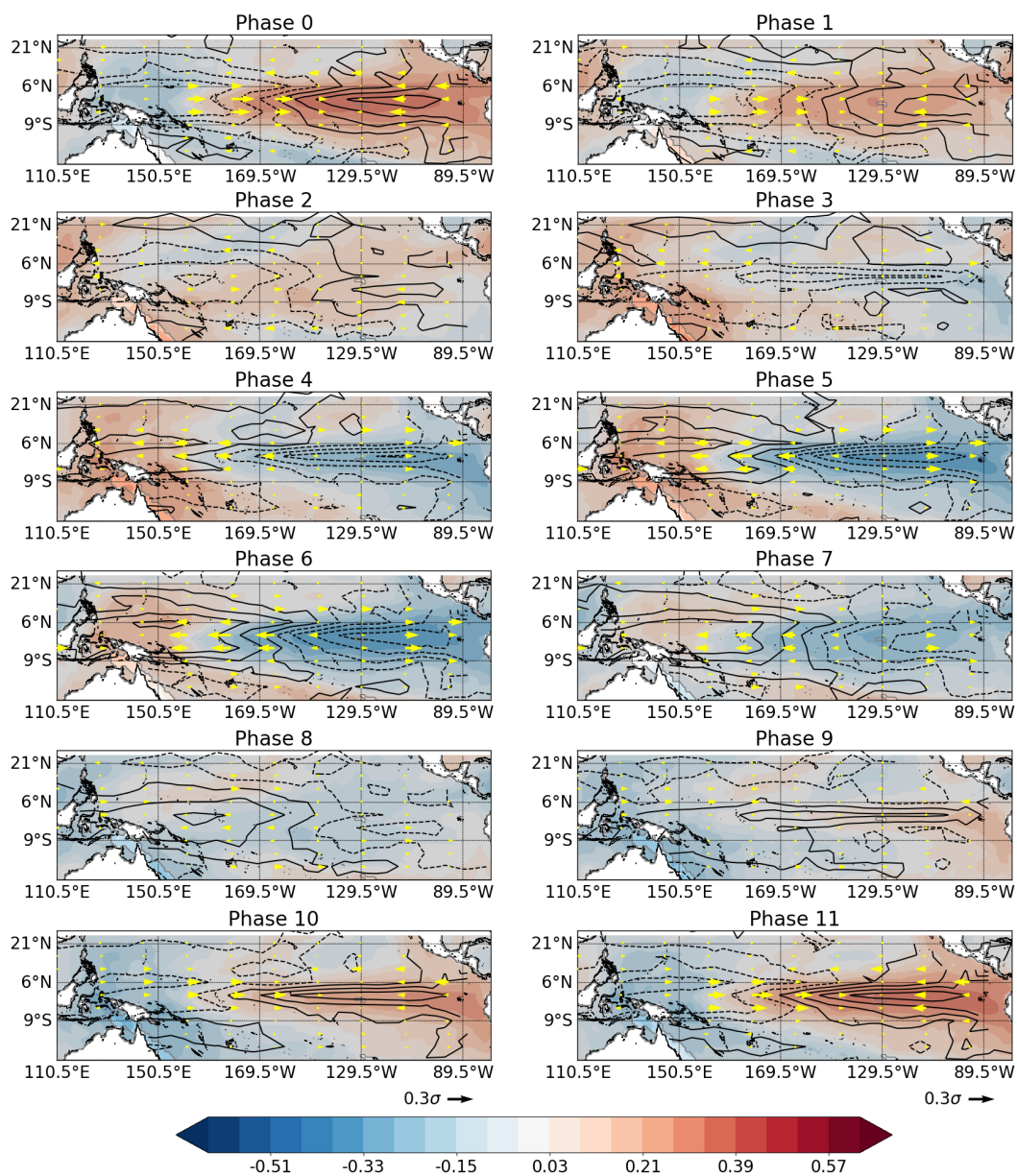
Figure S3. As Fig. 2 5 in the main text but for IMF14 from NorCPM1 model IMF11.
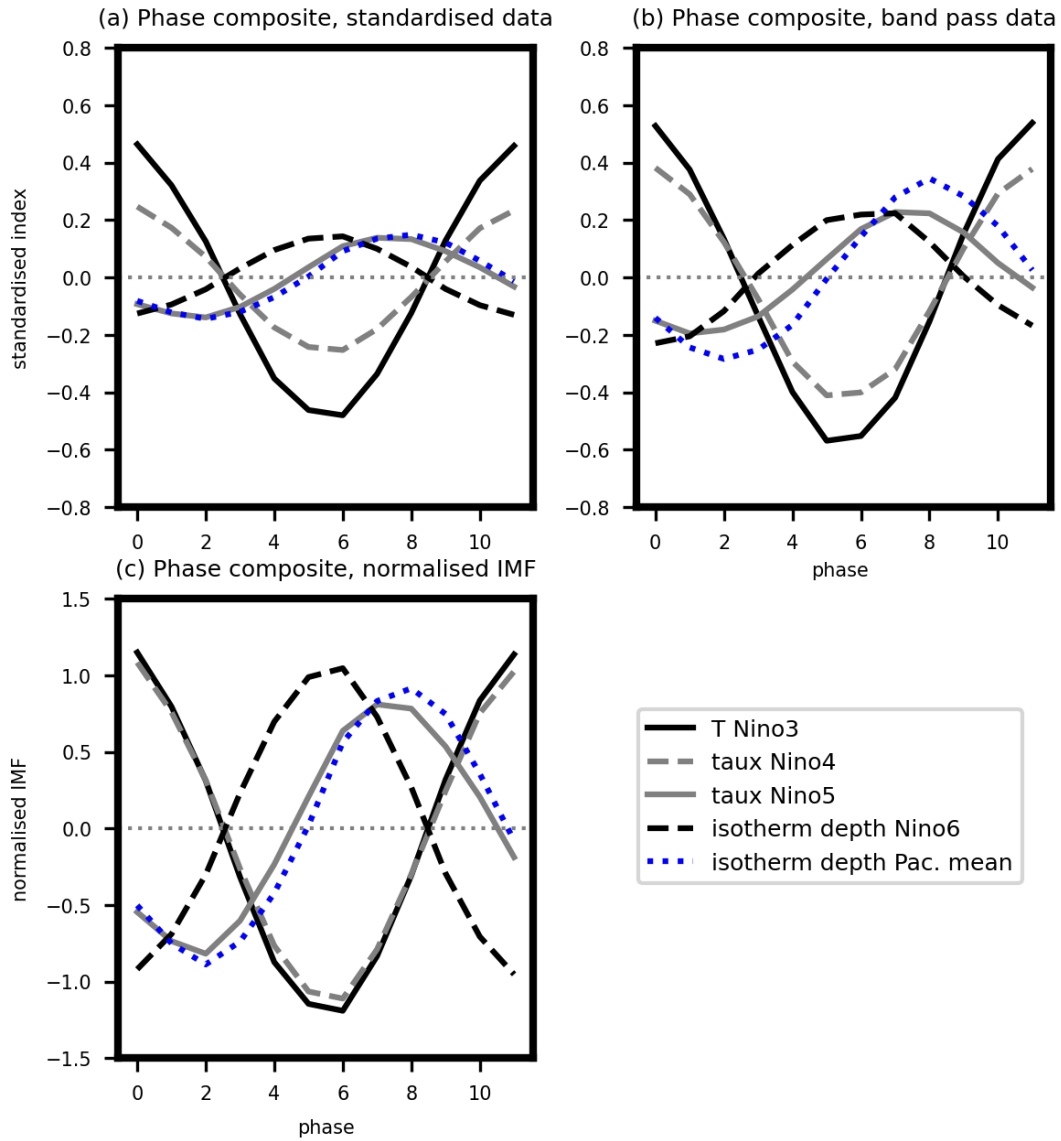
Figure S4. As Fig. ~~2~~ 6 in the main text but for ~~IMF15 from NorCPM1 model~~IMF11.
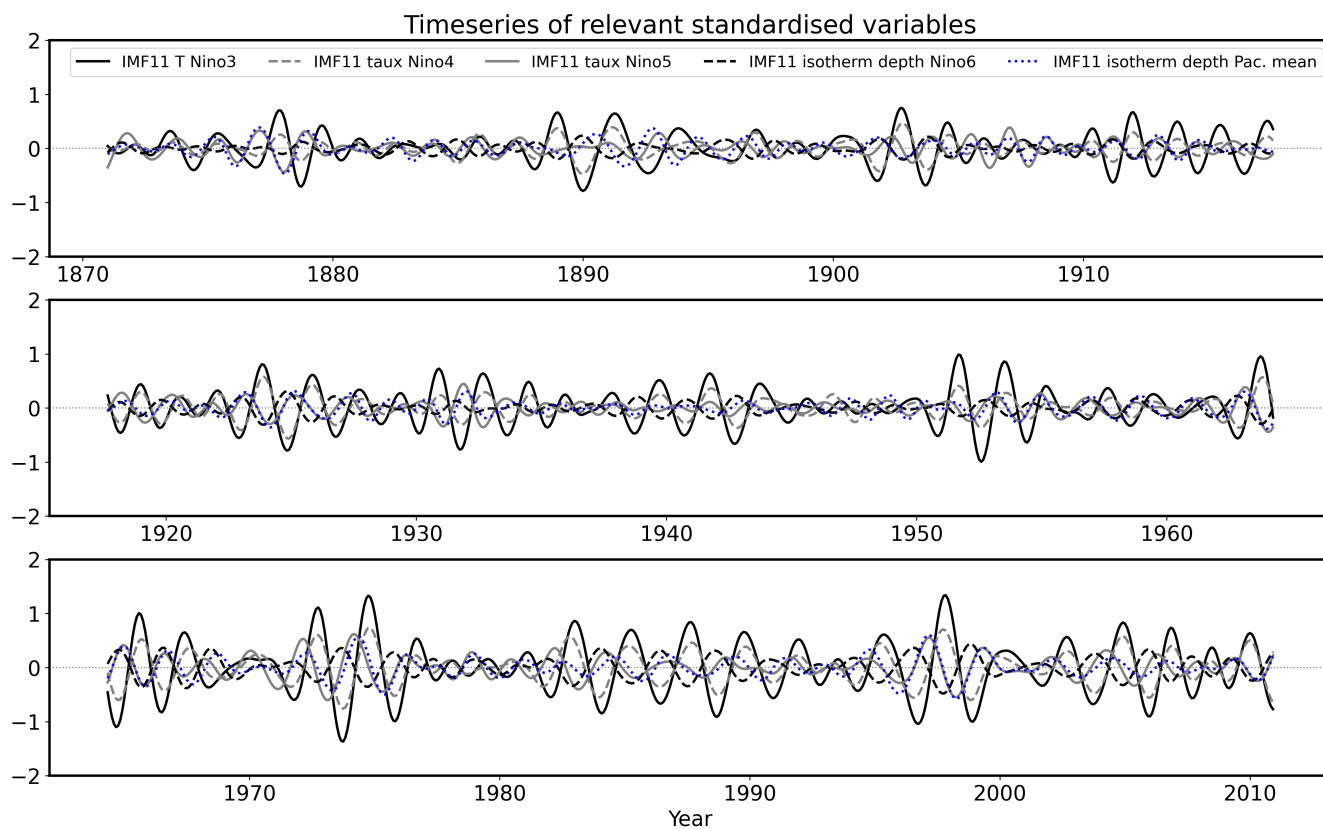
Figure S5. As Fig. ~~1~~7 in the main text but for ~~IMF14 and IMF15 from NorCMP1 model~~IMF11. ~~Here band-pass filter is computed for periods 22-54 months.~~
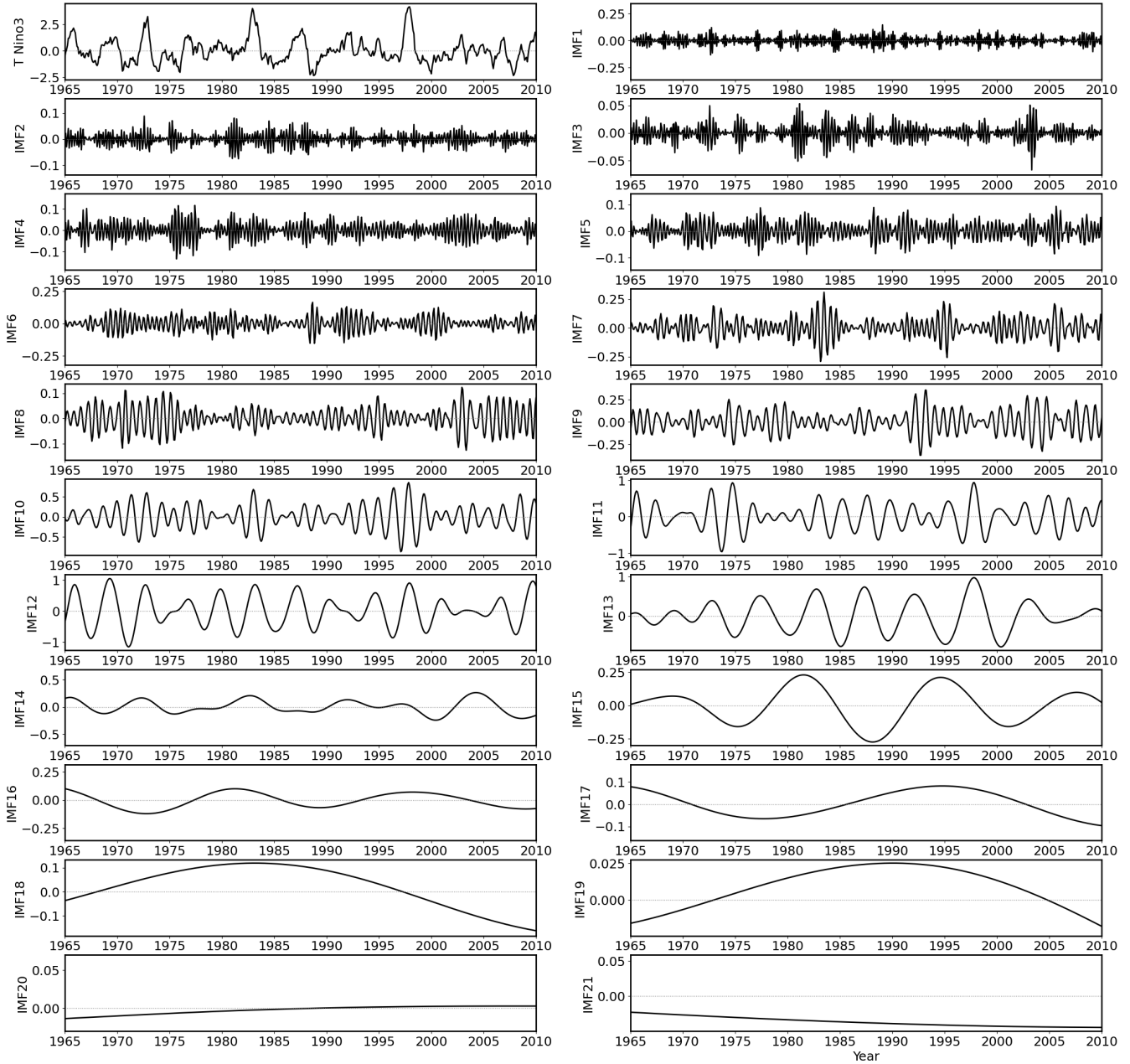
Figure S6. Timeseries of eastern Pacific SST (~~Nino3~~Niño3) from input data (top left panel) and IMFs as inferred via MEMD analysis for the same variable (see other panels as labelled). For clarity only values between 1965 and 2010 are shown. Note that amplitudes of different modes vary, i.e., y-axis is not the same in all panels. For characteristic periods of IMFs see Table S1 (second column).
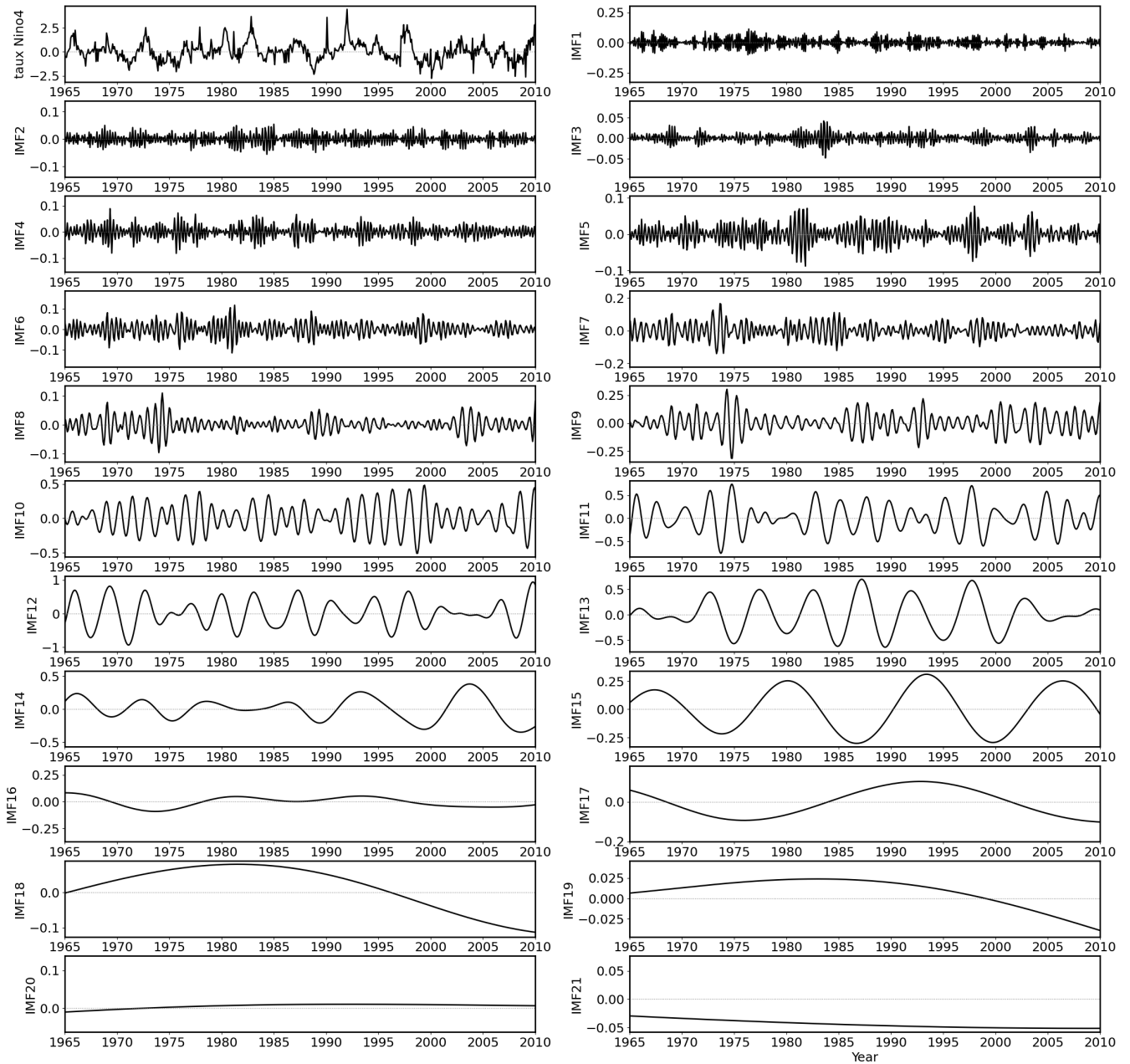
Figure S7. As Fig. S6 but for central Pacific τ_x (~~Nino4~~Niño4). For characteristic periods of IMFs see Table S1 (third column).
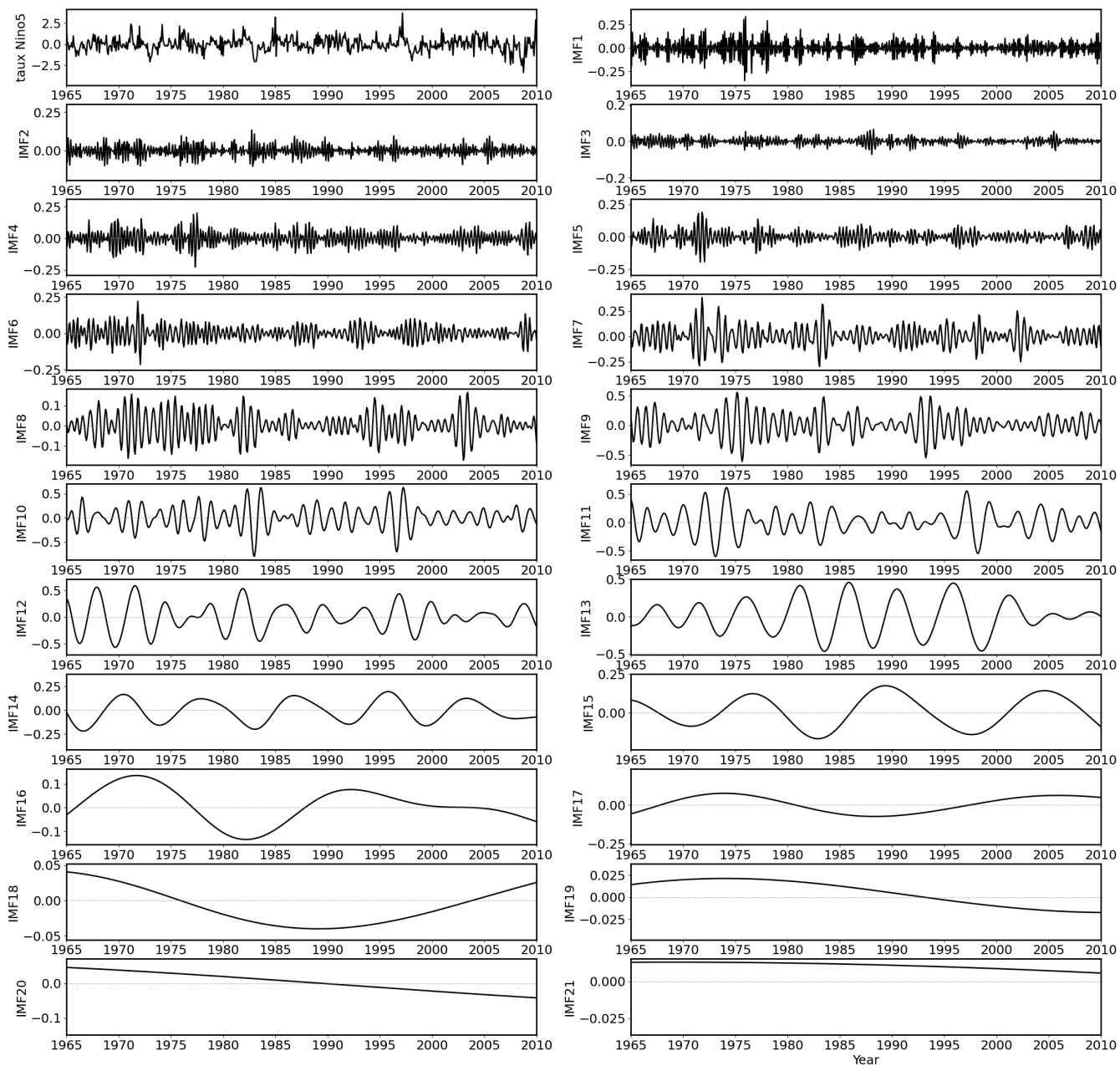
Figure S8. As Fig. S6 but for western Pacific τ_x (~~Nino5~~Niño5). For characteristic periods of IMFs see Table S1 (fourth column).
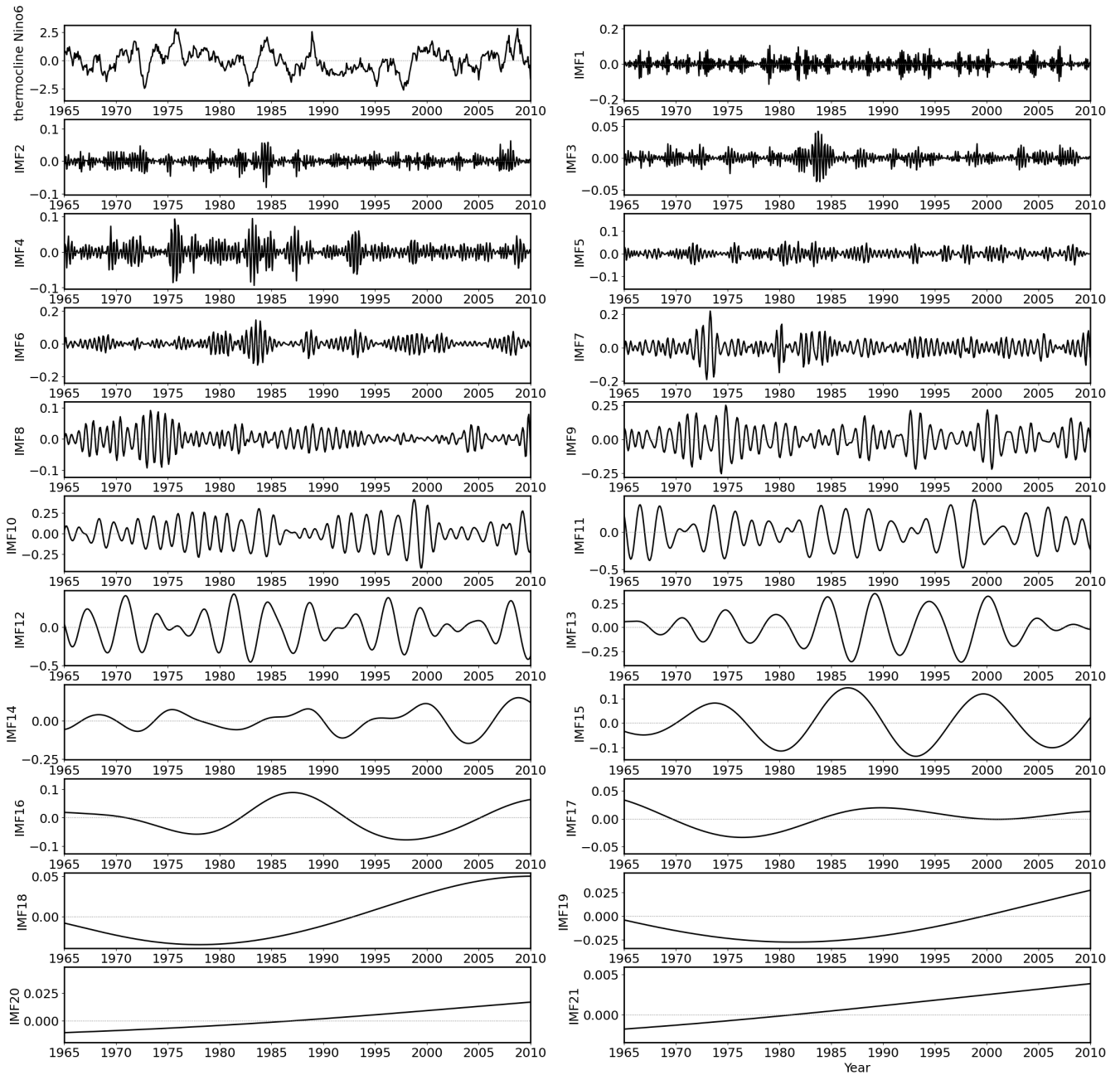
Figure S9. As Fig. S6 but for western Pacific off-equatorial thermocline depth (~~Nino6~~Niño6). For characteristic periods of IMFs see Table S1 (fifth column).
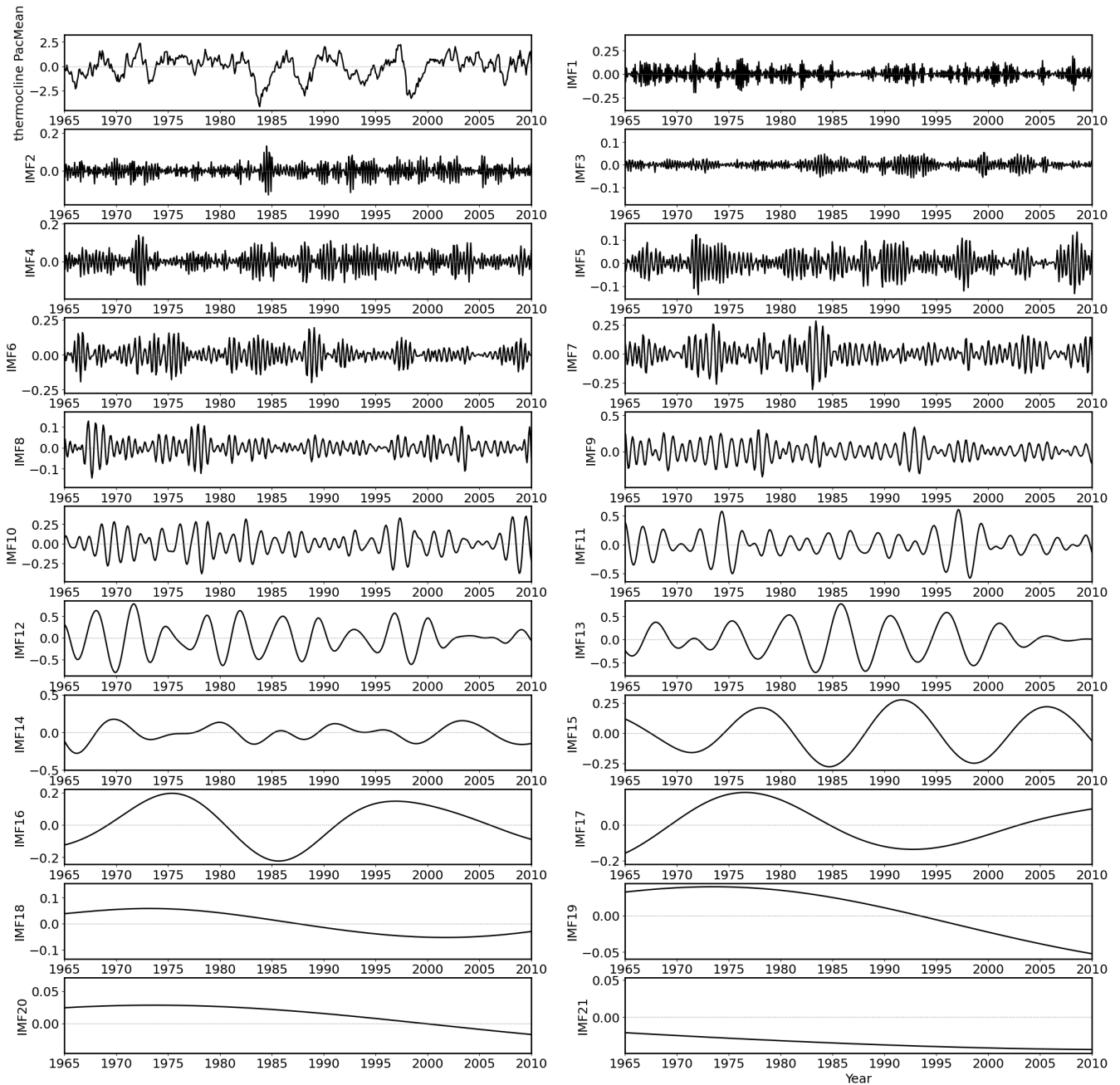
Figure S10. As Fig. S6 but for ~~thermocline depth (~~Pacific mean ~~)~~thermocline depth. For characteristic periods of IMFs see Table S1 (right column).

	~~SST (Nino3)~~ Niño3 SST	Niño4 τ_x ~~(Nino4)~~	Niño5 τ_x ~~(Nino5)~~	~~thermocline depth (Nino6)~~ Niño6 thermocline depth	~~thermocline depth (~~
IMF1	2.9	2.9	2.9	2.9	
IMF2	3.1	~~3.1~~ 3.2	3.0	~~3.1~~ 3.2	
IMF3	3.2	3.2	~~3.2~~ 3.3	~~3.2~~ 3.3	
IMF4	~~3.6~~ 3.8	~~3.5~~ 3.7	~~3.6~~ 3.8	~~3.6~~ 3.7	
IMF5	~~4.1~~ 4.3	~~4.1~~ 4.2	~~4.1~~ 4.3	~~4.0~~ 4.3	
IMF6	~~4.4~~ 5.2	~~4.5~~ 5.2	~~4.5~~ 5.2	~~4.5~~ 5.2	
IMF7	~~5.4~~ 6.3	~~5.4~~ 6.3	~~5.4~~ 6.4	~~5.4~~ 6.4	
IMF8	~~6.4~~ 7.5	~~6.5~~ 7.5	~~6.7~~ 7.5	~~6.6~~ 7.5	
IMF9	~~7.9~~ 9.8	~~8.0~~ 10	~~8.1~~ 9.8	~~8.2~~ 10	
IMF10	~~10~~ 15	~~10~~ 15	~~9.9~~ 14	~~10~~ 14	
IMF11	~~14~~ 23	~~14~~ 23	~~14~~ 24	~~14~~ 24	
IMF12	~~22~~ 39	~~22~~ 39	~~21~~ 37	~~22~~ 37	
IMF13	~~34~~ 58	~~35~~ 56	~~35~~ 54	~~35~~ 54	
IMF14	~~65~~ 89	~~62~~ 85	~~60~~ 90	~~58~~ 90	
IMF15	~~89~~ 141	~~95~~ 152	~~91~~ 152	~~95~~ 152	
IMF16	~~141~~ 206	~~141~~ 274	~~156~~ 239	~~153~~ 239	
IMF17	~~277~~ 370	~~274~~ 358	~~333~~ 336	~~286~~ 336	
IMF18	~~423~~ 590	~~448~~ 624	~~398~~ 582	~~409~~ 582	
IMF19	~~621~~ 713	~~648~~ 1120	~~744~~ 879	~~858~~ 879	
IMF20	~~964~~ 1965	~~1049~~ 1988	~~1008~~ 1844	~~886~~ 1845	
IMF21	~~1131~~ 2013	~~1587~~ 2013	~~1740~~ 1764	~~1754~~ 1764	
~~IMF22 2011 2006 1958 1973 2012~~					

Table S1. Characteristic timescales of all IMFs for eastern Pacific SST (~~Nino3~~Niño3), central Pacific τ_x (~~Nino4~~Niño4), western Pacific τ_x (~~Nino5~~Niño5), western Pacific off-equatorial thermocline depth (~~Nino6~~Niño6), ~~thermocline depth (~~Pacific mean ~~)~~thermocline depth. All values are given as approximate average periods in months. For corresponding timeseries of each variable's IMFs see Figs. S6-S10. Note that ~~IMF22~~ IMF21 is a trend by definition and similarly ~~IMF19-IMF21~~ IMF18-IMF20 have long timescale, thus periods of these IMFs are harder to establish using Hilbert transform (text below Eq. B4).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
IMF1	2.9	2.8	2.9	3.0	3.0	~~2.9~~3.0	2.9	~~3.1~~3.0	~~2.8~~3.0
IMF2	~~3.0~~3.2	~~3.1~~3.2	3.1	~~3.0~~3.1	3.1	3.2	3.1	3.2	3.1
IMF3	3.3	3.2	~~3.1~~3.3	~~3.2~~3.3	3.2	~~3.2~~3.3	3.3	3.3	3.2
IMF4	~~3.5~~3.7	~~3.6~~3.7	~~3.5~~3.8	~~3.6~~3.8	~~3.6~~3.8	~~3.5~~3.7	~~3.6~~3.7	~~3.5~~3.8	~~3.6~~3.7
IMF5	~~4.0~~4.2	~~4.1~~4.3	~~4.1~~4.3	~~4.1~~4.2	~~4.1~~4.3	~~4.0~~4.3	~~4.1~~4.4	~~4.2~~4.4	~~4.1~~4.4
IMF6	~~4.5~~5.2	~~4.5~~5.2	~~4.4~~5.2	~~4.5~~5.1	~~4.5~~5.2	~~4.5~~5.2	~~4.5~~5.1	~~4.5~~5.1	~~4.5~~5.2
IMF7	~~5.5~~6.3	~~5.4~~6.3	~~5.3~~6.4	~~5.4~~6.4	~~5.4~~6.3	~~5.4~~6.4	~~5.4~~6.3	~~5.4~~6.3	~~5.3~~6.3
IMF8	~~6.6~~7.6	~~6.5~~7.5	~~6.7~~7.5	~~6.6~~7.5	~~6.4~~7.5	~~6.5~~7.6	~~6.6~~7.5	~~6.6~~7.6	~~6.7~~7.7
IMF9	~~8.2~~10	~~8.3~~9.9	~~8.1~~9.5	8,0.8,2.7,9.8,0.8,0.7,9.8,0.8,2.8,1.7,7.8,1.8,1.8,0.7,9.8,1.8,0.8,1.~~IMF10~~10	~~10~~9.7	~~10~~9.9	~~10~~9.6	~~10~~9.9	~~10~~9.8
~~IMF11~~IMF10	15	15	~~14~~15	~~14~~15	14	14	~~14~~15	14	~~14~~15
~~IMF12~~IMF11	23	~~21~~22	~~21~~23	~~21~~24	22	22	22	22	21
~~IMF13~~IMF12	~~37~~38	36	~~36~~38	33~34	~~36~~37	~~36~~37	~~36~~38	~~33~~35	~~34~~37
~~IMF14~~IMF13	~~65~~54	~~59~~58	56	~~58~~54	60	~~58~~54	~~58~~60	~~58~~56	57
~~IMF15~~IMF14	~~95~~84	99	~~103~~97	96	~~90~~86	~~90~~87	~~104~~97	~~103~~91	97
~~IMF16~~IMF15	~~150~~138	146	~~161~~147	134	156	~~166~~150	~~173~~141	~~147~~133	151
~~IMF17~~IMF16	~~275~~267	~~282~~276	~~297~~247	~~275~~221	~~276~~210	~~314~~245	~~310~~245	~~286~~243	~~274~~229
~~IMF18~~IMF17	~~355~~361	~~403~~395	~~436~~454	~~434~~429	413	~~432~~337	~~419~~335	~~419~~356	~~548~~401
~~IMF19~~IMF18	~~677~~636	~~833~~725	~~628~~617	~~648~~576	~~621~~591	~~838~~555	~~694~~616	~~841~~525	~~576~~573
~~IMF20~~IMF19	~~1042~~610	~~972~~897	~~1037~~1056	~~905~~754	~~888~~907	~~1162~~890	~~994~~826	~~921~~890	~~946~~903
~~IMF21~~IMF20	~~1714~~1940	~~1738~~1928	~~1649~~2124	~~1734~~1869	~~1754~~1892	~~1748~~1864	~~1738~~1818	~~1723~~1791	~~1452~~1841
~~IMF22~~IMF21	~~2002~~2012	~~1999~~2012	~~2006~~2013	~~1958~~1851	~~1907~~1757	~~1975~~1982	~~1929~~1943	~~1734~~2005	~~2011~~1853

Table S2. Characteristic timescales of all IMFs for all 20 PCs. All values are given as approximate average periods in months. Note that ~~IMF22~~IMF21 is a trend by definition and similarly ~~IMF19~~IMF21, ~~IMF18~~IMF20 have long timescale, thus periods of these IMFs are harder to establish using Hilbert transform (text below Eq. B4).

S.3 Simple Example for MEMD

~~We can test MEMD method on simple data with periodic signals, like cosine and sine waves.~~ As an example, we describe how
MEMD works on a few simple periodic timeseries. We define four timeseries that have a shared angular frequency of $\pi/2$ with
other harmonics or phase shifts added on top. We input these timeseries into the MEMD and expect the MEMD to isolate the
shared mode with angular frequency of $\pi/2$ in all four timeseries, i.e., find the synchronised signal within the timeseries. We
also expect MEMD to find other harmonics in the timeseries.

To do this, we construct ~~4~~ the four timeseries as follows

$$x_{\text{inp}} = \sin\left(\frac{\pi t}{2}\right) \tag{S.1}$$

$$y_{\text{inp}} = \sin\left(\frac{\pi t}{2}\right) + \sin\left(\frac{2\pi t}{2}\right) + \sin\left(\frac{4\pi t}{2}\right) \tag{S.2}$$

$$z_{\text{inp}} = \sin\left(\frac{\pi t}{2}\right) + \sin\left(\frac{3\pi t}{2}\right) \tag{S.3}$$

$$w_{\text{inp}} = \sin\left(\frac{\pi t}{2} + \frac{\pi}{2}\right) \tag{S.4}$$

where t is time and $x_{\text{inp}}, y_{\text{inp}}, z_{\text{inp}}, w_{\text{inp}}$ are timeseries with a common periodic signal $\sin(\pi t/2)$ and a few additional timescales
or phase shifts. Thus, w_{inp} is the same as x_{inp} but 90-degrees phase shifted, whereas y_{inp} and z_{inp} have additional timesecales
that are double, tripple or quadruple of x_{inp}'s timescale. These four timeseries ~~are plotted in~~ (Fig. S11 ~~(~~, top left) ~~and are then~~
are input into MEMD algorithm. The algorithm then returns 5 IMFs. IMF3 (Fig. S11, middle right) can be considered as the
goal of this data, i.e., identification of common timescales across the 4 different timeseries/datasets, i.e., angular frequency
$\pi/2$ (as mentioned above). The algorithm identifies the same mode in all four timeseries despite phase shifting or presence of
other timescales in these simple timeseries. Such a mode is robustly identified across different parameter sweeps of MEMD
(not shown). Thus, IMF3 can be considered here as equivalent of the ENSO's LF/QQ mode that has been shown in the past to
exist across the tropical Pacific and a similar mode is identified again in the main text via MEMD as well.

IMF1 (Fig. S11, top right) represents the fastest 'waves' (shortest period/timescale) that we can find in y_{inp} and z_{inp}, i.e.,
related to angular frequencies $3\pi/2$ and $4\pi/2$. The latter two frequencies are identified by the MEMD as similar thus they
appear in the same IMF, although one could change the parameters of the MEMD algorithm to split the two modes into
separate IMFs. However, that can then lead to splitting up other modes as well (especially IMF2), leading to unrealistic results
(i.e., mode mixing; not shown). IMF2 (Fig. S11, middle left) shows intermediate angular frequency present in y_{inp}, i.e., $2\pi/2$,
but this IMF's output is not perfect, resulting in varying amplitudes of the wave throughout the analysis period, and thus IMF4
and IMF5 (bottom panels in Fig. S11) then compensate for the loss of amplitude in IMF2 in this case. Note that a longer
timeseries somewhat helps mitigating this issue as any timeseries analysis tool has issues at the edges of the data and thus
only data sufficiently far from the edges should be considered in analysis (there amplitude can be somewhat stable in IMF2).
This means that longer datasets are preferred for MEMD analysis to ensure stability. Also, IMF4 and IMF5 should technically
be zero (given the chosen input timeseries), but due to edge effects and other issues with (M)EMD method (see main text for

details) they are still present though their amplitudes are small. This suggests that IMFs of the longest periods can sometimes

35 be rather artificial constructs of the data and should be treated with caution especially when the trend of the data is essentially

zero (as here or in the main text where trend has been removed prior to MEMD analysis).

This example only shows that signals that are well synchronized across timeseries will show up clearly in MEMD analysis, however other signals that exist in, e.g., only one mode (e.g., y_{inp}'s $2\pi/2$ wave) can be problematic as the method may struggle with keeping zeros in other timeseries (see IMF2). Then, leaking can occur both within, e.g., IMF2 and into other

40 modes, causing mode-mixing again (like here IMF2 leaks into IMF4,5, especially at the edges). Similar issues can exist with trends as shown here. Thus, caution and verification with other methods is advised when using MEMD.
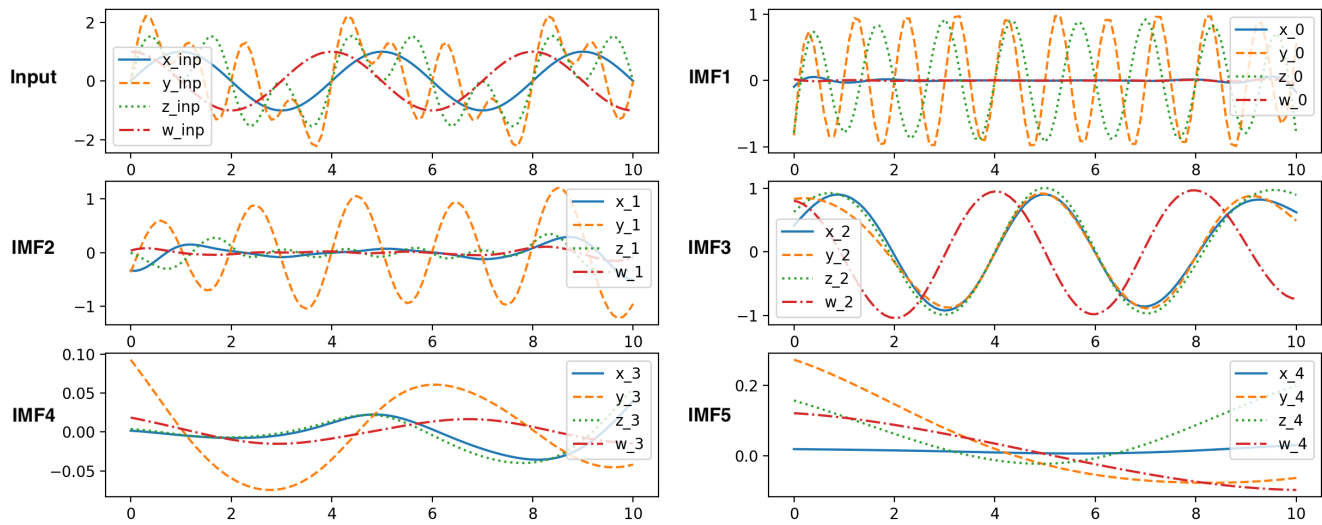
Figure S11. MEMD analysis of simple timeseries (Eqs. S.1-S.4). Top left panel shows input timeseries and the rest of the panels show the five IMFs that MEMD produces. IMF5 typically represents the trend of the data. See text for more details. Note that amplitudes of IMF4,5 are smaller than for IMF1,2,3 (i.e., y-axes are not the same across panels).