Lagged Teleconnections of Climate Variables Identified via Complex Rotated Maximum Covariance Analysis

NICLAS RIEGER,[a,b] ÁLVARO CORRAL,[a,c,d] ESTRELLA OLMEDO,[e] AND ANTONIO TURIEL[e]

[a] *Centre de Recerca Matemàtica, Bellaterra, Barcelona, Spain*
[b] *Departament de Física, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*
[c] *Departament de Matemàtiques, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*
[d] *Complexity Science Hub Vienna, Vienna, Austria*
[e] *Institute of Marine Sciences (ICM-CSIC) and Barcelona Expert Center, Barcelona, Barcelona, Spain*

(Manuscript received 30 March 2021, in final form 9 August 2021)

ABSTRACT: A proper description of ocean–atmosphere interactions is key for a correct understanding of climate evolution. The interplay among the different variables acting over the climate is complex, often leading to correlations across long spatial distances (teleconnections). On some occasions, those teleconnections occur with quite significant temporal shifts that are fundamental for the understanding of the underlying phenomena but that are poorly captured by standard methods. Applying orthogonal decomposition such as maximum covariance analysis (MCA) to geophysical datasets allows us to extract common dominant patterns between two different variables, but generally suffers from (i) the nonphysical orthogonal constraint as well as (ii) the consideration of simple correlations, whereby temporally offset signals are not detected. Here we propose an extension, complex rotated MCA, to address both limitations. We transform our signals using the Hilbert transform and perform the orthogonal decomposition in complex space, allowing us to correctly correlate out-of-phase signals. Subsequent varimax rotation removes the orthogonal constraints, leading to more physically meaningful modes of geophysical variability. As an example, we have employed this method on sea surface temperature and continental precipitation; our method successfully captures the temporal and spatial interactions between these two variables for (i) the seasonal cycle, (ii) canonical ENSO, (iii) the global warming trend, (iv) the Pacific decadal oscillation, (v) ENSO Modoki, and finally (vi) the Atlantic meridional mode. The complex rotated modes of MCA provide information on the regional amplitude and, under certain conditions, the regional time lag between changes on ocean temperature and land precipitation.

SIGNIFICANCE STATEMENT: Correlations between time series of different climate variables are often time-lagged and can appear over long spatial distances. Our goal was to develop a method that allows the simultaneous identification of the dominant spatial patterns and their time lags between two different climate variables. Using sea surface temperatures and continental rainfalls as an example, our method extracts the different dynamics of the seasonal cycle as well as five other well-known climate phenomena. Especially for cyclic time series like the seasonal cycle, the relative time lags at different locations can be determined precisely, whereas for acyclic time series only qualitative statements about time lags can be made. In future studies, we expect new insights into the dynamical structure of the Madden–Julian oscillation thanks to this method, which is readily available as a Python package.

KEYWORDS: Atmosphere-ocean interaction; Teleconnections; Precipitation; Sea surface temperature; Empirical orthogonal functions; Pattern detection; Principal components analysis; Dimensionality reduction

1. Introduction

Earth's climate system is extremely complicated and deciphering the web of interdependencies and influences of different climate subsystems is an involved challenge. As the quantity and quality of Earth observations increase thanks to the advances in remote sensing, so too does the amount of data that needs to be processed. Data-driven dimensionality reduction methods are therefore crucial for climate studies, as they allow high-dimensional spatiotemporally resolved signals to be disaggregated into the dominant patterns, while still capturing the subtle details of higher-resolution data. As such, principal component analysis (PCA), or empirical orthogonal function (EOF) analysis as it is often referred to in climate science, allows us to identify the dominant internal structure of the variability as expressed by the variance, with a variety of different available versions of PCA proving the popularity of such methods in climate science (e.g., Hannachi et al. 2007; Hannachi 2021).

Climate phenomena with different expression in oceanic and atmospheric variables, such as El Niño–Southern Oscillation (ENSO), however, require the simultaneous analysis of several variables for a more comprehensive description. In principle, multivariate PCA (Kutzbach 1967) makes it possible to extract the patterns of covariability of more than one variable. However, multivariate PCA accumulates the variance and the covariance of variables with very different variability in the same quantities. In consequence this may mask covarying patterns as low-variability patterns of one variable can be erroneously accumulated in very dominant structures of one of the other, large-variability variables (Bretherton et al. 1992).

Corresponding author: Niclas Rieger, nrieger@crm.cat

DOI: 10.1175/JCLI-D-21-0244.1

Maximum covariance analysis (MCA)[1] avoids this masking by taking into account only the covariance between two sets of variables. As such, it bears similarity to canonical correlation analysis (CCA; Hotelling 1936), which aims at maximizing the temporal correlation between both variables. When the number of grid points (i.e., number of time series) is higher than the number of observations (i.e., number of time steps) and the data exhibit multicollinearity, as is often the case for climate data, CCA fails as it requires the individual variance matrices to be nonsingular unless regularized (Vinod 1976; Cruz-Cano and Lee 2014; Hannachi 2016). If the two fields of variables are identical, MCA reduces to PCA, the former thus being a natural generalization of PCA.

Yet the methods discussed above maximize instantaneous correlation and do not consider time-delayed signals. To gain a deeper understanding of the dynamics of climate phenomena, however, it is necessary to systematically investigate time lags. A typical approach to tackle this problem is to consider one variable set with a time lag defined a priori followed by an MCA (e.g., Li et al. 2016). However, this requires knowledge of the time lag, which may vary from one location to another (Ballabrera-Poy et al. 2002).

In this paper, we propose *complex rotated MCA* to systematically investigate the phase shift of two variables. We generate complex time series known as the analytical signal, where the real and imaginary parts are related to each other by the Hilbert transform, and decompose the covariance matrix in complex space, in analogy to complex PCA (Horel 1984; Bloomfield and Davis 1994). We also effectively reduce spectral leakage inherent in the Hilbert transform of noncyclic signals by using an extrapolation method of the signal beyond its boundaries. Finally, to relax the orthogonality constraint of the obtained solutions, we apply varimax rotation to the spatial patterns, which leads to more localized solutions and thus facilitates their physical interpretation (Richman 1986; Cheng and Dunkerton 1995).

To make the method readily accessible as a tool, we provide it as a Python package, called *xmca* (Rieger 2021). Due to the power and popularity of NumPy (van der Walt et al. 2011) and xarray (Hoyer and Hamman 2017), both packages form the basis of xmca, so that their typical data format can be used directly as input for analysis. The package is modularized in a way that provides the user free choice whether standard, complex, rotated, or complex rotated MCA is to be performed. The user can also choose between varimax orthogonal rotation as well as promax oblique rotation. Further, if desired, standardization of the input data is computed on the fly. The different flavors work in the same way for PCA, if one instead of two fields is provided as input.

The remainder of the article is structured as follows. Section 2 introduces the methodology, where we briefly discuss MCA (section 2a), complex MCA (section 2b), and rotated MCA (section 2c). Section 3 describes the data used to test the method, using first synthetic data (section 3a) and then climatic variables (section 3b). Section 4 presents the results of both the synthetic (section 4a) and real-world analysis (section 4b). We conclude our study and provide directions for future research in section 5.

2. Methods

a. Maximum covariance analysis

Let us consider two spatiotemporal data fields $\mathbf{X}_A \in \mathbb{R}^{m\times n_A}$ and $\mathbf{X}_B \in \mathbb{R}^{m\times n_B}$ representing two different geophysical fields $s \in \{A, B\}$, both having temporal dimension m and spatial dimensions n_A and n_B, respectively. Throughout the text, the index s is used implicitly without further definition to represent one of the two fields. In the following, we will refer to the temporal dimensions as the number of observations while we denote the spatial dimensions by the number of grid points. Assuming each time series to have zero mean, MCA then aims at maximizing

$$\mathbf{v}_A^{\mathrm{T}}\mathbf{C}\mathbf{v}_B, \quad \text{s.t.} \quad \mathbf{v}_A^{\mathrm{T}}\mathbf{v}_A = \mathbf{v}_B^{\mathrm{T}}\mathbf{v}_B = 1, \tag{1}$$

where $\mathbf{C}$ denotes the temporal covariance matrix and $\mathbf{v}_A, \mathbf{v}_B$ the *spatial patterns*, of both fields, respectively. Mathematically, this can be achieved by applying the singular value decomposition (SVD) to the covariance matrix,

$$\mathbf{C} = \frac{1}{m-1}\mathbf{X}_A^{\mathrm{T}}\mathbf{X}_B = \mathbf{V}_A\mathbf{\Sigma}\mathbf{V}_B^{\mathrm{T}}, \tag{2}$$

with the columns of the obtained singular vector matrices ($\mathbf{V}_s$) representing pairs of spatial patterns describing the maximum amount of temporal covariance between both variables. The entries of $\mathbf{\Sigma} \in \mathbb{R}^{n_A\times n_B}$ along the main diagonal, the singular values σ_k, represent the covariance of each spatial pattern pair k, providing a mean of estimating the relative importance of each pair via the covariance fraction[2] γ_k:

$$\gamma_k = \sigma_k\left(\sum_{j=1}^{\min(n_A,n_B)}\sigma_j\right)^{-1}. \tag{3}$$

By projecting the data fields on their respective singular vectors, we obtain the corresponding temporal evolution for each spatial pattern given by the columns of $\mathbf{P}_s = \mathbf{X}_s\mathbf{V}_s$. Since the singular vectors are orthonormal (i.e., $\mathbf{V}_s^{\mathrm{T}}\mathbf{V}_s = \mathbf{1}_{n_s}$, with $\mathbf{1}_{n_s}$ being the identity matrix of rank n_s), the projections of the left and right field are uncorrelated [i.e., $\mathbf{P}_A^{\mathrm{T}}\mathbf{P}_B = \mathbf{\Sigma}/(m-1)$] while

[1] Sometimes referred to as singular value decomposition (SVD) analysis. This name is unfortunate and should not be confused with the actual factorization technique of a real/complex matrix.

[2] Typically the squared covariance fraction defined as $\gamma_k^* = \sigma_k^2/\left(\sum_{j=1}^{\min(n_A,n_B)}\sigma_j^2\right)$ is considered for the relative importance of each mode for MCA. However, we opt for the nonsquared covariance fraction since the total explained covariance is conserved under rotation, i.e., for r rotated modes $\sum_i^r\sigma_i = \sum_i^r\sigma_i^*$, where σ_i^* refers to the covariance associated to mode i after rotation. Furthermore, this measure is comparable to the solutions obtained by PCA, and in fact it is equivalent when $\mathbf{X}_A = \mathbf{X}_B$, for which MCA reduces to PCA and the singular values equal the eigenvalues in PCA.

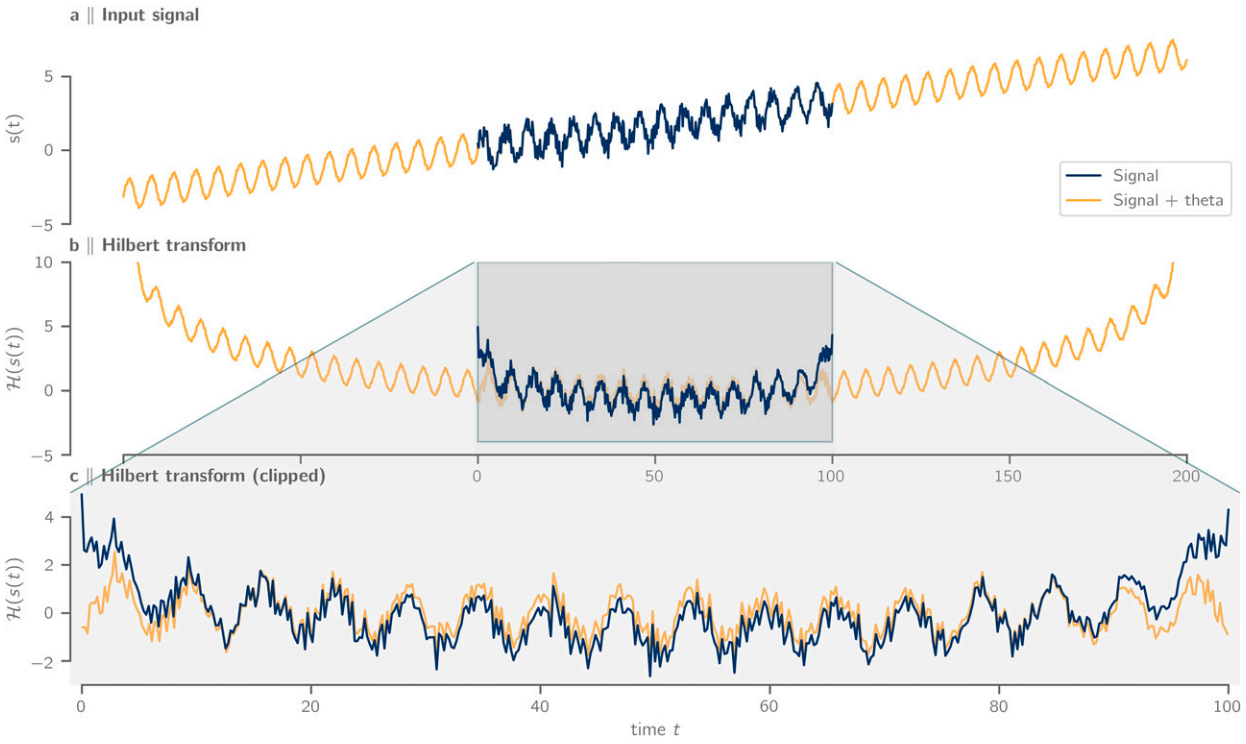

FIG. 1. Example illustrating the Hilbert transform using the theta extension. (a) Input signal $s(t) = \sin(t) + ct + \varepsilon(t)$ using an arbitrary constant c and Gaussian white noise ε with zero mean and unit variance (blue) as well as the extended time series via forecasting/hindcasting using the optimized theta model (Fiorucci et al. 2016). (b) Hilbert transform $\mathscr{H}(\cdot)$ of the original and extended signal, respectively. (c) Comparison of both Hilbert transforms over the domain of the original series.

the projections of the same field are correlated in general (i.e., $\mathbf{P}_s^{\mathrm{T}}\mathbf{P}_s$ is not a diagonal matrix). In this paper, we will refer to the spatial patterns and their corresponding time projections as *empirical orthogonal functions* (EOFs) and *principal components* (PCs), respectively, according to the usual convention in climate science. The EOFs and the PCs associated with a specific singular value σ_k are denoted as *mode k*.

b. Complex MCA

Propagating features or lagged signals could be detected by using a complex representation of the input fields. In analogy to complex PCA (Wallace and Dickinson 1972; Rasmusson et al. 1981; Horel 1984; Bloomfield and Davis 1994), we complexify the real input fields via the Hilbert transform to construct the *analytical signal* $\hat{\mathbf{X}}_s$ defined as

$$\hat{\mathbf{X}}_s = \mathbf{X}_s + i\mathscr{H}(\mathbf{X}_s), \tag{4}$$

where $\mathscr{H}(\cdot)$ denotes the column-wise applied Hilbert transform. The analytical signal constructed in that way is a unique complex representation of the real signal, but whether it also represents a physical reality depends on the frequency spectrum of the analyzed signal. By construction, the frequency components of the Hilbert transform are phase shifted by $-\pi/2$ with respect to those of the original signal. Therefore, for narrow-bandwidth signals, the Hilbert transform has a simple physical interpretation (i.e., it represents a signal that arrives with a lag of one-fourth of the typical period). If the signal consists of multiple dominating frequencies, however, the interpretation of the phase is more elusive, as it cannot be simply associated with a single frequency. Thus, the more narrow the signal bandwidth, the more directly we can relate the phase to specific timings of its Hilbert transform (Boashash 1992).

A fundamental issue in the computation of the Hilbert transform arises when nonstationary or drifting signals are processed. Such signals are noncyclic; therefore, when the Fourier coefficients are calculated, strong boundary effects can occur due to spectral leakage (Boashash 1992) (Fig. 1). This problem can be circumvented by detrending the time series and considering only integer cycles as well as by applying window functions to the time series (e.g., Hanning, Hamming). However, this comes at the cost of information loss. Additionally, such mitigation techniques are particularly ill suited to deal with nonstationarities associated with climate change (which include not only noticeable trends on the mean level, but also increases in the amplitude of some periodic phenomena). Therefore, it seems important to introduce techniques capable to dealing with nonstationary signals.

To mitigate spectral leakage across the boundaries of the time series, we extrapolate the time series at both boundaries, to the past and to the future, using the optimized theta model (Assimakopoulos and Nikolopoulos 2000; Fiorucci et al. 2016), a special case of an autoregressive integrated moving average

model with drift, ARIMA(0, 1, 1) (Hyndman and Billah 2003). The theta model is a relatively simple yet well-performing extrapolation method. While the forecast in itself actually works with noncyclic signals, seasonal features are considered via multiplicative classical decomposition, thus allowing cyclic and noncyclic signals to be extrapolated. This approach of handling the seasonal structure of a time series requires the user to specify the dominant period of the signal T_s beforehand. For discrete time series T_s represents the number of time steps needed to complete one cycle—for example, 365 for daily data considering an annual cycle, or 24 for hourly data with a daily cycle [for more details we refer the reader to Fiorucci et al. (2016)]. We then apply the Hilbert transform to the extended series, so the spectral leakage is only important on the backward and forward extensions of it. Finally, we extract the central part (removing the parts corresponding to the extension) that correctly corresponds to the Hilbert transform of the original series. Using this approach, we effectively reduce the edge effects of the Hilbert transform compared to a nonprocessed time series (Fig. 1). Notice that it is not necessary that the extrapolation faithfully reproduces the characteristics of the original series; it just suffices for our purpose that the extrapolated time series approximately continues the cyclic structure at the original time series boundaries in order to reduce spectral leakage. Apart from tracking trends, the exact extrapolation beyond the boundaries is not essential since its effects on the central part of the Hilbert transform are very marginal at most.

After the described complexification of the original time series, we follow the steps of standard MCA, with the difference that the transpose $(\cdot)^{\mathrm{T}}$ incorporates the complex conjugate $\overline{(\cdot)}$ and the obtained EOFs and projections PCs are complex and unitary. This allows us to calculate the spatial amplitude $\mathscr{A}_s$ and phase function $\mathbf{\Theta}_s$ for both fields:

$$\mathscr{A}_s = (\mathbf{V}_s \odot \overline{\mathbf{V}}_s)^{\odot 1/2}, \tag{5}$$

$$\mathbf{\Theta}_s = \arctan2[\mathrm{Im}(\mathbf{V}_s), \mathrm{Re}(\mathbf{V}_s)], \tag{6}$$

where $\odot$ denotes the element-wise multiplication/exponentiation, and arctan2 refers to the two-argument arctangent that is calculated element-wise (see appendix A). Although this matrix notation seems somewhat cumbersome compared to the more direct expression through scalar fields, it allows us to be coherent with the rest of the paper. The phase function can be interpreted directly as a time lag if the corresponding (real) PC has a narrowband spectrum with just one dominant frequency. If the spectrum is rather broadband or has several dominant frequencies, an interpretation of the phase function is usually not straightforward. We note that the complex EOFs derived from the SVD are only defined up to a phase shift of $\exp(i\theta)$ with $\theta \in [0, 2\pi]$. However, if complex EOFs are obtained with a different phase shift θ (e.g., due to another SVD algorithm), the change will also be reflected in the projected PCs, so that taking into account both, PC and phase function, the results are unambiguous.

c. Rotated MCA

While orthogonality is often a mathematically desirable property, it does not make a lot of sense from a purely geophysical standpoint. Therefore, standard EOFs are difficult to interpret in the case of geophysical data. The major drawbacks of EOFs due to orthogonality are twofold: First, EOFs are sensitive to the selected spatial domain; that is, including or removing some regions may change large parts of the EOFs. Second, EOFs tend to split certain geophysically meaningful patterns across several consecutive modes (Richman 1986). To relax the orthogonality constraint to better accommodate the geophysical reality, the EOFs can be *rotated*, which implies a linear transformation of the first r loaded[3] EOFs $\mathbf{L}_{s,r}$. This concept, which was originally developed in the context of PCA, can also be applied to MCA (Cheng and Dunkerton 1995). For this, we apply the rotation matrix $\mathbf{R} \in \mathbb{C}^{r\times r}$ to the *loading matrix* $\mathbf{L}_r \in \mathbb{C}^{n\times r}$:

$$\begin{aligned} n &= n_A + n_B, \\ \mathbf{L}_r &= \begin{pmatrix} \mathbf{L}_{A,r} \\ \mathbf{L}_{B,r} \end{pmatrix} = \begin{pmatrix} \mathbf{V}_{A,r} \\ \mathbf{V}_{B,r} \end{pmatrix} \mathbf{\Sigma}_r^{1/2}, \end{aligned} \tag{7}$$

where r reflects the respective submatrices containing only the first r columns. In addition, $\mathbf{\Sigma}_r$ is the diagonal submatrix containing only the first r columns and rows.

There are a number of different criteria for defining the rotation matrix $\mathbf{R}$ (e.g., Richman 1986), including the varimax orthogonal rotation (Kaiser 1958) and the promax oblique rotation (Hendrickson and White 1964), whose general aim is to regroup the obtained patterns by approximating simple structures (Thurstone 1947). Mathematically, varimax rotation seeks to maximize the summed variances of squared loadings which is achieved by (i) restricting rotated EOFs to be composed by only a few numbers of grid points with high loadings while the remaining grid points exhibit near-zero loadings and by (ii) limiting each grid point to contribute to only one rotated EOF while having near-zero loadings for the other EOFs. Since nonrotated EOFs are typically dense (i.e., consisting of mostly nonzero values), varimax rotation produces more sparse EOFs containing mostly zero or close-to-zero values, leading to spatially compact structures, which allow a clearer interpretation. Promax oblique rotation builds upon the varimax solution by raising the rotated, normalized EOFs to the power $p \geq 1$ while retaining the original sign, thus further reducing low loading compared to high loading of the EOFs. Promax can be understood as an oblique generalization, with $p = 1$ yielding a varimax orthogonal solution. Hendrickson and White (1964) provides a value for p, which the authors consider appropriate for most applications ($p = 4$). In an extensive review Richman (1986) points out, however, that the promax rotation using $p = 2$ consistently performs better, which is what we will use in this paper. To keep the paper self-contained, we provide a brief summary of both rotation criteria in appendix B.

The main difference between both rotation types is that promax allows rotated PCs to be correlated with each other, with higher values of p typically leading to stronger correlations.

[3] Loaded EOFs are weighted by the square root of the corresponding singular value.

In contrast, varimax solutions yield always uncorrelated PCs. For both varimax orthogonal and promax oblique rotation, the obtained EOFs are no longer orthogonal. The question of which rotation method is the most suitable for a given analysis remains unsettled in the literature. In reality, we do not expect geophysical signals to be perfectly uncorrelated, which generally argues for applying an oblique rotation. Nevertheless, Finch (2006) showed that varimax orthogonal and promax oblique solutions perform similarly, in particular when the PCs obtained by the oblique solution exhibit low linear Pearson correlation coefficients. In the presence of simple structures, however, promax oblique rotation performs better by effectively reducing the number of grid points that contribute to each mode, hence further simplifying the EOFs and increasing correlations among the PCs (Finch 2006). Therefore, the decision on how many EOFs to rotate and which rotation type to perform is a choice to be taken case by case; we will explore this in section 4b.

3. Data

To test our method, we apply it to artificial and real climate data. For the artificial datasets, we consider complex MCA without rotation, as studies already exist that demonstrate the better interpretability and lower sensitivity to sampling errors of the varimax-rotated solutions compared to the unrotated EOFs. (Lian and Chen 2012; Richman 1986; Cheng et al. 1995). By means of two synthetic experiments we seek to illustrate the advantages and caveats of complex MCA. In a first experiment (experiment I), we test the performance of complex MCA compared to standard MCA considering time-lagged signals. In a second experiment (experiment II), we investigate how the theta extension can improve the result of complex MCA to nonstationary processes. Finally, we apply complex MCA with rotation to climatic data that we expect to have intrinsic geophysical cycles but are also affected by the nonstationarity of climate change.

a. Synthetic data

We create two 2D spatiotemporal data fields X_A, X_B with coordinates representing longitude $\lambda \in [0, 359]$ and time $t \in [0, 364]$ days. The data generation model follows

$$X_s(t,\lambda) = c_s(\lambda)\zeta_s(t,\lambda) + \varepsilon(t,\lambda), \tag{8}$$

where $\zeta_s(t, \lambda)$ depends on the individual experiment design, $c_s(\lambda) = \cos^2 a_s\lambda$ represents a "zonal" modulation factor with scale factor a_s, and $\varepsilon(t, \lambda)$ denotes Gaussian white noise with zero mean and variance of 1.

The idea of experiment I is to highlight the advantage of using complexified fields compared to standard MCA in the presence of moving patterns and phase-shifted, stationary fields. Therefore, we define both signals as traveling waves $\zeta_s(t, \lambda) = \cos(\omega t + k_s\lambda + \phi_s)$. Our parameter choices are motivated by the Madden–Julian oscillation (MJO), which is an eastward propagating mode of deep convection and associated zonal wind circulation in the tropical atmosphere (Madden and Julian 1971; Zhang 2005). As such, the MJO is one of the dominant drivers of intraseasonal variability in the tropics characterized by a zonal propagation period of about 30–90 days. In general, MJO events tend to dominate over the Indian Ocean and the western Pacific before they decay toward the eastern part of the Pacific. Furthermore, when observing the MJO through different variables, the spatial scale of MJO events can vary. While the zonal wind circulation typically exhibits a wavenumber of about 1, the convective precipitation patterns may have zonal wavenumbers of 1–3. With this in mind, we fix the model parameter to $\omega = 2\pi/56$ days^{-1} (representing a period of 56 days), $k_A = -2\pi/360$, $k_B = -6\pi/360$ (representing wavenumbers 1 and 3, respectively) and phase shifts $\phi_A = 0$ and $\phi_B = \pi/2$. Finally, the modulation factor $c_s(\lambda)$ can be thought of describing regions of enhanced and suppressed MJO activity in the tropics. We therefore choose the scale factor $a_A = \pi/360$ and $a_B = 2\pi/360$ representing 1 and 2 regions of enhanced activity, respectively. (Fig. 2a).

In experiment II we investigate the response of complex MCA to noncyclical, nonstationary signals. Since MCA seeks to maximize covariance through a new set of linear combinations, we restrict ourselves here to linear trends (Table 1). To keep the experiment as simple as possible, we also consider static fields only. Then, the signal may be simply defined as $\zeta_s(t, \lambda) = \zeta_s(t) = h_s t$, where $h_A = 3/365$ day^{-1} and $h_B = 5/365$ day^{-1} describe linear trends. In this case, the modulation factor $c_s(\lambda)$ may describe, for instance, differential heating of the continents and the oceans due to global warming. The actual values of the scale factors remain unchanged compared to experiment I (Fig. 3a). For a better overview, the parameters of both experiments are summarized in Table 1.

b. Climate data

We analyze monthly means of global sea surface temperature (SST) and continental precipitation using the extended ERA5 dataset from 1950 to 2019 (Hersbach et al. 2019; Bell et al. 2020) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) as a state-of-the-art replacement of the ERA-Interim reanalysis (Dee et al. 2011). In general, trends and low-frequency variability of surface temperature and humidity are represented well for the period 1979–2019 (Hersbach et al. 2019; Simmons et al. 2021). However, Simmons et al. (2021) noticed strong biases in temperature and humidity records for central and East Africa in the 1950s. Via inspection, we found that East African precipitation records in particular are strongly biased during that period toward much higher values. To avoid providing our theta model extension with an incorrect starting point of the time series, we remove the first 10 years, providing us data from 1960 to 2019. Furthermore, the SVD of the covariance matrix is a rather memory intensive numerical operation, which is why we limit the domain of interest from 40°S to 60°N with a $1° \times 1°$ spatial resolution, guaranteeing that most of the continents will be included into the analysis. In total, the fields of SST and continental precipitation cover $n_A = 21\,816$ and $n_B = 14\,184$ grid points, respectively. To align the different temporal scales of highly variable precipitation and slow-varying SST and to filter out the high-frequency signals, we smooth both datasets with a 6-month moving average, for each month taking into account the three preceding and the two following months.

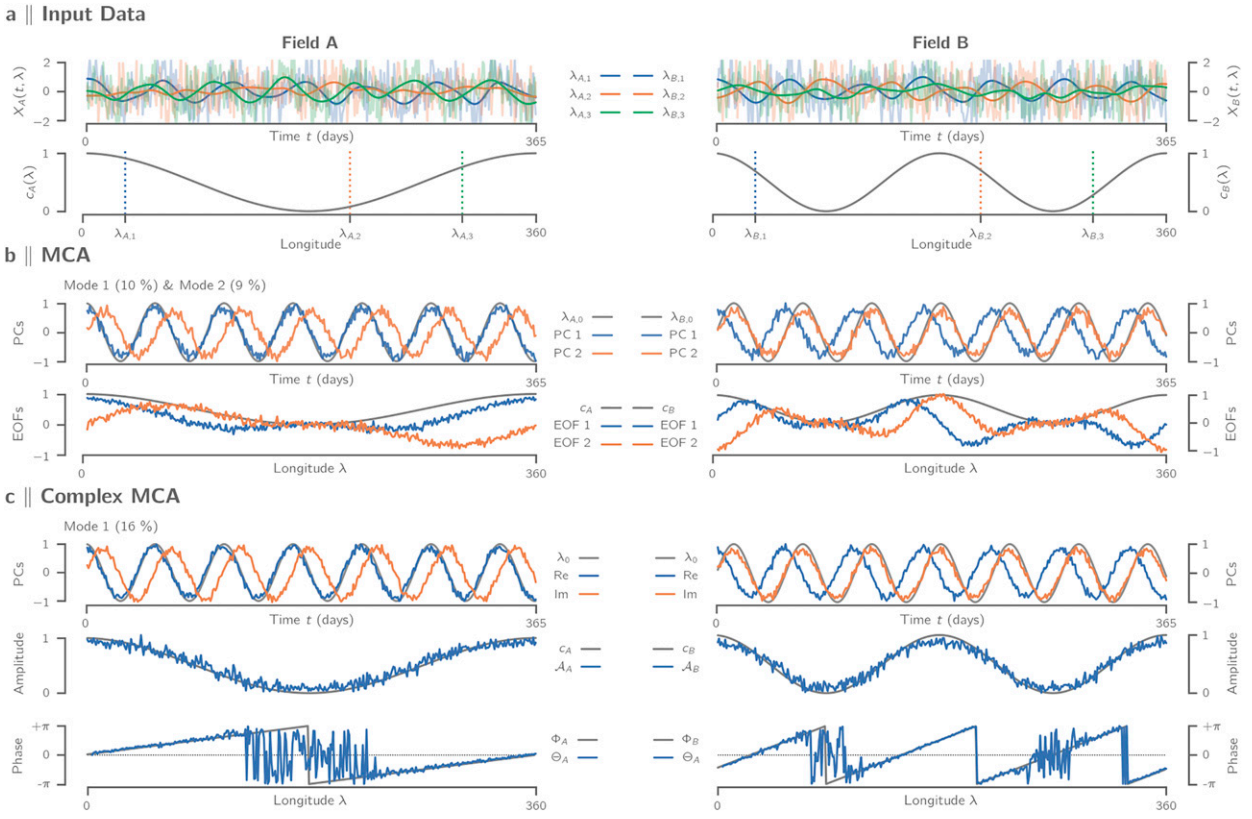

FIG. 2. Results of experiment I. (a) Illustration of synthetic data fields (left) A and (right) B showing (top) the temporal evolution $X_s(t, \lambda)$ for some selected longitudes λ smoothed via a weighted running average using a Gaussian kernel with a standard deviation of 7 days and (bottom) the associated modulation patterns $c_s(\lambda)$for both fields, respectively. (b) Standard MCA solution showing mode 1 (explaining 10% of the total covariance) and mode 2 (9%) with (top) the associated PCs and (bottom). (c) Complex MCA solution showing mode 1 (16%) with (top) the associated real (Re) and imaginary (Im) parts of the first PC, (center) the spatial amplitude functions, and (bottom) the phase functions. For comparison the denoised signal $X_s(t, \lambda_{s,0} = 0)$ (gray line; λ_0), the spatial modulation pattern $c_s(\lambda)$(gray line; c_s), and the theoretical phase shift $\Phi_s = k_s\lambda - \phi_s$ (gray line; Φ_s) are shown for both fields, respectively. The phase functions indicate the phase shift that has to be applied to the time series at a specific longitude λ in order to obtain the real part of the nonshifted PC (shown in top panel). Percentages in parentheses represent the covariance fraction γ_k. All EOFs and spatial amplitude functions are max-normalized for the sake of readability.

We choose this particular time window to effectively filter out the semiannual cycles present in the geophysical observations. We then normalize both data fields by dividing each grid point by its temporal standard deviation to ensure equal spatial importance. In a nonstandardized MCA analysis, regions of high rainfall, like the tropics, would dominate the covariance patterns, since annual rainfall in the midlatitudes and subtropics is typically much lower. After normalization, we weight the data points located on the regular $1° \times 1°$ grid according to their associated area on a sphere (North et al. 1982) by multiplying each grid point with $\sqrt{\cos(\varphi_j)}$, φ_j being the latitude at grid point j.

Climate indices of monthly means, used in this study for the sake of comparison, are downloaded from the websites of the National Oceanic and Atmospheric Administration (NOAA; https://psl.noaa.gov/data/climateindices/list/), namely the oceanic Niño index (ONI) provided by the NOAA Climate Prediction Center, the Pacific decadal oscillation (PDO; Mantua and Hare 2002), and the Atlantic meridional mode (AMM; Chiang and Vimont 2004). We calculate the ENSO Modoki index (EMI; Ashok et al. 2007) according to EMI = $R_A - (1/2)R_B - (1/2)R_C$, where R denotes the area-averaged ERA5 SST anomalies over the regions A (165°E–140°W, 10°S–10°N), B (110°–70°W, 15°S–5°N), and C (125°–145°E, 10°S–20°N), respectively. Furthermore, we define an oceanic warming index (OWI) as the area average of ERA5 SST anomalies (40°S–60°N) using the same latitude area weighting as for the MCA data preparation. Finally, all indices are smoothed using a 6-month moving average in alignment with our data preprocessing. We

TABLE 1. Models for synthetic data generation of X_s for experiments I and II (section 3a). For our final model, we consider the coordinate ranges $t = 0, 1, \ldots, 364$ days and $\lambda = 0, 1, \ldots, 359$ (without unit).

		Experiment I			Experiment II
Field	a_s	ω (day^{-1})	k_s	ϕ_s	h_s (day^{-1})
A	$\pi/360$	$2\pi/56$	$-2\pi/360$	0	3/365
B	$2\pi/360$	$2\pi/56$	$-6\pi/360$	$\pi/2$	5/365

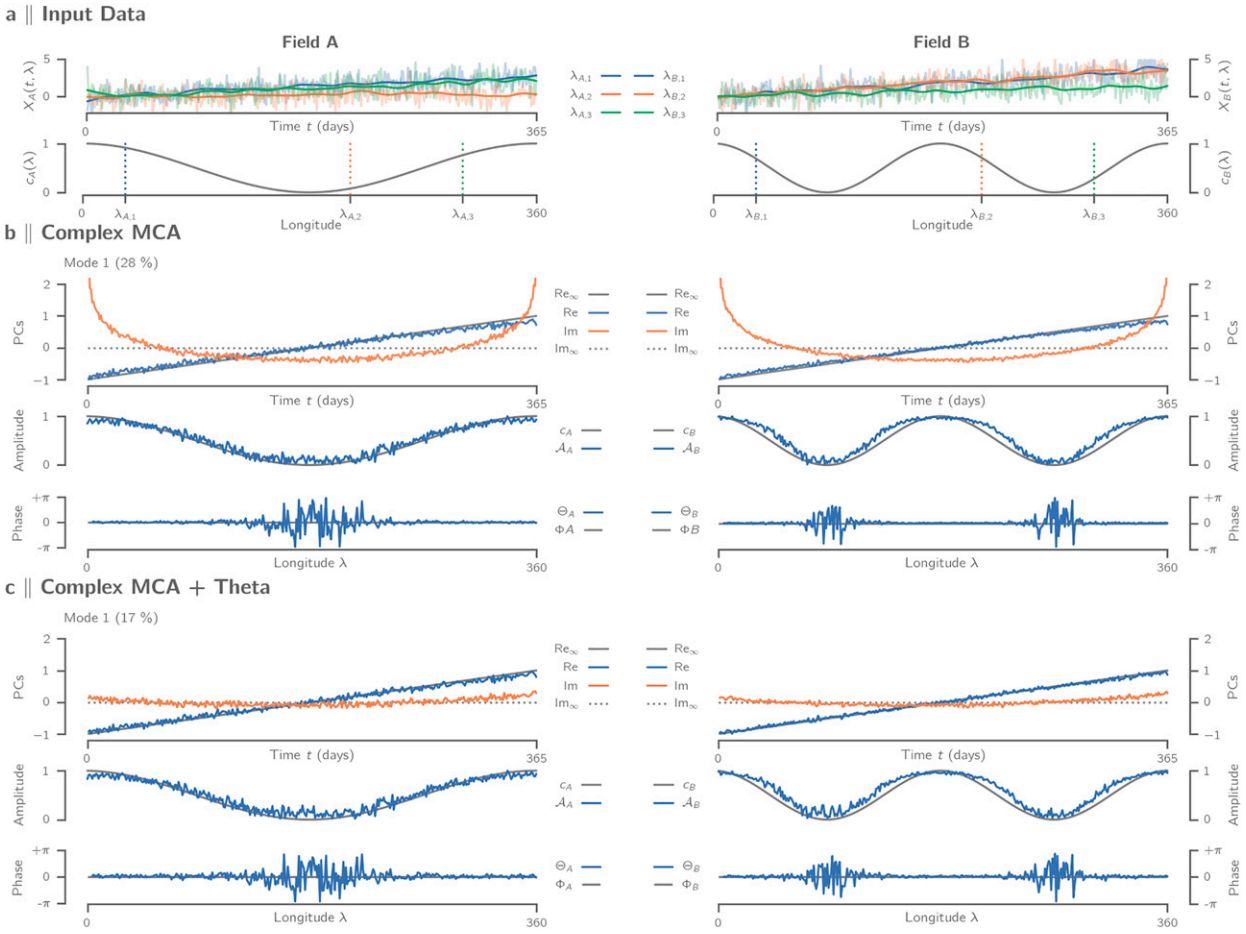

FIG. 3. Results of experiment II. (a) Illustration of synthetic data fields (left) A and (right) B showing (top) the temporal evolution $X_s(t, \lambda)$ for some selected longitudes λ smoothed via a weighted running average using a Gaussian kernel with a standard deviation of 7 days and (bottom) the associated modulation patterns $c_s(\lambda)$ for both fields, respectively. (b) Complex MCA solution showing mode 1 (explaining 28% of the total covariance) with (top) the associated real and imaginary parts of the first PC, (center) the spatial amplitude functions, and (bottom) the phase functions. (c) As in (b), but showing mode 1 (17%) for the theta extended complex MCA. For comparison theoretical real and imaginary part of a infinite linear trend (gray lines; Re_∞ and Im_∞), the spatial modulation pattern $c_s(\lambda)$ (gray line; c_s) and the theoretical phase shift $\Phi_s = 0$ (gray line; Φ_s) are shown for both fields, respectively. Percentages in parentheses represent the covariance fraction γ_k. All spatial amplitude functions are max-normalized for the sake of readability.

further center and max-normalize all indices for better comparison with our obtained PCs.

4. Results

In the following, we discuss the results from the synthetic experiments before investigating the results of the analysis of SST and continental precipitation.

a. Synthetic data

Before we start the investigation of the synthetic data, we apply standard and complex MCA as a baseline study on Gaussian white noise. By comparing the individual singular values as well as their sum obtained for each method (Table 2), we observe that complex MCA consistently yields higher covariance both for each mode individually and overall, as a result of the additional time-lagged cross-covariance between fields A and B. Therefore, when comparing standard and complex MCA in the following, we will make use of the singular values directly instead of the covariance fraction γ_k in order to assess the explained covariance by each mode.

1) EXPERIMENT I: LAGGED SIGNALS

Applying standard MCA (denoted by subscript std), the traveling wave is split into two modes the first explaining 10% of the shared covariance, and the second 9% (Fig. 2b). This is not surprising since the singular vectors are orthogonal and the associated PCs are uncorrelated, forcing the signals into two distinct modes following a sine and a cosine, which are perfectly uncorrelated for a full number of cycles and thus orthogonal. While the first two PCs correctly indicate the temporal evolution of the field following a sine/cosine, the

TABLE 2. Singular values of MCA (standard), complex MCA (complex), and theta extended complex MCA (complex + theta) obtained for purely white noise random fields (noise) and the synthetic experiments described in section 4a (experiments I and II). Covariance fraction γ_k for mode k (in % shown within parenthesis) is calculated according to Eq. (3). Some exemplary modes (in boldface) are depicted in Figs. 2 and 3. Total cumulated covariance is shown in the last row.

	Noise		Experiment I		Experiment II		
Mode	Standard	Complex	Standard	Complex	Standard	Complex	Complex + theta
σ_1	2.58 (1.0)	8.37 (1.3)	**27.43** (9.8)	**102.77** (16.1)	89.1 (31.4)	**179.82** (28.2)	**93.69** (17.0)
σ_2	2.53 (0.9)	8.11 (1.3)	**24.05** (8.6)	6.97 (1.1)	1.95 (0.7)	6.19 (1.0)	6.18 (1.1)
σ_3	2.50 (0.9)	7.84 (1.2)	2.21 (0.8)	6.82 (1.1)	1.92 (0.7)	5.96 (0.9)	5.97 (1.1)
⋮	⋮	⋮	⋮	⋮	⋮	⋮	
$\sum_j \sigma_j$	269.91	629.95	279.96	640.02	284.1	637.97	551.16

EOFs fail to provide a clear and easy interpretation of the associated spatial structures. In particular, without further a priori knowledge of the expected covarying structures, it seems a trying exercise to deduce the traveling wave from the obtained PCs and EOFs. Apart from that the remaining modes do not show any more distinct patterns, basically representing noise.

In comparison, complex MCA (denoted by subscript com) allows us to represent the traveling wave by a single mode (Fig. 2c). Additionally, the varying strength of the signal at different longitudes is clearly captured by the spatial amplitude for both fields. Another advantage in the interpretation compared to standard MCA is the spatial phase, which shows the migration of the signal to higher longitudes. Identical phase values at different longitudes indicate correlation while a phase shift of π represents anticorrelation. We note that we have not used the theta model extension here, since the signals in this experiment are stationary and spectral leakage due to non-integer number of periods can be assumed to be marginal. Furthermore, comparing the singular values $\sigma_{\text{com},1} > \sigma_{\text{std},1} + \sigma_{\text{std},2}$ (Table 2) indicates that the complex mode 1 accounts for more covariance and thus captures more of the traveling wave than the combined mode 1 and 2 of the standard solution. This makes sense, since the contributions of longitudes to the individual EOFs 1 and 2 of the standard solution will decrease for phase shifts which are not equal to 0 (i.e., sine) or $\pi/2$ (i.e., cosine). As a pitfall of complex MCA, however, it should be noted that longitudes whose spatial amplitude function is very low tend to have noisy phase function values. In our experiment, this is obvious for longitudes of both fields where $c_s(\lambda) \approx 0$ and the associated phase function does no longer follow the positive linear trend. At these longitudes, the max-normalized spatial amplitude function falls below 0.2 for both fields, serving as a general orientation for the consideration of phase function values in the following analysis. In general, it is therefore advisable not to consider regions with low amplitudes.

2) EXPERIMENT 2: TRENDS AND NONCYCLIC SIGNALS

As discussed in section 2b, nonperiodic behavior usually leads to undesired boundary effects caused by spectral leakage in the frequency domain. In our experiment, this effect is clearly evident for both PCs of the complex MCA (denoted by subscript com) (Fig. 3b). Note that although the boundary effects seem to occur only in the imaginary parts of the PCs, this may not be the case in general. The theta extended complex MCA (denoted by subscript thc), on the other hand, successfully mitigates the boundary effects of the PCs (Fig. 3c), where we have set the theta period $T_s = 1$ as the time series have no seasonality. Nevertheless, the spatial patterns are similar for both methods and can hardly be distinguished visually. Considering only spatially static fields in this experiment, the phase function is constant for both fields, with exceptions at longitudes of low-amplitude values as already mentioned for experiment I. Examining the singular values, we notice that $\sigma_{\text{com},1} \gg \sigma_{\text{thc},1} \approx \sigma_{\text{std},1}$ (Table 2), indicating the increased covariance of the standard Hilbert transform due to boundary effects compared to the theta model extended Hilbert transform. This implies that boundary effects created by the Hilbert transforms can strongly affect correlations and, depending on their magnitude, lead to a severe ''inflation'' of the singular values. As a consequence, the boundary effects can appear as parts of the first modes, thus completely misleading the interpretation of the results.

More generally, the frequency spectrum of a linear trend on a given interval is typically broadband and thus the analytical signal constructed by the Hilbert transform has not a physical interpretation in terms of characteristic frequencies. Therefore, the phase function cannot be interpreted in terms of a physical phase shift. The only exception is for a phase shift of $\theta = 0; \pm\pi$ since these relative phase shifts represent correlating and anticorrelating signals disregarding the mathematical nature of the phase. A fundamental consequence of this is that for modes whose PC is broadband (e.g., a trend), only correlating ($\theta \approx 0$) as well as anticorrelating ($\theta \approx \pm\pi$) patterns should be considered.

b. *Climate data: SST and continental precipitation*

We apply theta extended, complex MCA to SST and continental precipitation using a theta period $T_s = 12$ to account for the seasonal cycle. The dimensionality of the problem can be greatly reduced with the first 72 modes explaining more than 99% of the lagged covariance (Fig. 4a). To simplify the obtained patterns and to increase the physical meaning of the results (section 2c), we rotate the first 150 modes explaining 99.82% of the lagged covariance. Our decision to rotate 150 modes is motivated by the idea of retaining as much information as possible without including the noise that dominates the higher modes. Since we observe a marked drop in singular values at

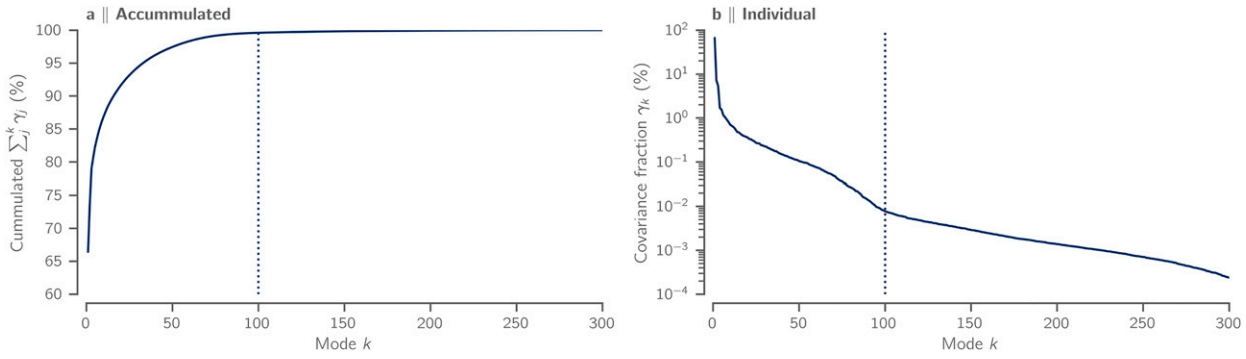

FIG. 4. Described covariance of the first 300 modes of complex MCA applied to SST and continental precipitation showing the (a) accumulated and (b) individual share. Dashed line estimates the boundary of suspected ''noisy'' modes.

about the mode number 100 followed by an exponential decrease in singular values (Fig. 4b), we guess that there is a negligible information content in higher modes. Moreover, the quality of the reconstructed signal is rather independent of the exact number of rotated modes, with 150 ± 50 yielding basically identical results for the first six modes for both varimax orthogonal and promax $p = 2$ oblique rotation. Therefore, we will restrict our discussion in the following to the first six modes. To address the question of the rotation method to be chosen, we note that promax oblique rotation performs better when simple structures are present and correlations among PCs are high (Finch 2006). We expect our first mode to be dominated by the shared dynamics of the seasonal cycle of both SST and continental precipitation, which is indeed what we find for both promax (not shown here) and varimax solution (Fig. 5). This mode, however, is fairly global and as such does not represent a simple structure. Furthermore, we observe that the correlations among the first six promax-rotated PCs range from −0.15 to 0.13 only, underlining that the promax oblique solution is close to orthogonal and differences from the

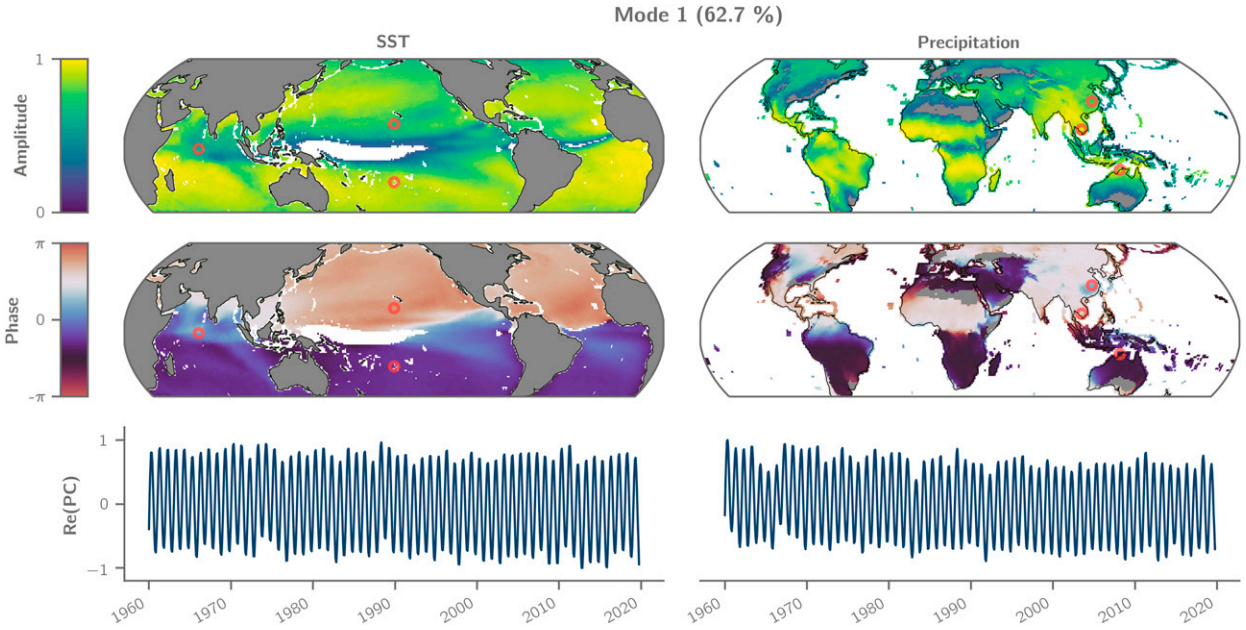

FIG. 5. Complex varimax-rotated MCA of (left) SST and (right) continental precipitation showing mode 1 and its relative importance indicated by the covariance fraction. (top) The amplitude functions show the regions predominantly contributing to the mode. (middle) The phase functions depict the relative phase shifts with respect to (bottom) the corresponding real part of the PC, where 0 (blue) means correlation and $\pm\pi$ (red) shows anticorrelation. For each grid point, the corresponding PC can be computed from the given phase value by applying the negative phase shift to the PC of phase 0 (bottom). The amplitude functions, PCs and indices shown are all max-normalized for the sake of comparability. In both amplitude and phase function regions with a max-normalized amplitude below 0.25 are masked out. Red circles mark North Pacific (15°N, 170°E), Indian Ocean (0°, 70°W), and South Pacific (20°S, 160°E) for SST and Poyang Lake, China (29.1°N, 116.3°W), Phnom Penh, Cambodia (11.6°N, 104.9°W), and Darwin, Australia (12.5°S, 130.8°W) for precipitation. The phase shifted PCs of these locations are examined in Fig. 6.

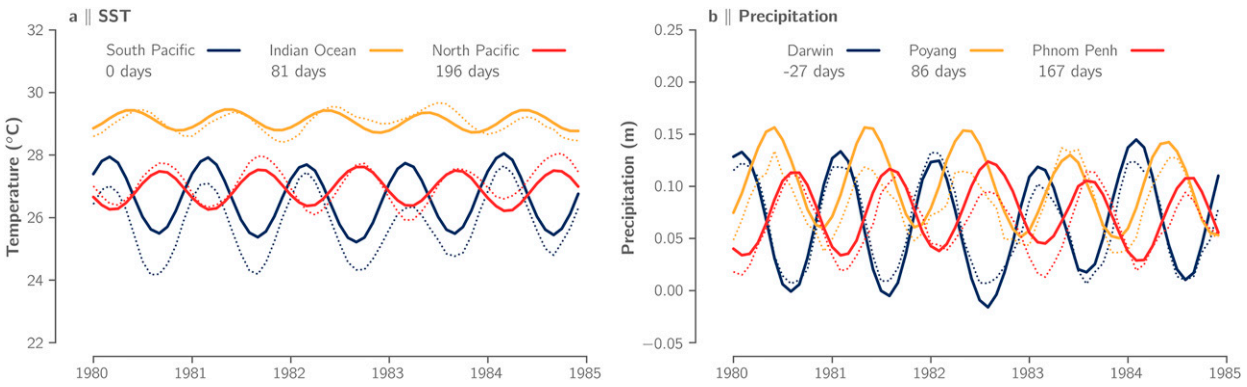

FIG. 6. Comparison between 6-month moving average of ERA5 SST and continental precipitation (dotted lines) and the reconstructed time series based on mode 1 only (thick lines) for the exemplary locations defined in Fig. 5 covering the years 1980–85. Days in legend refer to the time shift τ of the individual locations with respect to the SST of the South Pacific derived from the spatial phase functions (Fig. 5).

varimax solutions only marginal, at least for the first six modes. Since the promax oblique solution seemingly does not provide a better result, we opt for the somewhat simpler varimax orthogonal rotation. In the following, we will discuss the varimax-rotated modes by investigating the real part of the PCs, the spatial amplitude and phase functions for both fields, SST and continental precipitation, respectively.

In our representations of the modes, we remove nonsignificant, ''noisy'' phase values (see section 4a) by masking out regions in the spatial amplitude and phase function exhibiting a max-normalized amplitude of <0.25.

Mode 1 describes 62.7% of the covariance between SST and continental precipitation, clearly showing the annual cycle (Figs. 5 and 12a). As expected, the annual cycle shows itself in both variables on a global scale, with the exception of the equatorial ocean, where the seasonal SST variations are only weak. The phase function correctly identifies the anticorrelation between the northern and southern oceans. It also suggests that the eastern equatorial Pacific and the equatorial Indian Ocean nevertheless show a weak seasonal signal, which is, however, positively phase shifted relative to the rest of the southern ocean. On the continents, precipitation dominates mainly in monsoon areas and the tropics. The phase function illustrates the division into June–August (white) and December–February (black) dominated rainfall systems and identifies corresponding transition zones, as for example over the South American rain forest and central North America. It also highlights some interesting dynamical regions that stand out from their respective environments, namely, the west and east coasts of North America, the Mediterranean region, and, to a lesser degree, the East Asian monsoon region in China.

The obtained mode provides an instructive example to highlight the benefits of the complexified approach due to the mode's corresponding narrowband frequency spectrum (Fig. 12a). Using the dominant periodicity of mode 1, $T_1 = 12$ months, the interpretation of a phase shift θ, given by the spatial phase function, as time lag τ is straightforward and can be computed via $\tau = T_1\theta/(2\pi)$. Doing this for some exemplary locations (denoted by red circles in Fig. 5), we observe that the seasonal SST maximum of the North Pacific follows the one in the South Pacific by 196 days, that is, being almost perfectly anticorrelated (Fig. 6). The relative time shifts of the SST and precipitation maximum of 81 days in the Indian Ocean and 86 days at Poyang Lake, China, is close to 3 months and as such translates to a phase shift of about $-\pi/2$, something that cannot be picked up as feature within one mode when using standard MCA. It should also be noted that although the sampling frequency of SST and precipitation is monthly, the phase function is continuous and thus allows us to infer time lags on shorter time scales. This is, for instance, the case for the seasonal precipitation maximum in Darwin, Australia, which precedes the seasonal SST maximum of the South Pacific by 27 days.

Mode 2 (7.0%) can be clearly associated with the ocean–atmosphere phenomenon of ENSO (Figs. 7 and 12b). During El Niño, higher SST in the central and eastern Pacific and lower SST in the western Pacific positively correlate with heavier rainfall within a narrow band along the west coast and the southeastern coast of South America (Tedeschi et al. 2013), the east and west coasts of North America (Ropelewski and Halpert 1986), the East Asian monsoon region (Wen et al. 2019), and the Horn of Africa (Indeje et al. 2000). At the same time, precipitation decreases in northern South America (Tedeschi et al. 2013), Oceania (Dai and Wigley 2000), South Africa (Gaughan et al. 2016), and the Indian monsoon region (Cherchi and Navarra 2013). During La Niña (the counterphase of El Niño), these correlations are reversed. Recently, similar teleconnections have been identified via event coincidence analysis (Wiedermann et al. 2021). For a current summary of established ENSO-related rainfall patterns during El Niño and La Niña see Lenssen et al. (2020).

Our result also shows the dynamical link between ENSO in the Pacific Ocean and the Indian Ocean (Krishnamurthy and Kirtman 2003), the South China Sea (Klein et al. 1999), and the tropical North Atlantic (Enfield and Mayer 1997; Saravanan and Chang 2000; Alexander and Scott 2002; Chiang and Sobel 2002) in accordance with previous studies. Interestingly, it was found that the ENSO-related SST teleconnections in the remote oceans often occur with some delays, with the Indian

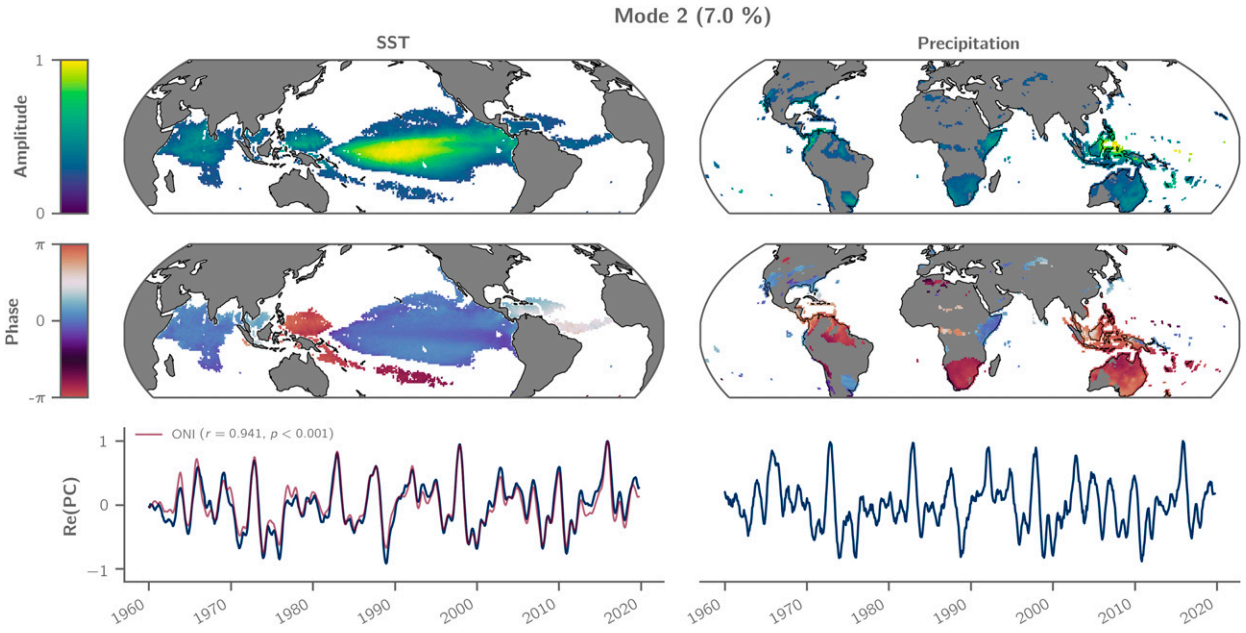

FIG. 7. As in Fig. 5, but showing mode 2 corresponding to the oceanic Niño index (ONI) provided by the NOAA Climate Prediction Center as described in section 3b. The Spearman correlation between ONI and the real part of SST PC2 is $r = 0.941$ with p value < 0.001.

Ocean typically peaking ~3 months and the South China Sea and tropical North Atlantic ~5 months after the SST peak in the Pacific ENSO region (Enfield and Mayer 1997; Klein et al. 1999; Saravanan and Chang 2000). The mechanism behind these lagged responses, known as the atmospheric bridge, is based on the characteristic atmospheric circulation during El Niño, which causes changes in cloud cover and evaporation over the remote oceans, leading to increased net heat flux and SSTs (Lau and Nath 1996; Klein et al. 1999). However, due to the broadband frequency spectrum of the SST PC (Fig. 12b), with most of the energy contained at four different peaks around 2.5, 3.5, 5, and 11 years, the phase cannot simply be translated into a time shift. Nevertheless, the tropical North Atlantic clearly exhibits more positive phase values compared to the Pacific El Niño region, therefore indicating to be positively time shifted relative to the Pacific.

Mode 3 (6.9%) represents global warming and the associated changes in precipitation patterns (Fig. 8). The warming SST patterns clearly emerge in all major ocean basins, although more pronounced in the Northern Hemisphere due to the asymmetric response of the northern and southern trade winds to global warming (Xie et al. 2010). We also note a pronounced warming of the western part of both the Pacific and the Atlantic basins, both regions of enhanced ocean heat transport (Xie et al. 2010; Hannachi and Trendafilov 2017). Similar to these oceanic trends, we also observe global trends in the precipitation patterns, with decreasing rainfall over the Mediterranean, South Africa, Australia, South America, and parts of western North America. There seems to be also a decrease of rainfall over the west Asian monsoon region. In contrast to that, the results suggest increased precipitation over the Indian monsoon region as well as some localized regions in Europe and in South and North America. These results are largely in agreement with studies based on observational data (Gu and Adler 2015) and, more recently, on CMIP5 climate simulations (Giorgi et al. 2019). Finally, it should be stressed that both PCs, SST and precipitation, only provide a meaningful interpretation for phases $\theta \approx 0; \pm\pi$ (correlation, anticorrelation), due to their broadband frequency spectra (Fig. 12c). For phase shifts different from that, no clear conclusions can be drawn, as for example is the case for equatorial Africa where the phase shift is approximately $+\pi/2$.

Mode 4 (1.6%) shows a slow-oscillating pattern of SST in the northeastern Pacific correlating with localized precipitation patterns distributed over all continents (Fig. 9). The typical spatial SST pattern is known as the Pacific decadal oscillation (PDO) (Mantua et al. 1997) and a well-established climate index. A combination of various processes originating in the tropics and extratropics has been proposed as the physical source of the PDO (Newman et al. 2016), with ENSO and PDO likely responding to the same forcing function (Pierce 2002). Our analysis, however, provides a means to disentangle ENSO- and PDO-related precipitation patterns, which are often similar for western North America (Hu and Huang 2009), although it also reveals important differences (e.g., for Australia, the Indian subcontinent, or the African Sahel region). Yet care must be taken when interpreting regions that have a phase shift different from 0, $\pm\pi$. Although the PDO exhibits a strong spectral energy at about 35 years, the mode contains also important features at about 1–4 years (Fig. 12d), making the phase interpretation physically less clear.

Mode 5 (1.6%) describes an oscillating SST anomaly mainly limited to the central Pacific (Figs. 10 and 12e) describing El Niño Modoki (Ashok et al. 2007) and represented by the EMI. Higher SST in the central Pacific and lower SST in the eastern Pacific correlate with reduced precipitation in the East

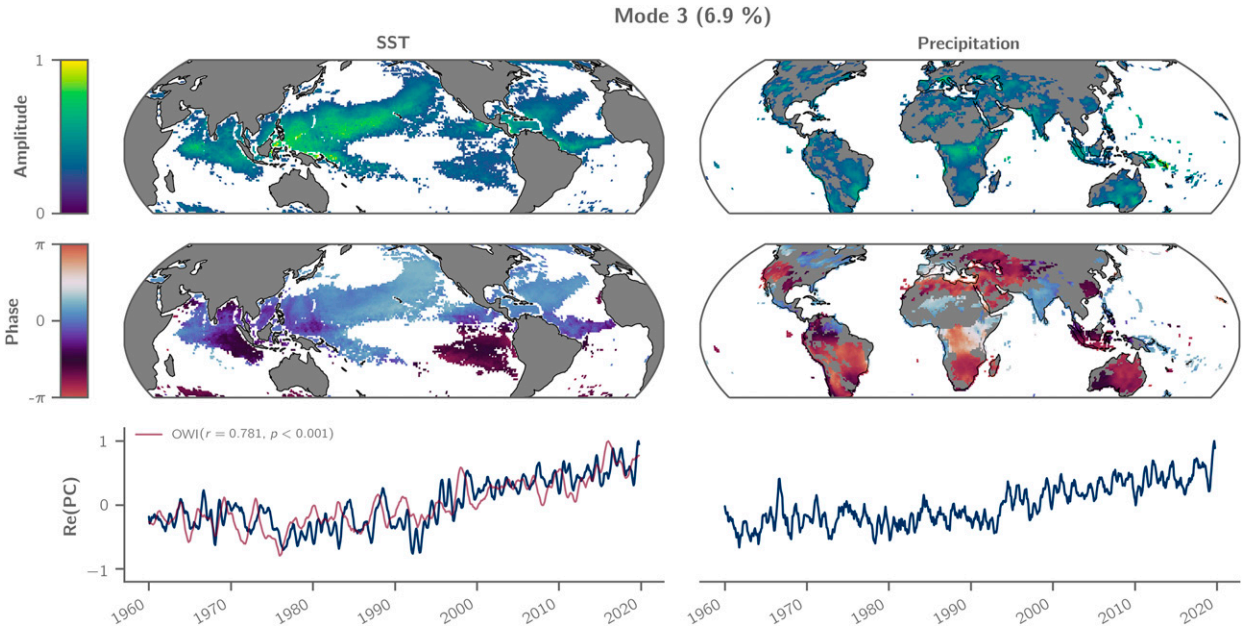

FIG. 8. As in Fig. 5, but showing mode 3 corresponding to the oceanic warming index (OWI) as described in section 3b. The Spearman correlation between the OWI and the real part of SST PC3 is $r = 0.781$ with a p value < 0.001.

Asian monsoon region (Feng et al. 2011), Australia (Taschetto and England 2009), parts of South America (Tedeschi et al. 2013), and South Africa (Ratnam et al. 2014) and vice versa. Some of these teleconnections have recently been revealed by event coincidence analysis (Wiedermann et al. 2021) although important links to the East Asian monsoon region, for example, were missing. Although our result suggests that continental rainfall in certain regions of Africa, Arabia, and the Americas are phase-shifted El Niño Modoki expressions, future work has to show if these weak amplitudes are significant.

Mode 6 (1.1%) is characterized by an SST pattern concentrated in the tropical and subarctic North Atlantic (Figs. 11 and 12f). The SST pattern, representing the Atlantic meridional mode (AMM), is the dominant coupled ocean–atmospheric

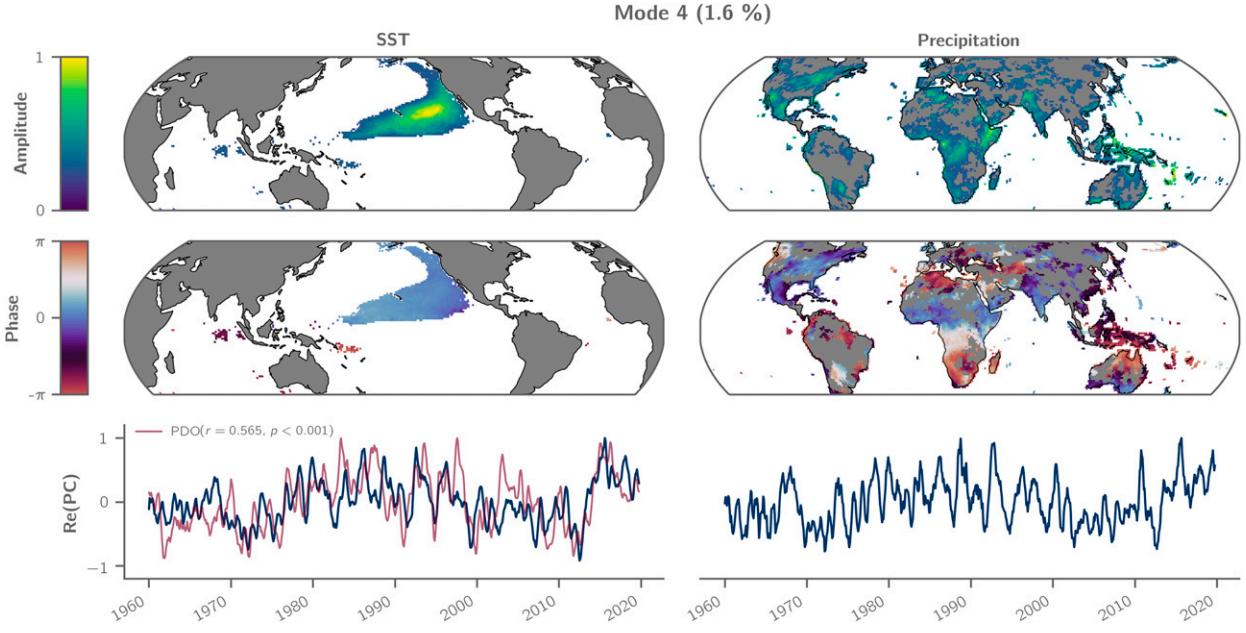

FIG. 9. As in Fig. 5, but showing mode 4 corresponding to the Pacific decadal oscillation (PDO; Mantua and Hare 2002) as described in section 3b. The Spearman correlation between the PDO and the real part of SST PC4 is $r = 0.565$ with a p value < 0.001.

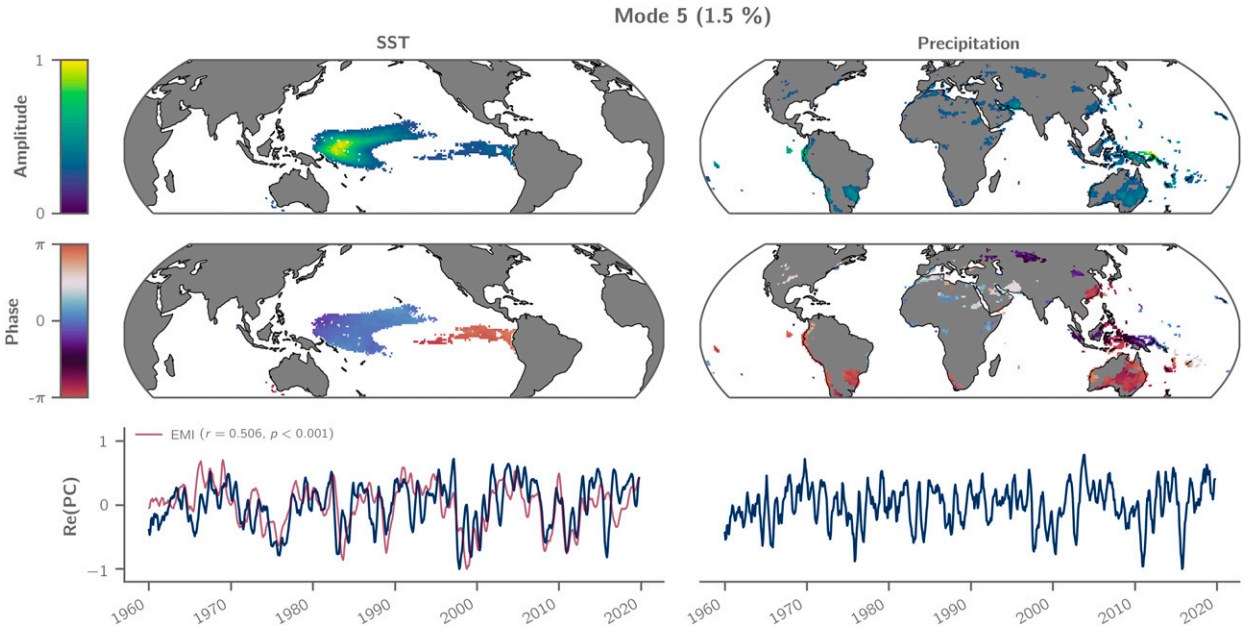

FIG. 10. As in Fig. 5, but showing mode 5 corresponding to the ENSO Modoki index (EMI; Ashok et al. 2007) as described in section 3b. The Spearman correlation between the EMI and the real part of SST PC5 is $r = 0.506$ with a p value < 0.001.

phenomenon in the tropical Atlantic (Servain et al. 1999) and its impacts on precipitation of the African Sahel zone, Central America, and northern South American continent are well known (Lamb et al. 1986; Martin et al. 2014). Only very recently, Vittal et al. (2020) found a link between the AMM and Indian summer monsoon rainfall, which they used to improve precipitation forecast. In addition, our result suggests more teleconnections between the AMM and precipitation variability over the Mediterranean, East Africa, the Congo basin, North America, and the East Asian monsoon region, providing much potential for advancing local rainfall predictions in those areas. Due to the missing intrinsic time scale with no clear periodicity (Fig. 12f), only correlating and anticorrelating patterns should be interpreted. However, most of

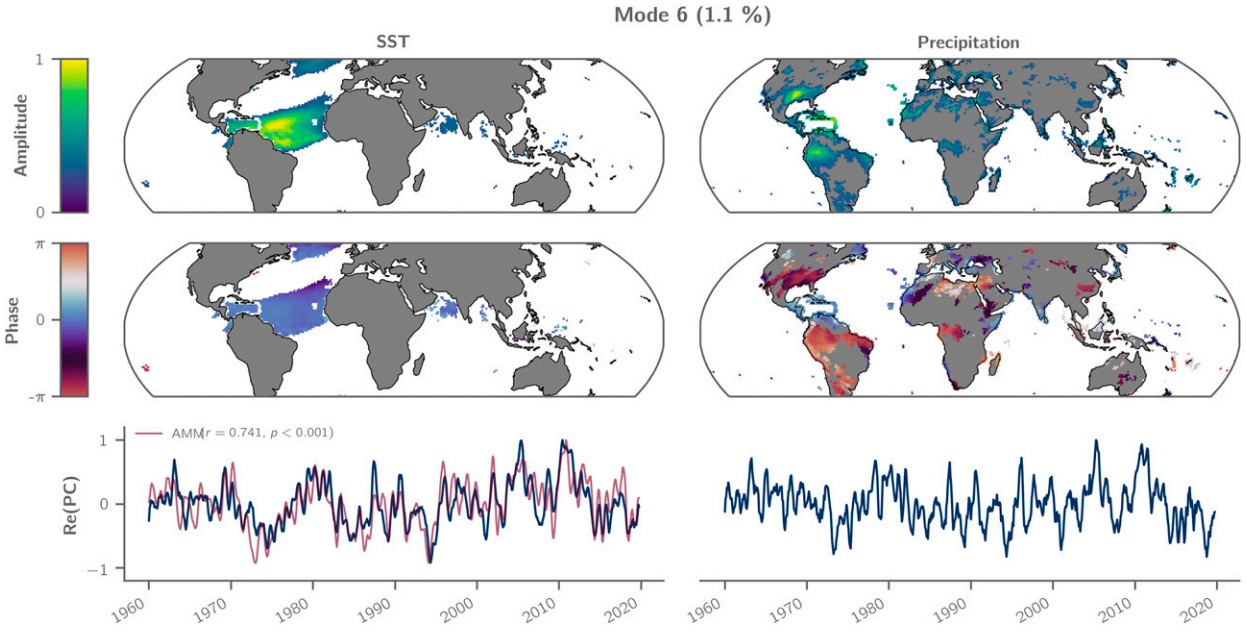

FIG. 11. As in Fig. 5, but showing mode 6 corresponding to the Atlantic meridional mode (AMM; Chiang and Vimont 2004) as described in section 3b. The Spearman correlation between the AMM and the real part of SST PC6 is $r = 0.741$ with a p value < 0.001.

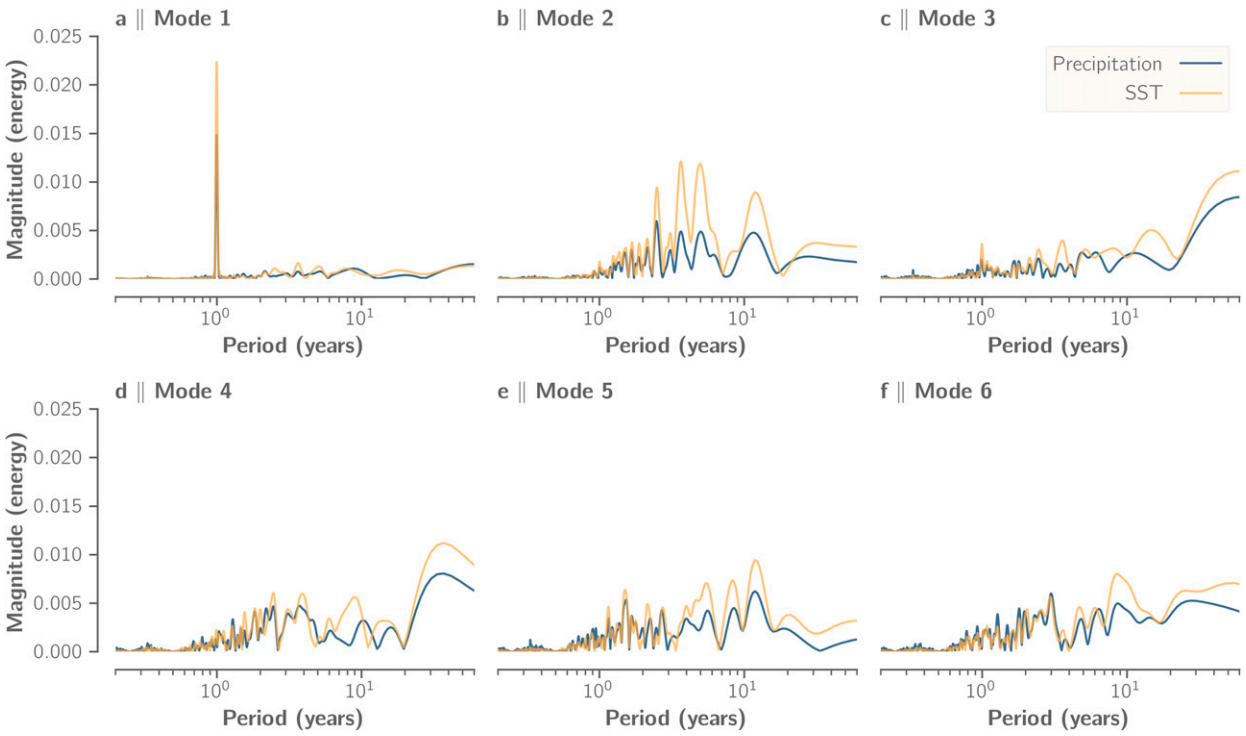

FIG. 12. Magnitude spectrum of the first six PCs considering only the real part and using a Hanning window.

the regions identified by this mode satisfy being correlated or anticorrelated.

5. Conclusions

Understanding the intricate climate system is a challenging task that requires advanced statistical methods. Finding correlations among a set of different climate variables is complicated by the frequently present lagged responses of different variables to the same forcing. We show that complex rotated MCA provides a practical tool to single out such modes from a high-dimensional data space, even when the narrowband assumption of the input signals is only partially satisfied.

By taking into account the spatial amplitude and phase function of the obtained complex modes, we obtain a simple approach to examine otherwise complicated spatial and temporal structures. Our synthetic experiments highlight that, in the case of phase-shifted signals, complex MCA can capture a more comprehensive and complete picture of the correlations present. However, they also show the sensitivity of the Hilbert transform to spectral leakage caused by boundary conditions of the given time series, such as when the time series clearly consists of a noninteger number of cycles and/or in the presence of trends. Since the stationarity assumption does often not hold (due to climate change), time series should generally be preprocessed when applying complex MCA.

Extending the time series via the optimized theta model mitigates the effect of spectral leakage and produces physically reasonable PCs. This procedure allows us to resolve trends and noncyclic signals, although for trends the obtained spatial phase function has a simple physical interpretation only for correlating and anticorrelating patterns. Nevertheless, this approach provides a means of applying complex MCA without the need to detrend the time series of interest. Moreover, excluding the first mode, the original fields can be reconstructed without the seasonal cycle, providing an advanced tool to preprocess time series containing nonstationary and nonlinear seasonal features.

A general caveat in complex MCA is the fact that the phase function loses its interpretation for PCs with a broadband frequency spectrum. But although the spatial phase function does not always have a simple physical interpretation for most of the modes due to their broadband frequency spectrum, complex MCA nevertheless can always be interpreted for the correlating and anticorrelating patterns.

Applying complex rotated MCA to SST and continental precipitation, we clearly identify the main shared dynamics in both variables, namely (i) the seasonal cycle, (ii) the canonical ENSO, (iii) the trends associated with global warming, (iv) the PDO, (v) ENSO Modoki, and (vi) the AMM. We also retrieve phase shifted signals between the two climate variables. While for the seasonal cycle these phase shifts can directly be translated into a time shift, the remaining modes generally do not lend themselves to such a simple interpretation due to their broadband frequency spectra. But even without a precise equivalent as time lag, the phase function provides a means to identify regions of lagged correlations, for instance between the SST of the Pacific and the tropical North Atlantic during El Niño events. In addition, by focusing on narrow frequency

bands only, one may detect frequency ranges over which the phase is not random, thus potentially uncovering more of the dynamic structure of the individual modes.

Many of the obtained correlation patterns in SST and continental precipitation had already been evidenced by a multiplicity of different, partly regional studies. The great advantage of complex rotated MCA is that it allows us to obtain all those patterns by a single analysis of the correlation of two geophysical variables at global scale in a more compact and easy-to-interpret way. Besides, our results also point out to new ocean–atmospheric teleconnections that, to our knowledge, have not been reported, most notably for the PDO and the AMM.

Regarding future applications of complex rotated MCA, this method has the potential for shedding light in the investigation of seasonal and subseasonal phenomena, as well as for spatially propagating patterns. As future work, we plan to analyze the Madden–Julian oscillation. Additionally, complex rotated MCA could be used to evidence other connections between less studied variables (e.g., sea surface salinity, sea surface height, soil moisture, winds, etc.), which has the potential to reveal new phenomena and novel aspects of existing or new teleconnections.

Acknowledgments. This work is part of the Climate Advanced Forecasting of Sub-Seasonal Extremes (CAFE) project and has been prepared in the framework of the doctorate in Physics of the Autonomous University of Barcelona. The authors gratefully acknowledge funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement 813844, and also from the Spanish government through the "Severo Ochoa Centre of Excellence" accreditation (CEX2019-000928-S). This work is a contribution to CSIC Thematic Interdisciplinary Platform TELEDETECT. The authors acknowledge the Copernicus Climate Change Service and the Physical Sciences Laboratory of NOAA for providing the data. The authors would also like to thank J. Ballabrera-Poy for his valuable comments on a previous version of the paper.

Data availability statement. The ERA5 data used in this study are publicly available from the Copernicus Climate Data Store at https://doi.org/10.24381/cds.f17050d7 as described in Hersbach et al. (2019). The ERA5 preliminary extension from 1950 to 1978 is described in Bell et al. (2020). All climate indices used in this study are openly accessible and can be found on the website of the Physical Sciences Laboratory of NOAA at https://psl.noaa.gov/data/climateindices/list/.

APPENDIX A

Definition of Arctan2

Every nonzero complex number in Cartesian coordinates, $z = x + iy$, can be transformed into polar complex coordinates, $z = ae^{i\theta}$, where $a = (x^2 + y^2)^{1/2}$ is the amplitude, and θ the phase of z. For each $z \neq 0$, the phase is only defined up to an integer multiple of 2π, resulting in an infinite number of possible values. To construct a well-defined function $\theta(x, y)$, one typically limits the phase θ to $(-\pi, \pi]$. Then, the two-argument arctangent function arctan2(y, x) converts the values of y and x to the polar phase via

$$\theta = \operatorname{arctan2}(y,x) = \begin{cases} \arctan\left(\frac{y}{x}\right), & \text{if } x>0 \\ \arctan\left(\frac{y}{x}\right)+\pi, & \text{if } x<0 \text{ and } y\geq 0 \\ \arctan\left(\frac{y}{x}\right)-\pi, & \text{if } x<0 \text{ and } y<0 \\ +\frac{\pi}{2}, & \text{if } x=0 \text{ and } y>0 \\ -\frac{\pi}{2}, & \text{if } x=0 \text{ and } y<0 \\ \text{undefined}, & \text{if } x=0 \text{ and } y=0 \end{cases}.$$

APPENDIX B

Finding the Rotation Matrix R

Let us assume a complex loading matrix $\mathbf{L}_r \in \mathbb{C}^{n\times r}$ containing only the first r columns (modes) that are to be rotated and n denoting the number of grid points. The number of grid points may be the sum of the grid points of two different fields $n = n_A + n_B$ as is the case for MCA and described by Eq. (7) or simply the total number of grid points if only one field is considered as it is for PCA. In the following, we will drop the subscript r in order to keep the notation simple and let $\tilde{\cdot}$ denote the rotated solutions. Furthermore, * refers to the conjugate transpose of a matrix and $|\cdot|$ denotes the absolute value of a complex number.

a. Varimax orthogonal rotation

The goal of varimax rotation is to approximate simple structures (Thurstone 1947) of the EOFs, which is achieved by simplifying the columns of $\mathbf{L}$ via an orthogonal rotation $\mathbf{R}$. For this purpose, Kaiser (1958) defines the simplicity S_k:

$$S_k = \frac{1}{n}\sum_{j=1}^{n}(|\tilde{l}_{jk}|^2)^2 - \frac{1}{n^2}\left(\sum_{j=1}^{n}|\tilde{l}_{jk}|^2\right)^2, \quad k = 1,\ldots,r, \tag{B1}$$

which measures the variance of the squared amplitude of the rotated loadings $\tilde{l}_{jk}$. With increasing variance, the squared rotated amplitudes $|\tilde{l}_{jk}|^2$ either become low or large, thus increasing simplicity. The normalized simplicity S then reads

$$S = \sum_{k=1}^{r}\left[\frac{1}{n}\sum_{j=1}^{n}\left(\frac{|\tilde{l}_{jk}|^2}{h_j^2}\right)^2 - \frac{1}{n^2}\left(\sum_{j=1}^{n}\frac{|\tilde{l}_{jk}|^2}{h_j^2}\right)^2\right], \tag{B2}$$

where $h_j = \left(\sum_{k=1}^{r}|l_{jk}|^2\right)^{1/2}$ represents the *communality* of grid point j, which is the amount of variance of the jth grid point accounted for by the r retained modes. Subsequently, the normalized varimax-rotated EOFs, $\tilde{\mathbf{L}}_{V,\text{norm}}$, are the solution to the varimax criterion

$$\tilde{\mathbf{L}}_{V,\mathrm{norm}} = \mathbf{H}^{-1}\mathbf{L}\mathbf{R} \quad \text{s.t.} \quad \arg\max_{l}(S) \text{ and } \mathbf{R}^{*}\mathbf{R} = 1, \tag{B3}$$

with the communality matrix $\mathbf{H} \in \mathbb{R}^{n\times n}$, whose elements are given by $\mathrm{diag}(h_1, \ldots, h_n)$. Equation (B3) can be solved by an iterative process, in which the EOFs are rotated in pairs in order to maximize S. Finally, the denormalized varimax-rotated EOFs can be computed via $\tilde{\mathbf{L}}_V = \mathbf{H}\tilde{\mathbf{L}}_{V,\mathrm{norm}}$.

b. Promax oblique rotation

Achieving simple structures with promax is done via an oblique Procrustes transformation (Hurley and Cattell 1962). Every target matrix of rotated EOFs $\mathbf{T}$ can always be approximated from a base matrix $\mathbf{B}$ via a linear transformation $\mathbf{R}$,

$$\mathbf{T} = \mathbf{B}\mathbf{R} + \mathbf{E}, \tag{B4}$$

where $\mathbf{E}$ is an error matrix. Minimizing trace$(\mathbf{E}^{*}\mathbf{E})$ yields the complex Procrustes equation

$$\mathbf{R} = (\mathbf{B}^{*}\mathbf{B})^{-1}\mathbf{B}^{*}\mathbf{T}. \tag{B5}$$

The basic assumption of promax is that a varimax orthogonal rotation is a reasonable approximation to an optimal oblique solution. Therefore, the base matrix is chosen to be $\mathbf{B} = \tilde{\mathbf{H}}^{-1}\tilde{\mathbf{L}}_V$ whose entries are normalized by the varimax communalities $\tilde{h}_j = \sum_{k=1}^{r} |\tilde{l}_{jk}|^2$. Then, the promax equation defines the elements of the target matrix $\mathbf{T}$,

$$t_{jk} = |b_{jk}^{+}|^{p+1}/b_{jk}^{+}, \tag{B6}$$

where $+$ denotes the max-normalized entries given by $b_{jk}^{+} = b_{jk}/\max_j|b_{jk}|$. The power parameter p thus defines the strength of the promax operation, while the sign remains unchanged. Using Eq. (B5), the denormalized promax-rotated EOFs are given by

$$\tilde{\mathbf{L}}_P = \tilde{\mathbf{L}}_V\mathbf{R}\mathbf{D} = \tilde{\mathbf{H}}^{-1}\tilde{\mathbf{L}}_V[(\tilde{\mathbf{L}}_V^{*}\tilde{\mathbf{H}}^{-2}\tilde{\mathbf{L}}_V)^{-1}\tilde{\mathbf{L}}_V^{*}\tilde{\mathbf{H}}^{-1}\mathbf{T}]\mathbf{D}, \tag{B7}$$

where the normalization matrix is given by $\mathbf{D}^2 = \mathrm{diag}(\mathbf{R}^{*}\mathbf{R})^{-1}$.

REFERENCES

Alexander, M., and J. Scott, 2002: The influence of ENSO on air–sea interaction in the Atlantic. *Geophys. Res. Lett.*, **29**, 46-1–46-4, https://doi.org/10.1029/2001GL014347.

Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007, https://doi.org/10.1029/2006JC003798.

Assimakopoulos, V., and K. Nikolopoulos, 2000: The theta model: A decomposition approach to forecasting. *Int. J. Forecasting*, **16**, 521–530, https://doi.org/10.1016/S0169-2070(00)00066-2.

Ballabrera-Poy, J., R. Murtugudde, and A. J. Busalacchi, 2002: On the potential impact of sea surface salinity observations on ENSO predictions. *J. Geophys. Res.*, **107**, 8007, https://doi.org/10.1029/2001JC000834.

Bell, B., and Coauthors, 2020: ERA5 monthly averaged data on single levels from 1950 to 1978 (preliminary version). Copernicus Climate Data Store, accessed 20 January 2021, https://doi.org/10.24381/cds.f17050d7.

Bloomfield, P., and J. M. Davis, 1994: Orthogonal rotation of complex principal components. *Int. J. Climatol.*, **14**, 759–775, https://doi.org/10.1002/joc.3370140706.

Boashash, B., 1992: Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. *Proc. IEEE*, **80**, 520–538, https://doi.org/10.1109/5.135376.

Bretherton, C., C. Smith, and J. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate*, **5**, 541–560, https://doi.org/10.1175/1520-0442(1992)005<0541:AIOMFF>2.0.CO;2.

Cheng, X., and T. J. Dunkerton, 1995: Orthogonal rotation of spatial patterns derived from singular value decomposition analysis. *J. Climate*, **8**, 2631–2643, https://doi.org/10.1175/1520-0442(1995)008<2631:OROSPD>2.0.CO;2.

——, G. Nitsche, and J. M. Wallace, 1995: Robustness of low-frequency circulation patterns derived from EOF and rotated EOF analyses. *J. Climate*, **8**, 1709–1713, https://doi.org/10.1175/1520-0442(1995)008<1709:ROLFCP>2.0.CO;2.

Cherchi, A., and A. Navarra, 2013: Influence of ENSO and of the Indian Ocean Dipole on the Indian summer monsoon variability. *Climate Dyn.*, **41**, 81–103, https://doi.org/10.1007/s00382-012-1602-y.

Chiang, J. C. H., and A. H. Sobel, 2002: Tropical tropospheric temperature variations caused by ENSO and their influence on the remote tropical climate. *J. Climate*, **15**, 2616–2631, https://doi.org/10.1175/1520-0442(2002)015<2616:TTTVCB>2.0.CO;2.

——, and D. J. Vimont, 2004: Analogous Pacific and Atlantic meridional modes of tropical atmosphere–ocean variability. *J. Climate*, **17**, 4143–4158, https://doi.org/10.1175/JCLI4953.1.

Cruz-Cano, R., and M.-L. T. Lee, 2014: Fast regularized canonical correlation analysis. *Comput. Stat. Data Anal.*, **70**, 88–100, https://doi.org/10.1016/j.csda.2013.09.020.

Dai, A., and T. M. L. Wigley, 2000: Global patterns of ENSO-induced precipitation. *Geophys. Res. Lett.*, **27**, 1283–1286, https://doi.org/10.1029/1999GL011140.

Dee, D., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Enfield, D. B., and D. A. Mayer, 1997: Tropical Atlantic sea surface temperature variability and its relation to El Niño–Southern Oscillation. *J. Geophys. Res.*, **102**, 929–945, https://doi.org/10.1029/96JC03296.

Feng, J., W. Chen, C.-Y. Tam, and W. Zhou, 2011: Different impacts of El Niño and El Niño Modoki on China rainfall in the decaying phases. *Int. J. Climatol.*, **31**, 2091–2101, https://doi.org/10.1002/joc.2217.

Finch, H., 2006: Comparison of the performance of Varimax and Promax rotations: Factor structure recovery for dichotomous items. *J. Educ. Meas.*, **43**, 39–52, https://doi.org/10.1111/j.1745-3984.2006.00003.x.

Fiorucci, J. A., T. R. Pellegrini, F. Louzada, F. Petropoulos, and A. B. Koehler, 2016: Models for optimising the theta method and their relationship to state space models. *Int. J. Forecasting*, **32**, 1151–1161, https://doi.org/10.1016/j.ijforecast.2016.02.005.

Gaughan, A. E., C. G. Staub, A. Hoell, A. Weaver, and P. R. Waylen, 2016: Inter- and intra-annual precipitation variability and associated relationships to ENSO and the IOD in southern Africa. *Int. J. Climatol.*, **36**, 1643–1656, https://doi.org/10.1002/joc.4448.

Giorgi, F., F. Raffaele, and E. Coppola, 2019: The response of precipitation characteristics to global warming from climate

projections. *Earth Syst. Dyn.*, **10**, 73–89, https://doi.org/10.5194/esd-10-73-2019.

Gu, G., and R. F. Adler, 2015: Spatial patterns of global precipitation change and variability during 1901–2010. *J. Climate*, **28**, 4431–4453, https://doi.org/10.1175/JCLI-D-14-00201.1.

Hannachi, A., 2016: Regularised empirical orthogonal functions. *Tellus*, **68A**, 31723, https://doi.org/10.3402/tellusa.v68.31723.

——, 2021: *Patterns Identification and Data Mining in Weather and Climate*. Springer, 624 pp.

——, and N. Trendafilov, 2017: Archetypal analysis: Mining weather and climate extremes. *J. Climate*, **30**, 6927–6944, https://doi.org/10.1175/JCLI-D-16-0798.1.

——, I. Jolliffe, and D. Stephenson, 2007: Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.*, **27**, 1119–1152, https://doi.org/10.1002/joc.1499.

Hendrickson, A. E., and P. O. White, 1964: Promax: A quick method for rotation to oblique simple structure. *Br. J. Stat. Psychol.*, **17**, 65–70, https://doi.org/10.1111/j.2044-8317.1964.tb00244.x.

Hersbach, H., and Coauthors, 2019: ERA5 monthly averaged data on single levels from 1979 to present. Copernicus Climate Data Store, accessed 20 January 2021, https://doi.org/10.24381/cds.f17050d7.

Horel, J., 1984: Complex principal component analysis: Theory and examples. *J. Climate Appl. Meteor.*, **23**, 1660–1673, https://doi.org/10.1175/1520-0450(1984)023<1660:CPCATA>2.0.CO;2.

Hotelling, H., 1936: Relations between two sets of variates. *Biometrika*, **28**, 321–377, https://doi.org/10.2307/2333955.

Hoyer, S., and J. Hamman, 2017: xarray: N-D labeled arrays and datasets in Python. *J. Open Res. Software*, **5**, 10, https://doi.org/10.5334/jors.148.

Hu, Z.-Z., and B. Huang, 2009: Interferential impact of ENSO and PDO on dry and wet conditions in the U.S. Great Plains. *J. Climate*, **22**, 6047–6065, https://doi.org/10.1175/2009JCLI2798.1.

Hurley, J. R., and R. B. Cattell, 1962: The Procrustes program: Producing direct rotation to test a hypothesized factor structure. *Behav. Sci.*, **7**, 258–262, https://doi.org/10.1002/bs.3830070216.

Hyndman, R. J., and B. Billah, 2003: Unmasking the theta method. *Int. J. Forecasting*, **19**, 287–290, https://doi.org/10.1016/S0169-2070(01)00143-1.

Indeje, M., F. H. Semazzi, and L. J. Ogallo, 2000: ENSO signals in East African rainfall seasons. *Int. J. Climatol.*, **20**, 19–46, https://doi.org/10.1002/(SICI)1097-0088(200001)20:1<19::AID-JOC449>3.0.CO;2-0.

Kaiser, H. F., 1958: The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187–200, https://doi.org/10.1007/BF02289233.

Klein, S. A., B. J. Soden, and N.-C. Lau, 1999: Remote sea surface temperature variations during ENSO: Evidence for a tropical atmospheric bridge. *J. Climate*, **12**, 917–932, https://doi.org/10.1175/1520-0442(1999)012<0917:RSSTVD>2.0.CO;2.

Krishnamurthy, V., and B. P. Kirtman, 2003: Variability of the Indian Ocean: Relation to monsoon and ENSO. *Quart. J. Roy. Meteor. Soc.*, **129**, 1623–1646, https://doi.org/10.1256/qj.01.166.

Kutzbach, J. E., 1967: Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. *J. Appl. Meteor. Climatol.*, **6**, 791–802, https://doi.org/10.1175/1520-0450(1967)006<0791:EEOSLP>2.0.CO;2.

Lamb, P. J., R. A. Peppler, and S. Hastenrath, 1986: Interannual variability in the tropical Atlantic. *Nature*, **322**, 238–240, https://doi.org/10.1038/322238a0.

Lau, N.-C., and M. J. Nath, 1996: The role of the ''atmospheric bridge'' in linking tropical Pacific ENSO events to extratropical SST anomalies. *J. Climate*, **9**, 2036–2057, https://doi.org/10.1175/1520-0442(1996)009<2036:TROTBI>2.0.CO;2.

Lenssen, N. J. L., L. Goddard, and S. Mason, 2020: Seasonal forecast skill of ENSO teleconnection maps. *Wea. Forecasting*, **35**, 2387–2406, https://doi.org/10.1175/WAF-D-19-0235.1.

Li, L., R. Schmitt, C. Ummenhofer, and K. Karnauskas, 2016: North Atlantic salinity as a predictor of Sahel rainfall. *Sci. Adv.*, **2**, e1501588, https://doi.org/10.1126/sciadv.1501588.

Lian, T., and D. Chen, 2012: An evaluation of rotated EOF analysis and its application to tropical Pacific SST variability. *J. Climate*, **25**, 5361–5373, https://doi.org/10.1175/JCLI-D-11-00663.1.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708, https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

Mantua, N. J., and S. R. Hare, 2002: The Pacific decadal oscillation. *J. Oceanogr.*, **58**, 35–44, https://doi.org/10.1023/A:1015820616384.

——, ——, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1080, https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2.

Martin, E. R., C. Thorncroft, and B. B. B. Booth, 2014: The multidecadal Atlantic SST—Sahel rainfall teleconnection in CMIP5 simulations. *J. Climate*, **27**, 784–806, https://doi.org/10.1175/JCLI-D-13-00242.1.

Newman, M., and Coauthors, 2016: The Pacific decadal oscillation, revisited. *J. Climate*, **29**, 4399–4427, https://doi.org/10.1175/JCLI-D-15-0508.1.

North, G., T. L. Bell, R. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110**, 699–706, https://doi.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2.

Pierce, D. W., 2002: The role of sea surface temperatures in interactions between ENSO and the North Pacific Oscillation. *J. Climate*, **15**, 1295–1308, https://doi.org/10.1175/1520-0442(2002)015<1295:TROSST>2.0.CO;2.

Rasmusson, E. M., P. A. Arkin, W.-Y. Chen, and J. B. Jalickee, 1981: Biennial variations in surface temperature over the United States as revealed by singular decomposition. *Mon. Wea. Rev.*, **109**, 587–598, https://doi.org/10.1175/1520-0493(1981)109<0587:BVISTO>2.0.CO;2.

Ratnam, J. V., S. K. Behera, Y. Masumoto, and T. Yamagata, 2014: Remote effects of El Niño and Modoki events on the austral summer precipitation of southern Africa. *J. Climate*, **27**, 3802–3815, https://doi.org/10.1175/JCLI-D-13-00431.1.

Richman, M. B., 1986: Rotation of principal components. *J. Climatol.*, **6**, 293–335, https://doi.org/10.1002/joc.3370060305.

Rieger, N., 2021: xmca v0.3.3: Maximum covariance analysis for climate science. Zenodo, https://doi.org/10.5281/zenodo.4751735.

Ropelewski, C. F., and M. S. Halpert, 1986: North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Wea. Rev.*, **114**, 2352–2362, https://doi.org/10.1175/1520-0493(1986)114<2352:NAPATP>2.0.CO;2.

Saravanan, R., and P. Chang, 2000: Interaction between tropical Atlantic variability and El Niño–Southern Oscillation. *J. Climate*, **13**, 2177–2194, https://doi.org/10.1175/1520-0442(2000)013<2177:IBTAVA>2.0.CO;2.

Servain, J., I. Wainer, J. P. McCreary, and A. Dessier, 1999: Relationship between the equatorial and meridional modes of climatic variability in the tropical Atlantic. *Geophys. Res. Lett.*, **26**, 485–488, https://doi.org/10.1029/1999GL900014.

Simmons, A., and Coauthors, 2021: Low frequency variability and trends in surface air temperature and humidity from ERA5 and other datasets. ECMWF Tech. Memo. 881, ECMWF, 97 pp., https://doi.org/10.21957/ly5vbtbfd.

Taschetto, A. S., and M. H. England, 2009: El Niño Modoki impacts on Australian rainfall. *J. Climate*, **22**, 3167–3174, https://doi.org/10.1175/2008JCLI2589.1.

Tedeschi, R. G., I. F. Cavalcanti, and A. M. Grimm, 2013: Influences of two types of ENSO on South American precipitation. *Int. J. Climatol.*, **33**, 1382–1400, https://doi.org/10.1002/joc.3519.

Thurstone, L., 1947: The simple structure concept. *Multiple Factor Analysis: A Development and Expansion of The Vectors of Mind*. University of Chicago Press, 319–346.

van der Walt, S., S. C. Colbert, and G. Varoquaux, 2011: The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.*, **13**, 22–30, https://doi.org/10.1109/MCSE.2011.37.

Vinod, H. D., 1976: Canonical ridge and econometrics of joint production. *J. Econom.*, **4**, 147–166, https://doi.org/10.1016/0304-4076(76)90010-5.

Vittal, H., G. Villarini, and W. Zhang, 2020: Early prediction of the Indian summer monsoon rainfall by the Atlantic meridional mode. *Climate Dyn.*, **54**, 2337–2346, https://doi.org/10.1007/s00382-019-05117-0.

Wallace, J. M., and R. E. Dickinson, 1972: Empirical orthogonal representation of time series in the frequency domain. Part I: Theoretical considerations. *J. Appl. Meteor.*, **11**, 887–892, https://doi.org/10.1175/1520-0450(1972)011<0887:EOROTS>2.0.CO;2.

Wen, N., Z. Liu, and L. Li, 2019: Direct ENSO impact on East Asian summer precipitation in the developing summer. *Climate Dyn.*, **52**, 6799–6815, https://doi.org/10.1007/s00382-018-4545-0.

Wiedermann, M., J. F. Siegmund, J. F. Donges, and R. V. Donner, 2021: Differential imprints of distinct ENSO flavors in global patterns of very low and high seasonal precipitation. *Front. Climate*, **3**, 618548, https://doi.org/10.3389/fclim.2021.618548.

Xie, S.-P., C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg, 2010: Global warming pattern formation: Sea surface temperature and rainfall. *J. Climate*, **23**, 966–986, https://doi.org/10.1175/2009JCLI3329.1.

Zhang, C., 2005: Madden-Julian Oscillation. *Rev. Geophys.*, **43**, RG2003, https://doi.org/10.1029/2004RG000158.