A Feasibility Study of Three-Dimensional Empirical Orthogonal Functions From the NASA JPL Ocean General Circulation Model: Computing, Visualization and Interpretation

DANIELLE LAFARGA (iD)

THOMAS BUI

Y. TONY SONG** (iD)

THOMAS M. SMITH** (iD)

SAMUEL S. P. SHEN (iD)

*Author affiliations can be found in the back matter of this article

**These authors contributed equally to this work

ABSTRACT

Existing oceanic studies on either data reconstruction or dynamics often used 2-dimensional empirical orthogonal functions (EOF) for sea surface temperature (SST) and for deep layers. However, large-scale oceanic dynamics, such as equatorial ocean upwelling and arctic ocean ventilation, imply the existence of strong covariance among the temperatures and other parameters between different layers. These ocean dynamics are not best represented in the isolated 2-dimensional layer-by-layer calculations, while the 3-dimensional EOFs have a clear advantage. The purpose of this paper is to demonstrate 3D EOF calculations based on the NASA Jet Propulsion Laboratory (JPL) ocean general circulation model (OGCM) from surface to 5,500 meters depth, with 33 depth layers, 1-degree latitude and longitude spatial resolution, and monthly temporal resolution. We also present visualizations of the 3D EOFs and make physical interpretations of the first two EOFs. Our 3D EOF results demonstrate that (i) the 3D spatial pattern of equatorial ocean upwelling is mainly reflected in the first EOF mode and has its most variabilities within the depth layer between 100 and 400 meters, (ii) the 3D El Niño Southern Oscillation (ENSO) dynamic pattern is mainly reflected in the second EOF mode and is mostly confined from surface to the depth of 150 meters, and (iii) the lead eigenvalue from the 3D EOF calculation appears to contain some signal of oceanic warming. Additionally, our method of weighted 3D EOF computation and our 3D visualization Python code may be useful tools for both climate professionals and students.

1 INTRODUCTION

Empirical orthogonal functions (EOFs), or spatial modes, are often used to analyze general climate model (GCM) output to explore spatial patterns of climate variability or for reconstructing historical climate data (Zhang and Moore (2015), Houser et al (2020)). EOFs show which spatial patterns explain the most variance in the data, and can consequently give insight into how our data behaves. Although GCMs and climate dynamics are 3-dimensional in space and 1-dimensional in time, EOF analyses have often been performed in 2D layers in space (Shen et al. (2017), Meyers et al. (1999), Mu et al. (2018)). The primary purposes of our paper are twofold: (a) to show a computational algorithm of 3D EOFs, and (b) to interpret the climate dynamics of the first two EOFs. We do so by using the output data of a NASA JPL ocean general circulation model (OGCM).

Numerically, EOFs can be computed by finding the eigenvectors of the spatial covariance matrix from the weighted anomalies. This spatial covariance method often has convenient physical interpretations for leading modes. Most EOF calculations in recent years use singular value decomposition (SVD) with packages available in different computer languages, e.g., Python, R, and MatLab, when the size of the SVD matrix is not too big, say less than half of your computer memory (Shen and Somerville (2019)). Our 1-degree spatial resolution dataset is not very big and SVD computing can be done. However, we also use temporal covariance in this paper which can address the issue of memory limitations for future datasets of higher resolutions. This memory limitation is perhaps one of many reasons that EOFs were used to be computed layer-by-layer. Our method details in this paper are presented in the way of covariance.

Papers such as Shen et al. (2017), Meyers et al. (1999), and Mu et al. (2018) calculated 2D EOFs, and indicated that the strongest signal was El Niño Southern Oscillation (ENSO). ENSO is an ocean-atmosphere climate phenomenon characterized by the warming of sea surface temperatures in the central and/or east-central equatorial Pacific (NOAA CPC (2005)). The pattern is often described as a warm tongue in the eastern pacific. Despite its aforementioned clear spatial pattern in sea surface temperature (SST), ENSO is a phenomenon involving not only the surface but also the deep ocean, up to 150 meters, to be discussed later in this paper. Therefore, 2D EOFs lack the ability to describe the interactions between surface and deeper ocean layers. The need to consider such significant deep ocean dynamics motivated the idea of computing EOFs for all layers at once. The resulting 3D EOFs show that ENSO accounts for a smaller fraction of variability than equatorial upwelling. This validates the significance of calculating 3D EOFs.

Our paper will demonstrate 3D EOF calculations based on the NASA JPL ocean general circulation model (OGCM)

from surface to 5,500 meters depth, with 32 depth layers, 1-degree latitude and longitude spatial resolution, and monthly temporal resolution. Our mode visualization and physical interpretation will focus on the first two EOFs. Our 3D EOF results demonstrate that (i) the 3D spatial pattern of equatorial ocean upwelling is mainly reflected in the first EOF mode and has its most variabilities within the depth layer between 100 and 400 meters, (ii) the 3D El Niño Southern Oscillation (ENSO) dynamic pattern is mainly reflected in the second EOF mode and is mostly confined from surface to the depth of 150 meters, and (iii) the lead eigenvalue from the 3D EOF calculation appears to contain some signal of oceanic warming.

Relevant literature already exists on (i) multiway table data analysis theory, e.g, Structuration des Tableaux A Trois Indices de la Statistique (STATIS) (Acar and Yener (2009), Vallejo-Arboleda et al. (2007), Sabatier and Traissac (1994)), and (ii) computational algorithms for big data EOFs (Gittens et al. (2016)). A special case of multiway table is a 3-way table that includes locations, parameters, and time (Stanimirova et al. (2004)). In our case, we only look at only one parameter: temperature; our locations are three-dimensional; and our time is indexed for monthly data from the NASA JPL OGCM.

In principle, multiway data analysis algorithms such as STATIS should be able to deliver the results of PCs and EOFs when the multiway data table is properly non-dimensionalized. In practice, the problem is much more complex, such as the large size of the multiway data table, and the physical meanings of the PCs and EOFs. For example Gittens et al. (2016) contributed to the development of innovative computational algorithms to handle multi-terabytes of data for 3D EOF computing. Their paper considered the cases: (a) Climate Forecast System Reanalysis (CFSR) data with a 1/2-degree, 41 layers, and 6-hour resolution; (b) Community Atmosphere Model 5.0 (CAM5) data with 1/4-degree, 30 layers, and 6-hour resolution. Their CFSR dataset has 2.2 TB and their CAM5 dataset has 16 TB. They concluded that most of the run time was spent on In/Out process. Their EOFs from their multi-hour data by default have no clear physical interpretations. Accordingly, their paper included neither EOF figures nor physics interpretations. Our paper shows both the successful application of our 3D EOF algorithm that utilizes a temporal covariance matrix designed for future applications on multi-terabytes of data, and physics interpretations of the 3D EOFs from the NASA JPL OGCM. Therefore, our study has a modest purpose to show the feasibility of (i) computing 3D EOFs from a climate model output, and (ii) making physics interpretations from the point of view of ocean dynamics.

The remainder of this paper is organized as follows: Section 2 describes the data set, Section 3 describes the proposed method, section 4 includes results of performing the method on OGCM data, and section 5 contains the conclusion. Our method of weighted 3D EOF

computation and our 3D graphics Python code may be useful tools for both climate professionals and students. The comprehensive code can be accessed at https://github.com/dlafarga/calc_3D_EOFs.

2 DATA

2.1 TEMPERATURE DATA FROM THE NASA JPL OGCM

We use the monthly temperature data from the NASA JPL non-Boussinesq ocean general circulation model (OGCM) from January 1950 to December 2003 (Song and Hou (2006), Song et al. (2010)). The horizontal resolution of the JPL OGCM is –degree latitude-longitude and was later aggregated to 1-degree as in Shen et al. (2017). After aggregating the data, the files were separated into MatLab files according to their month and year. Figure 1 shows one file of 3D ocean temperature data for January 1950.

The data encompasses the entire global ocean with land represented by a nonnumber (NaN) grid cell and temperatures as a °C grid cell. There are a total of 32 depth layers whose depth boundaries are given in Table 1. As an example take the first boundary from Table 1 as 5. This means the top layer extends from 0 to 5 meters. Similarly, the second layer is from 5 to 10 meters, etc.

This paper uses the JPL non-Boussinesq (mass conserving) OGCM as it well portrays ocean dynamics. The model was based on both the S-Coordinate Rutgers University Model (Song and Haidvogel (1994)), and the Regional Ocean Model System (Shchepetkin and McWilliams (2003), Haidvogel et al. (2000)). The JPL non-Boussineq OGCM served as a resolution to the uncertainties in altimetry sea-surface-height (SSH) and ocean bottom pressure (OBP) in Boussineq approximations (Song and Hou (2006), Song et al. (2010)) identified by Huang and Jin (2002). Both SSH and OBP's consistency within the model have been validated by Moon et al. (2013) and Song and Colberg (2011). Because the OGCM output is dependable in its representation of ocean dynamics, we find it appropriate for our EOFs.
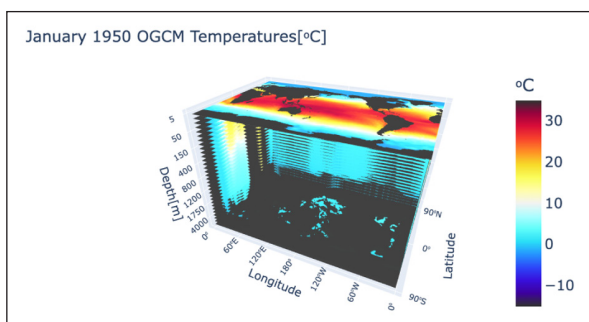

Figure 1 Ocean water temperature from the NASA JPL OGCM output for January 1950. A movable version of the figure can be found at https://ogcm-3d-visualization.herokuapp.com/ where the figure can be rotated and resized.

5	10	20	30	50	75	100
125	150	200	250	300	400	500
600	700	800	900	1000	1100	1200
1300	1400	1500	1750	2000	2500	3000
3500	4000	5000	5500			

Table 1 Depths of 32 layers of the NASA JPL OGCM.

Note: All depths are in meters.

2.2 FORMULATION OF THE SPACE-TIME DATA MATRIX

The OGCM output is an array written in 3D space grid boxes separated monthly for each year. Our time spans from January 1950 to December 2003. If we take January as an example it will have 54 years from 1950 to 2003; the same can be said for February, March, ..., December. Because of seasonal variation, we compute EOFs for each month from January, February, ..., December therefore we will compute 12 sets of EOFs. Let us first consider the data on a single layer with a given depth and time. We use January in the following description.

Denote the data matrix of one year and layer by $\mathbf{T}_{year,layer}$ and the final space-time data matrix by $\mathbf{X}$. The gridded temperature data ($T_{longitude,latitude}$) for the first year (Y_{1950}) and first layer (l_1) form a 360×180 matrix shown below:

$$\mathbf{T}_{Y_{1950},l_1} = \begin{pmatrix} T_{0,-89.5} & T_{0,-88.5} & \cdots & T_{0,89.5} \\ T_{1,-89.5} & T_{1,-88.5} & \cdots & T_{1,89.5} \\ T_{2,-89.5} & T_{2,-88.5} & \cdots & T_{2,89.5} \\ \cdots & \cdots & \cdots & \cdots \\ T_{359,-89.5} & T_{359,-88.5} & \cdots & T_{359,89.5} \end{pmatrix}_{360 \times 180} \quad (1)$$

The row data of the matrix are ordered from south to north starting at latitude –89.5° and ending at latitude 89.5° totaling 180 columns. The column data are ordered according to longitude from east to west starting at longitude 0° ending at longitude 359° totaling 360 rows. Namely, the first 180 points in the first row of the matrix represent 89.5°*S* to 89.5°*N* for a longitude 0° as shown by the first row of the matrix. The next row will show the same latitudes, 89.5°*S* to 89.5°*N*, for the next longitude 1°. This continues until longitude reaches 359°, the last row of the data matrix. Flattening this matrix by rows gives us a row vector of length $360 \times 180 = 64,800$.

$$\vec{T}_{Y_{1950},l_1} = (T_{0,-89.5} \quad T_{0,-88.5} \quad \cdots \quad T_{359,89.5})_{1 \times 64800} \quad (2)$$

We transpose this row vector $\vec{T}_{Y_{1950},l_1}$ to a column vector $\vec{T}_{l_1,Y_{1950}}$ and aggregate all the column vectors from 1950 to 2003 to form a space-time data matrix for the first layer of January. The 2D space-time data matrix for a single layer, denoted by $\mathbf{X}_{Jan_{2D}}$, looks like the following:

$$X_{Jan_{2D}} = \left(\vec{T}_{l_1,Y_{1950}} \quad \vec{T}_{l_1,Y_{1951}} \quad \vec{T}_{l_1,Y_{1952}} \quad \cdots \quad \vec{T}_{l_1,Y_{2003}}\right)_{64800\times54} \tag{3}$$

Anomalies can be computed from this matrix and multiplied by the area weight of the grid box *i* depending on its latitude in radians ϕ_i. These area weights reflect the geometric sizes of their corresponding grid boxes.

$$area_i = \sqrt{\cos\phi_i} \tag{4}$$

The result of the multiplication is the area-weighted space-time anomaly matrix. Conveniently the structure of this matrix allows for the removal of land/NaN data by removing rows as every year contains the same grid point of land data. For example, if January of 1950 had a land point at $T_{0,-89.5}$, then January 1951–2003 will have the same land point at $T_{0,-89.5}$. Only after the removal of land data can the area-weighted space-time anomaly matrix be fed into SVD, or any other method of computing EOFs, and produce the 2D EOFs layer by layer.

We can build on the idea of computing 2D EOFs to compute 3D EOFs by stacking 32 layers on top of each other as shown below.

$$X_{Jan_{3D}} = \begin{array}{c} \\ l_1 \\ l_2 \\ l_3 \\ \cdots \\ l_{32} \end{array} \begin{pmatrix} \overset{Y_{1950}}{\vec{T}_{l_1,Y_{1950}}} & \overset{Y_{1951}}{\vec{T}_{l_1,Y_{1951}}} & \overset{Y_{1952}}{\vec{T}_{l_1,Y_{1952}}} & \overset{\cdots}{\cdots} & \overset{Y_{2003}}{\vec{T}_{l_1,Y_{2003}}} \\ \vec{T}_{l_2,Y_{1950}} & \vec{T}_{l_2,Y_{1951}} & \vec{T}_{l_2,Y_{1952}} & \cdots & \vec{T}_{l_2,Y_{2003}} \\ \vec{T}_{l_3,Y_{1950}} & \vec{T}_{l_3,Y_{1951}} & \vec{T}_{l_3,Y_{1952}} & \cdots & \vec{T}_{l_3,Y_{2003}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \vec{T}_{l_{32},Y_{1950}} & \vec{T}_{l_{32},Y_{1951}} & \vec{T}_{l_{32},Y_{1952}} & \cdots & \vec{T}_{l_{32},Y_{2003}} \end{pmatrix}_{2,073,600\times54} \tag{5}$$

Here,

$$2,073,600 = 64,800 \times 32 \tag{6}$$

is the total number of 3D grid boxes for the OGCM output. Following the same way, we formulate our February space-time data matrix, and so on, since our EOFs are computed for each month from January to December. This space-time data matrix for January is one of the 12 matrices used to compute 3D EOFs as outlined in the next section (see equation 7).

3 METHODOLOGY

There are two methods described in this section: SVD and temporal covariance. We will first describe the SVD method to better understand the relationship between each method. We start by calculating our climatology. We use the entire data history from 1950 to 2003 as our climatology period. Let us use the following simple space-time data matrix notation $[T_{it}]$ to explain our calculation of climatology and anomalies:

$$X = [T_{it}]_{N\times Y} \tag{7}$$

where *i* is a grid box simplified from the previous latitude-longitude notation, t is time, *N* is the total number of grid boxes, and *Y* is the total number of years. Specifically, *N* = 2,073,600, and *Y* = 54. This *X* can be the space-time data matrix for a particular month, January, February, or another month.

The climatology is computed as the time means of the data matrix, i.e., the mean for each row:

$$\overline{T}_i = \frac{1}{Y}\sum_{t=1}^{Y} T_{it} \tag{8}$$

We compute the climatology for each month: January, February, ..., December. Thus, 12 sets of climatology are computed. Figure 2 shows the climatology of January and that of August.

Our anomalies are computed as follows:

$$A = [a_{it}]_{N\times Y} = [T_{it} - \overline{T}_i]_{N\times Y} \tag{9}$$

These anomalies are multiplied by their spatial weights to, as mentioned previously, better reflect the geometric sizes of their corresponding grid boxes. A 3D grid box in the equator region is much larger than one in the polar region. In 2D the weight is related to the area of a grid box (see equation 4). Naturally, in our 3D case the weight for a 3D grid box *i* should be related to the volume of the grid box:
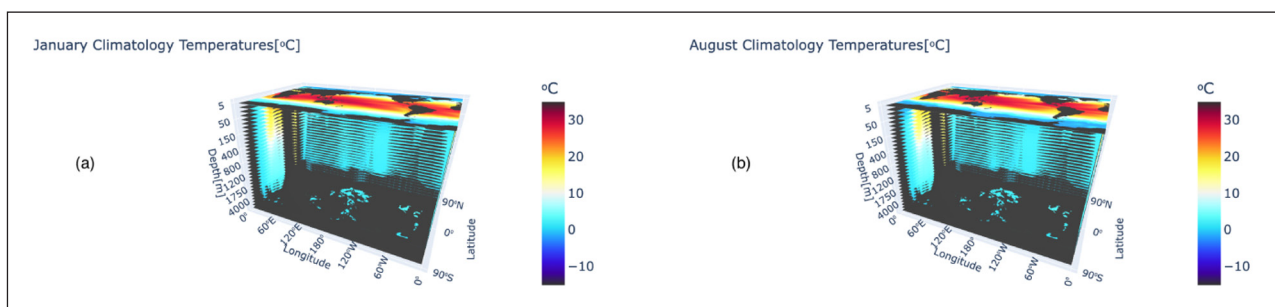
$$w_i = \sqrt{\Delta d_i \cos\phi_i} \tag{10}$$

Figure 2 Climatology for January and August computed using the equation above from the NASA JPL OGCM data. Movable figures can be viewed at the website https://climatology-3d-vis.herokuapp.com/apps/clim_Jan_plot.

where Δd_i is the thickness of the layer for the grid box i and ϕ_i is the latitude of the grid box i in radians. The square root is needed due to the symmetry of the formulation of the eigenvalue problem of a covariance matrix (Shen et al.(2017)).

Often, anomalies are standardized by dividing by the standard deviation. However, in the deeper ocean layers, temperature has a very small standard deviation. Dividing by the very small standard deviation would only amplify noise and cause large errors for the EOFs in the top layers. To avoid such errors, we simply use non-standardized volume-weighted anomalies. Up to now, the NaN missing data are still in the anomaly data matrix. We remove the rows of NaN and produce a volume-weighted space-time anomaly matrix without missing data. We denote this weighted anomaly matrix without missing data by

$$B = \left[b_{jt} \right]_{N' \times Y} \qquad (11)$$

where N' is the number of grid boxes with data. Then, we can feed this space-time data matrix to SVD to generate EOFs to demonstrate spatial patterns and principal components (PCs) to show temporal patterns. The SVD decomposition of B is as follows:

$$B = UDV^T \qquad (12)$$

where the superscript T indicates the transpose of a matrix. In this SVD formula, U is an $N' \times Y$ ($N' > Y$) spatial matrix. Each column vector of U is an EOF. Matrix V has dimension $Y \times Y$ and is the temporal matrix. Each column is a PC. Matrix D is a diagonal matrix, whose elements $d_k (k = 1, 2, ..., Y)$ correspond to the standard deviations of the EOFs and PCs. The next sections will explain these EOF and PC patterns in U and V, as well as the standard deviations in D.

Although SVD is a convenient way to compute EOFs and PCs, the covariance matrix method is often helpful in interpreting the meaning of EOFs and PCs from the point of view of variance and covariance in climate dynamics. Further, this method was traditionally used in the EOF calculations before using SVD in recent years. The traditional method often uses the eigenvalue problem for a spatial covariance matrix to compute EOFs. However, in our case, the temporal dimension Y is much smaller than the spatial dimension N', hence the temporal covariance matrix is much smaller than the spatial covariance matrix. It is preferred to solve the eigenvalue problem for the temporal covariance matrix which is much smaller than the spatial covariance matrix. The resulting eigenvectors of the temporal covariance matrix are PCs: $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_Y$. The EOFs can be obtained by the projection of the weighted anomaly matrix B to the PCs:

$$\vec{u}_k = \frac{B\vec{v}_k}{|B\vec{v}_k|}, \ k = 1, 2, \cdots, Y \qquad (13)$$

where $|B\vec{v}_k|$ is the Euclidean norm of the vector $B\vec{v}_k$.

The temporal covariance matrix, denoted as $[\Sigma_{tt'}]_{Y \times Y}$, is defined as follows:

$$[\Sigma_{tt'}]_{Y \times Y} = \left[\frac{1}{N'} \sum_{i=1}^{N'} b_{it} b_{it'} \right]_{Y \times Y} = \frac{1}{N'} B_{Y \times N'}^T B_{N' \times Y} \qquad (14)$$

An eigenvector $\vec{v}$ of this temporal matrix $[\Sigma_{tt'}]$ is defined as a vector whose orientation is not changed after the action of $[\Sigma_{tt'}]$. Namely, the matrix action vector $[\Sigma_{tt'}]\vec{v}$ is parallel to the original vector. Thus, there exists a scalar, denoted by λ so that one vector is the scalar multiplication of the other, i.e.,

$$[\Sigma_{tt'}]\vec{v} = \lambda \vec{v}. \qquad (15)$$

This eigenvalue problem has Y solutions whose Y eigenvectors ($\vec{v}_k, k = 1, 2, \cdots, Y$) are the same PCs computed from SVD (see the rigorous mathematical proof in Shen and Somerville (2019)). The eigenvalues ($\lambda_k, k = 1, 2, ..., Y$) are proportional to the $d_k^2, k = 1, 2, \cdots, Y$ computed from SVD. Therefore, $\lambda_k, k = 1, 2, ..., Y$ correspond to variances since $d_k, k = 1, 2, ..., Y$ correspond to standard deviation.

Eigenvalues represent the amount of variance explained by each respective eigenvector (National Center for Atmospheric Research Staff (Eds) (2013)) and are sorted from the greatest variance to the smallest. Variance is a measure of how far a set of numbers is spread out from their average value (Shen and Somerville (2019)). Thus, the first eigenmode will show us the most significant pattern, as it has the most variance in climate dynamics. The last modes are usually less accurate and less important as they contain much noise and small-scale variations (Liang et al. (2012)). The details of the EOF and PC derivations can be found in Shen et al. (2017).

We have computed the EOFs, PCs, standard deviations, and variances for each mode using both the methods of SVD and temporal covariance. We have compared their results as a cross-check and explored the physical interpretation of their meaning. These results are described in the next section.

4 RESULTS

4.1 EIGENVALUES AND VARIANCES EXPLAINED

EOFs, PCs, SVD eigenvalues, and eigenvalues of the temporal covariance matrix were computed for each month from January to December. Altogether 12 sets of EOFs were generated. Figure 3(a) shows the first eigenvalue of the temporal covariance matrix for the weighted temperature anomalies for each month from January to December. These eigenvalues represent variances of the space-time data projection onto their corresponding EOF1 for each month. The peaks in January

and August imply large variances in both months, relative to their neighboring months. Our examination of all the EOF patterns has shown that those two peaks are likely the manifestation of the relatively stronger equatorial upwelling in January and August.

Another important feature of Figure 3(a) is the obvious difference between January and December. The difference is around 3% and is likely an indication of the presence of a temporal trend in the space-time temperature data. If there were no temporal trend the annual cycle would prevail showing a 12-month cycle instead, and hence December's variance would be similar in size to January's. Visually this would be observed as a smoother transition between December to January. Later on, we further quantitatively verify our assumption here by solving the eigenvalue problem for the de-trended data. Indeed, the annual cycle of λ_1 appears.

Figure 3(b) and (c) show the scree plot for both peak months January and August. The scree plot plots every eigenvalue for their respective month as a percentage variance or cumulative variance. The purpose is to quantify the variance explained by each mode and to identify how many modes are needed to explain a significant amount of variance. The blue dotted line explains the percentage variance of each EOF mode relative to the total variance, i.e.,

$$p_k = \frac{\lambda_k}{\sum_{\ell=1}^{Y} \lambda_\ell} \times 100. \tag{16}$$

The red dotted line is the cumulative percentage variance relative to the total variance, defined as follows

$$q_k = \frac{\sum_{\ell=1}^{k} \lambda_\ell}{\sum_{\ell=1}^{Y} \lambda_\ell} \times 100. \tag{17}$$

Comparing Figure 3(b) and (c), January shows about $p_1 =$ 30% variance explained by the first mode while August shows around 33%. Their second modes drop to 11% and 6% for January and August respectively. Their third modes are around 5% for both January and August. Figure 3(b) shows that for January, eigenvalues λ_4 and λ_5 are close to each other and have a very small difference $p_4 - p_5$. EOF errors are large because they are proportional to the inverse of the difference of two neighborhood eigenvalues, i.e., $1/(\lambda_4 - \lambda_5)$, according to North's rule of thumb (North et al. (1982)). This implies that the mode mixture between EOF4 and EOF5 is stronger, and the geometric pattern of EOF4 may not have a meaningful physical interpretation due to the uncertainty caused by the mode mixing. For August, this mode mixing occurs from EOF3. Therefore, in this paper, we focus on the physical interpretations of EOF1 and EOF2. Up to mode 2, the cumulative percentage variance q_2 reaches slightly
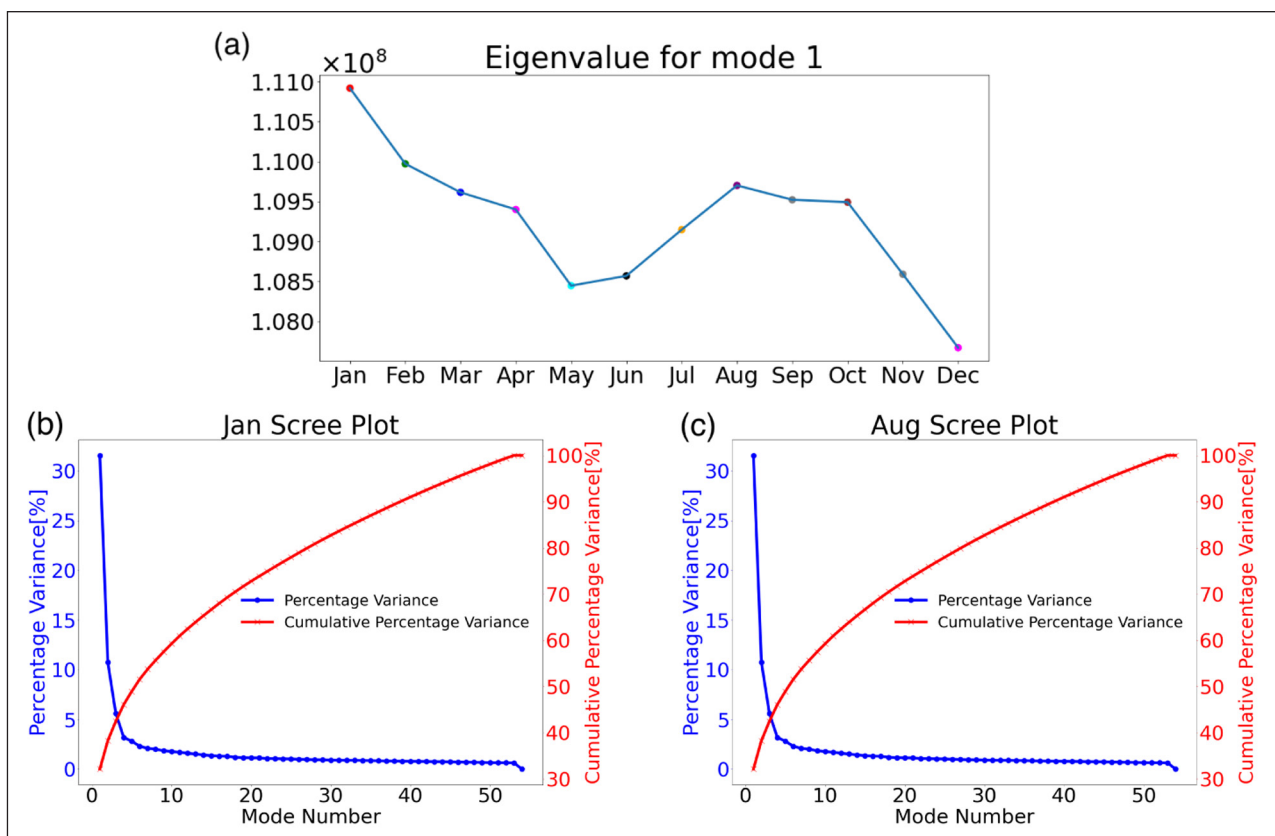

Figure 3 (a) The first eigenvalue λ_1 of the temporal covariance matrix for the weighted temperature anomalies for each month from January to December. Units of the first eigenvalue are the same as the covariance matrix in equation 14. **(b)** The scree plot for January based on data p_k and q_k. **(c)** The scree plot for August is based on data p_k and q_k.

over 40% for January and slightly below 40% for August. Meaning a large number of variances are explained by EOF1 and EOF2. Considering the case of historical data reconstruction, like Shen et al. (2017), a number of modes may be used such that 80% or 90% of variances is represented. For our study here on the NASA JPL OCGM, it takes about 42 modes to represent 90% of cumulative variance for both January and August.

4.2 PHYSICAL INTERPRETATION OF EOF1: EQUATORIAL UPWELLING

Figure 4 shows our 3D EOF1 for January and August at different depth levels. We intend to use these EOF patterns to explain the ocean dynamic features associated with

EOF1, PC1, and λ_1. Recall that the first mode of January and August explains 30% and 33% variance respectively. This means that a third of the cumulative variance in the data for each month is explained by the first mode alone. 2D EOF results from area-weighted anomaly data often show that the first mode is either a manifestation of a temporal trend or ENSO (Shen et al. (2017)). Interestingly enough, the first mode of our 3D EOF1 is not only related to the temporal trend, but also the vertical structure of ocean dynamics known as the equatorial upwelling.

Equatorial upwelling can be observed in Figure 4 as the horizontal divide of cool temperatures a little north of the equator with warm temperatures straddling it. Weisberg and Qiao (2000) and American Meteorological
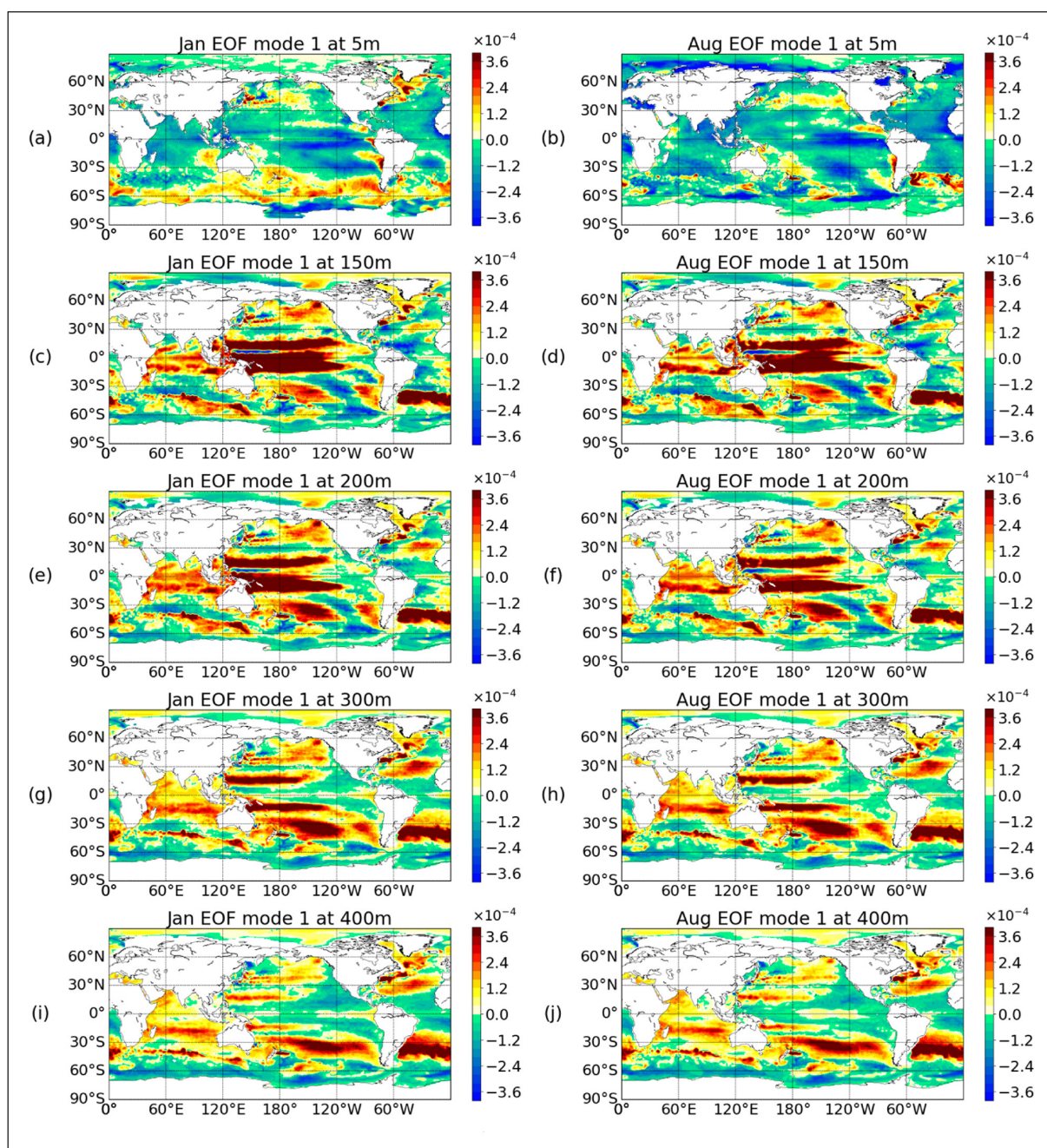

Figure 4 EOF1 for January and August showing the equatorial upwelling. Depths are chosen to show that the strong upwelling occurs at 150 meters and that the upwelling stops at a deeper level around 400 meters.

Society (2012) define equatorial upwelling as the rising of water along the equator occurring typically from 200 meters to the surface. Near the surface, the equatorial surface waters are being deflected by wind and Ekman flow. Thus, the surface waters are pushed away from the equator, and colder deep ocean waters are pumped up or more formally the deep ocean water upwells (Wyrtki and Eldyn (1982)). Our EOF1 provides a more precise quantitative description of the upwelling. Figures 4 and 5 show a strong upwelling starting around the depth of 100 meters that stops around 400 meters (specifically, see Figure 4(c) – (h)), and Figure 5(c)–(d)).

It is important to mention that a variety of signals can be explored in Figure 4. Despite the focus on equatorial upwelling in this section, there are other notable signals such as the significant warming pattern from surface to 400 meters in the Southern Ocean between 60∘S and 30°S. This is a widely known result of ocean warming due to global warming (see the satellite altimetry data from Song and Colberg (2011)). This is one of many signals to dissect from the first 3D EOF mode, but we will mainly focus on equatorial upwelling for the remainder of this subsection.

The signal strength of the equatorial upwelling process varies with depths. Looking at the surface layer for both January and August in Figure 4(a)–(b), the signal seems relatively weak, if we treat our 3D EOF1 as the upwelling mode. January shows a small line of cooler waters a little west of 120°E to approximately 170°E.

August is less obvious with a smaller diffused line of cool water between 130°E to 170°E. Although the line of cool waters is narrow and small on the map, the region still covers a large equatorial area because Figure 4 is in the 2D Mercator map projection.

The upwelling region gets larger in deeper layers, particularly at 150 meters we see a very clear divide between 120°E and 180°E in Figure 4. The strong upwelling continues at 200 meters for both months, but January shows a stronger signal as expected from its higher eigenvalue (see Figure 3). At 300 meters, the two warmer strips off the equator begin to fade away. The warm-cool-warm sandwich pattern of the upwelling signal starts to disappear and eventually has little reminiscence at 400 meters suggesting the vertical sandwich has a depth of about 400 meters. The length in the tropical Pacific in the east-west direction stretches about 13,000 kilometers from 120°E to 120°W. The "bread pieces" of warm waters in both south and north are thicker than the thin "cheese filling" in the middle, as indicated by Figures 4 and 5.

To show how the upwelling signal changes by depth from the perspective of cross-section, we have made a north-south cross-sectional cut along a meridional line of 160°E, shown in Figure 5. To first identify upwelling in Figure 5, we look for a cold section of water just north of the equator with warmer waters around it. Both January and August show cold waters at the surface between the equator and 10° north. Figure 5(a) and (b) suggest that upwelling stops before 500 meters as the 3D EOF1 is quite
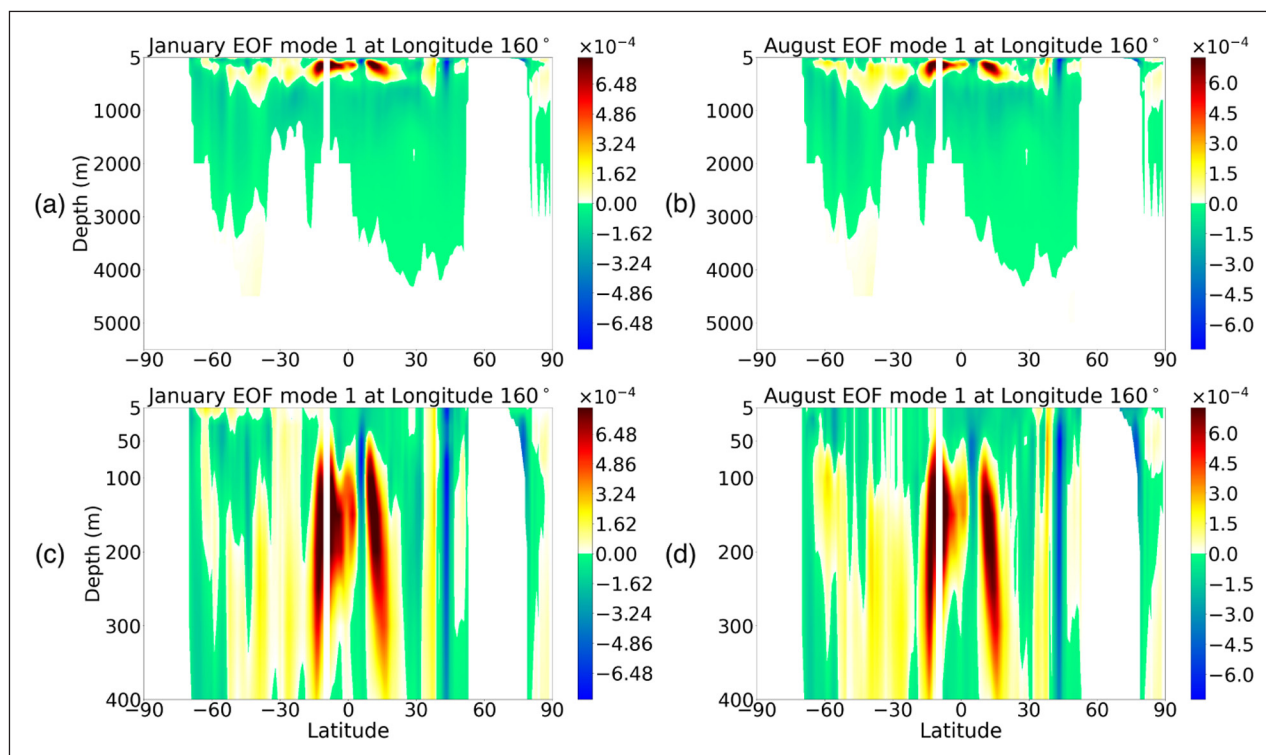

Figure 5 A north-south cross-section of the ocean taken at 160° E to show at what layers upwelling occurs. Panels **(a)** and **(b)** are the entire scope of all ocean layers. Panels **(c)** and **(d)** show the top layers up to 400 meters as this will better show where upwelling starts and stops.

uniform and shows little spatial variation for layers deeper than 500 meters. These observations motivated us to focus instead on depths from the surface to 400 meters where we can see more clearly a sandwich pattern of an equatorial cold (the narrow blue region at the equator) wrapped by two pieces of off-equatorial warm waters (the red regions). At about 100 meters, the warmer sections are distinct, continuing until they reach their strongest reading at around 150 meters. The warmer patches start to become thinner at 200 meters with the southern half between 10°S and 0° almost entirely gone for August at 250 meters and January at 300 meters. There is a small region in the southern patch around 20°S that persists to 400 meters along with the northern warm patch at 20°N. The entire upwelling region is limited to a zonal bend between 20°S and 20°N. However, this is only the EOF1 pattern, which reflects a temporally mean state. For a particular month, the equatorial upwelling region may be much larger, reaching 30°S and 30°N (Figure 6).

Recall EOFs are just a statistical tool used to show spatial patterns. We can identify the dynamics within the spatial patterns, and quantify their variance using EOFs. Equatorial upwelling has not been widely explored in previous EOF calculations therefore it is difficult to compare these results to other calculations such as Meyers et al. (1999), Mu et al. (2018), and Shen et al. (2017), whose results are favorable to ENSO. However, equatorial upwelling is an observable dynamic as seen in Figure 6. By visualizing the OGCM data we are able to identify what dynamics to expect of equatorial upwelling in EOF1.

Figure 6 plots the OGCM output for the strongest upwelling months and years: January 1993, January 2003, August 1990, and August 2003. We can observe that upwelling for all dates occurs in both the Pacific and Atlantic oceans. The upwelling signal is the strongest in the western and the central tropical Pacific, in contrast to the ENSO warm pattern over the eastern and central tropical Pacific. The upwelling over the western tropical Pacific in January 1993 (see Figure 6(a)) has a quantified warm-cold-warm sandwich pattern as 20°C – 12°C – 20°C, covering from 30°S to 35°N.

Focusing on January 1993 in Figure 6 the equatorial upwelling patterns in the Atlantic are also limited to the western part of the Atlantic ocean. The southern part of the sandwich bread piece encompasses a small area. In the Indian Ocean, the northern part of the sandwich bread piece almost does not exist. The upwelling signal is much weaker than that in the Pacific and Atlantic. The same can be said for the upwelling in all other months shown in Figure 6 when examining the water temperature distribution at the depth of 200 meters.

The high-latitude green regions of Figure 6 show another vertical structure of ocean dynamics: Ocean ventilation. This is the process in which surface water, recently in contact with the atmosphere, sinks, transporting carbon and oxygen to deeper ocean layers (Khatiwala et al. (2012) Naveira Garabato et al. (2017). Regions such as the Labrador Sea are more sensitive to ocean ventilation as its cold winter air will cause surface waters to lose heat to the atmosphere becoming cooler, denser, less buoyant, and consequently sinking deeper (Marshall and Schott (1999) Koelling et al. (2022)). These Labrador sea waters (LSW) that have undergone such convection can be seen in Figure 7.

This cross-sectional figure is a zonal mean, i.e., an average from longitude 0 to 360°. We should expect to see homogeneous values for depths up to 2500 meters at latitudes relative to the Labrador sea according to Yashayaev and Loder (2017). Observing Figure 7 from 55°N to 60°N, we can identify LSW occurring from 50 meters to 3000 meters. This range is much deeper than what was suggested from recent papers like Koelling et al. (2022) and Yashayaev and Loder (2017).
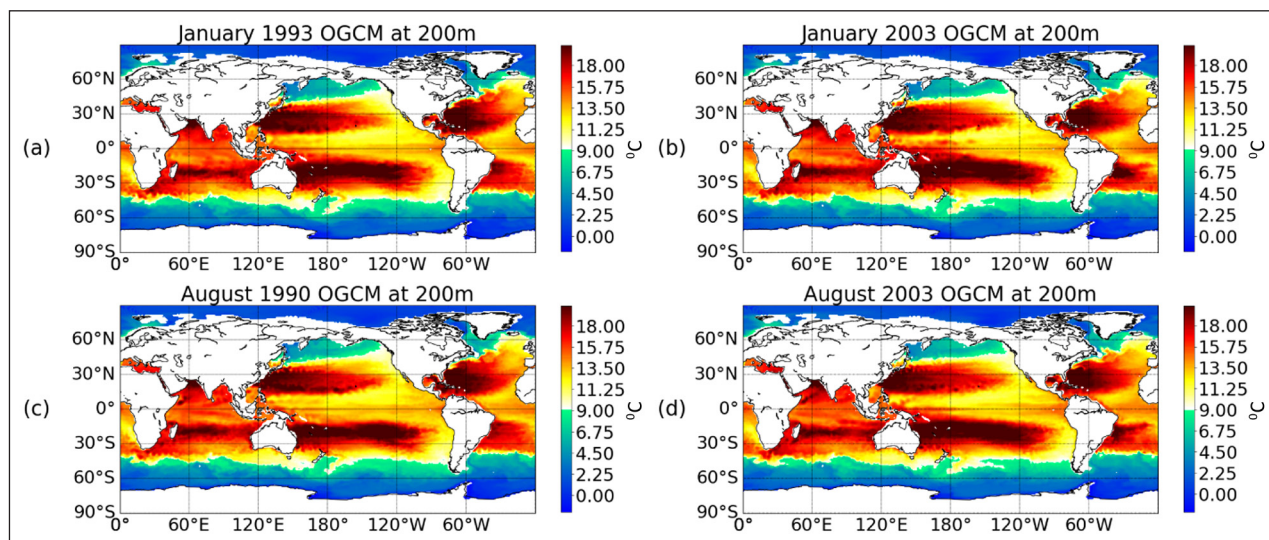

Figure 6 NASA JPL OGCM temperature for January and August at 200 meters showing equatorial upwelling. Years are chosen to reflect time with strong upwelling.
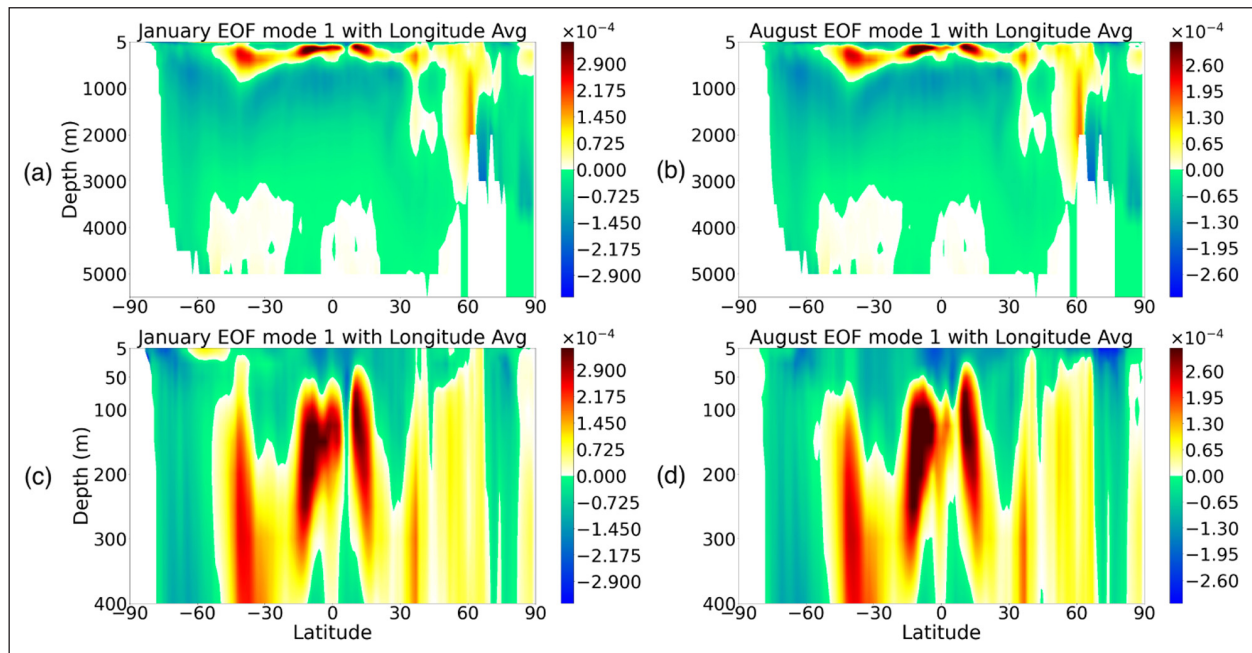
Figure 7 Cross-sectional map based on the zonal mean from 0 to 360° longitude degrees to show ocean ventilation at the high latitude regions.

The cross-sectional zonal mean map of Figure 7 also shows equatorial upwelling. The August upwelling signal at about 100 meters depth appears to be stronger than that of January according to Figure 7(c) and (d).

4.3 EOF2 AND THE SPATIAL PATTERN OF EL NIÑO

Figure 8 shows the eigenvalue λ_2 of the temporal covariance matrix for every month. In contrast to λ_1 in Figure 3, there is no longer a large jump from December to January in λ_2. As previously mentioned the smooth transition is due to an annual cycle.
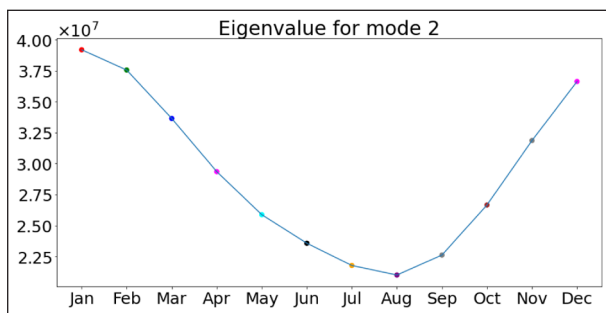
Figure 8 The second eigenvalue λ_2 of the temporal covariance matrix for the weighted temperature anomalies. Units of the first eigenvalue are the same as the covariance matrix in equation 14. January has the largest value, while August has the smallest value.

Figure 9 shows the EOF2 pattern at five selected layers: 5, 50, 100, 150, and 200 meters. The warm pattern in the eastern tropical Pacific is the well-known ENSO mode, a phenomenon that occurs mostly in the winter months of the northern Hemisphere as seen in NOAA CPC (2005). The winter ENSO pattern corresponds to the large eigenvalue λ_2 in winter months. Particularly December, November,

and January have the three largest λ_2 values according to Figure 8. This is expected as higher eigenvalues imply larger variance and vice versa. The figure also shows the three months with the smallest λ_2 values are August, July, and September indicating there is little chance for an ENSO signal to occur in these months.

As expected from SST studies and textbooks, such as Shen and Somerville (2019) and Shen et al. (2017), Figure 9 shows January has an apparent ENSO signal at its surface. The January ENSO pattern vertically exists at the surface and continues to deeper layers. The pattern starts to wither at 150 meters and is entirely dispersed at and deeper than 200 meters (see Figure 9(g) and (i)).

At 100 meters (see Figure 9(e)), we see a large continuous blue region over the western tropical Pacific that corresponds to the cold anomalies of ENSO underneath the surface as shown in Shen et al. (2017) and similar to Roemmich and Gilson (2011). This cold anomaly stretches east-west over 7,000 km and north-south over 2,000 km, almost twice the area of the United States. This body of cold anomaly water continues to exist at 200 meters (see Figure 9(i)). Our EOF2 data shows that the body begins around 50 meters and ends around 250 meters. Also at the 100 meters (see Figure 9(e)), we see the warm tongue in the eastern tropical Pacific split by a small strip of cool waters just north of the equator. This might be part of the signal of equatorial upwelling.

Although the SST ENSO events rarely occur in August, our 3D EOF2 for August still shows a warm tongue pattern over the eastern tropical Pacific. Despite that, if we were to compare either column in Figure 9 it is evident the ENSO signal in August is less coherent compared to January. This may mean that the 3D ENSO dynamics continuously exist throughout the year and do not completely die out in the summer.

For the August EOF2, in the region south of the equator, the cold waters are more persistent than in January. To better understand what depth these cooler waters reach, a zonal cross-section from the surface to the bottom of the ocean (see Figure 10) is taken at 10°S. This cross-sectional map shows consistency in the behavior displayed by Figure 9(j).

The vertical map (Figure 10) cuts through the conventional SST ENSO pattern over the tropical Pacific. Figure 10(c) shows that the eastern tropical Pacific's warm water body is limited to the top 150 meters, but the western tropical Pacific's cold water body approximately spans 50 to 250 meters. In January there is also a cold-

warm-cold pattern appearing in the Indian Ocean located within the top 250 meters as seen in Figure 10(c). However, this Indian ocean signal is not obvious for the August EOF2.

To show a more robust signal of ENSO, we have plotted the meridional mean from latitude 10°S to 10°N in Figure 11. Figure 11(c) shows a clear blue-red separation marked by the slanted yellow strip. The slope of the slanted yellow line is approximately –2 meters per longitude. This helps quantify the 3D structure of the El Niño dynamics, such as the thickness of the eastern tropical Pacific's warm water to the surface, and the thickness of the western tropical Pacific's cold water underneath the surface. January ENSO's warm pattern
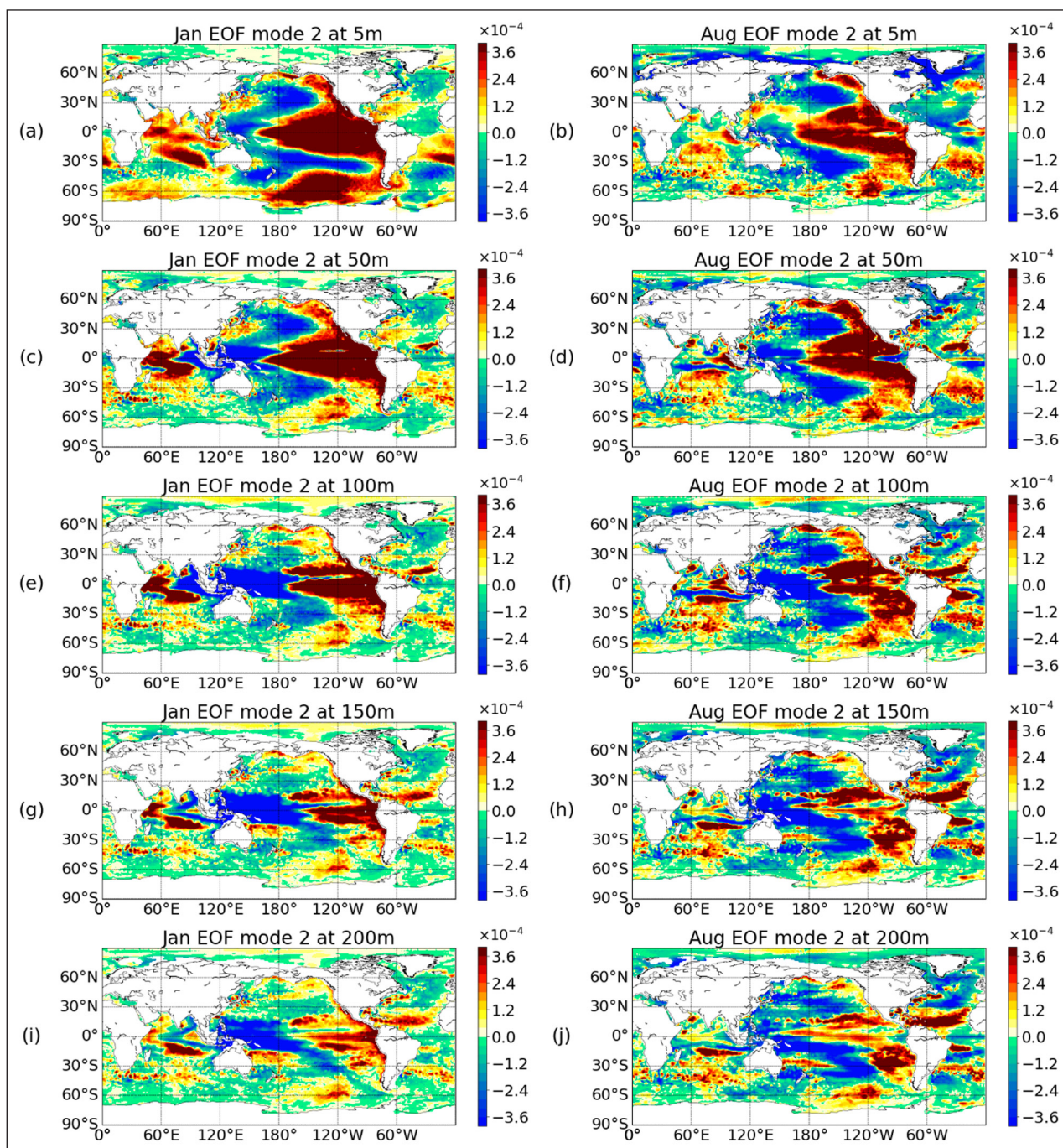

Figure 9 EOF2 for January and August showing the ENSO pattern. Depths are chosen to show the ENSO regions in different depths from surface to 100 meters. The ENSO pattern dies out approximately at levels deeper than 150 meters.
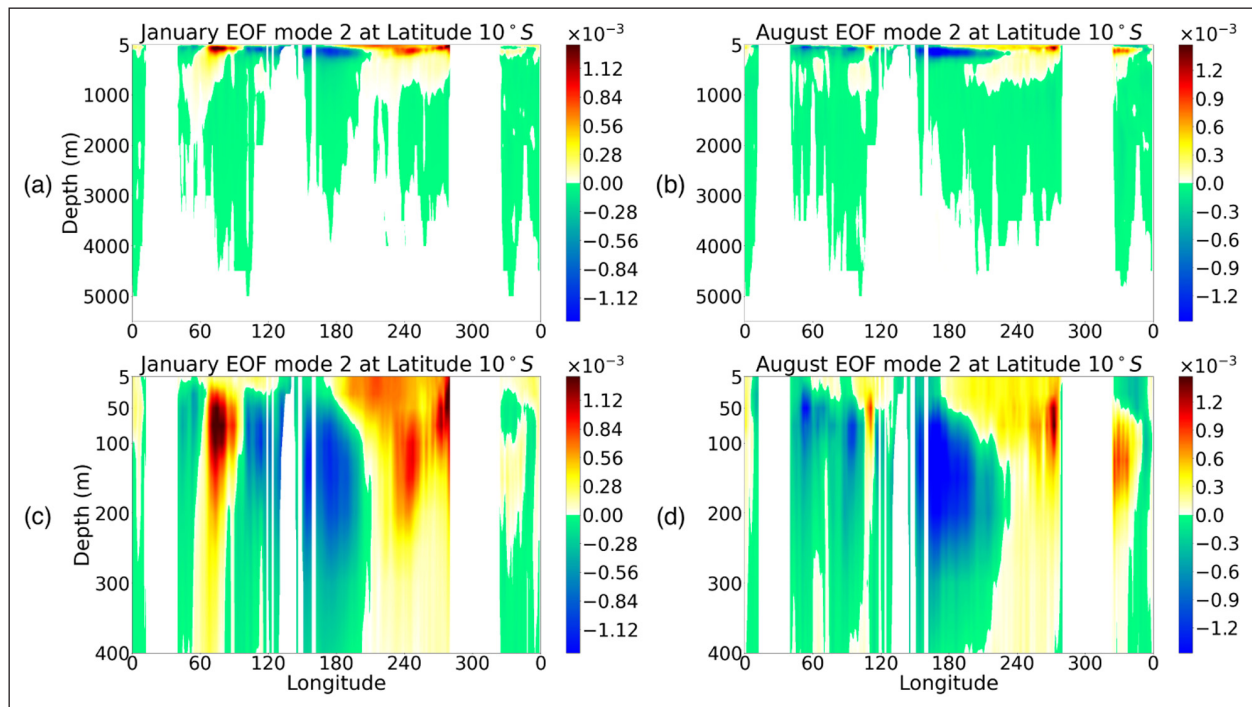
Figure 10 Cross section at latitude of 10°S that shows the separation of the warm and cold anomaly regions in the Pacific.
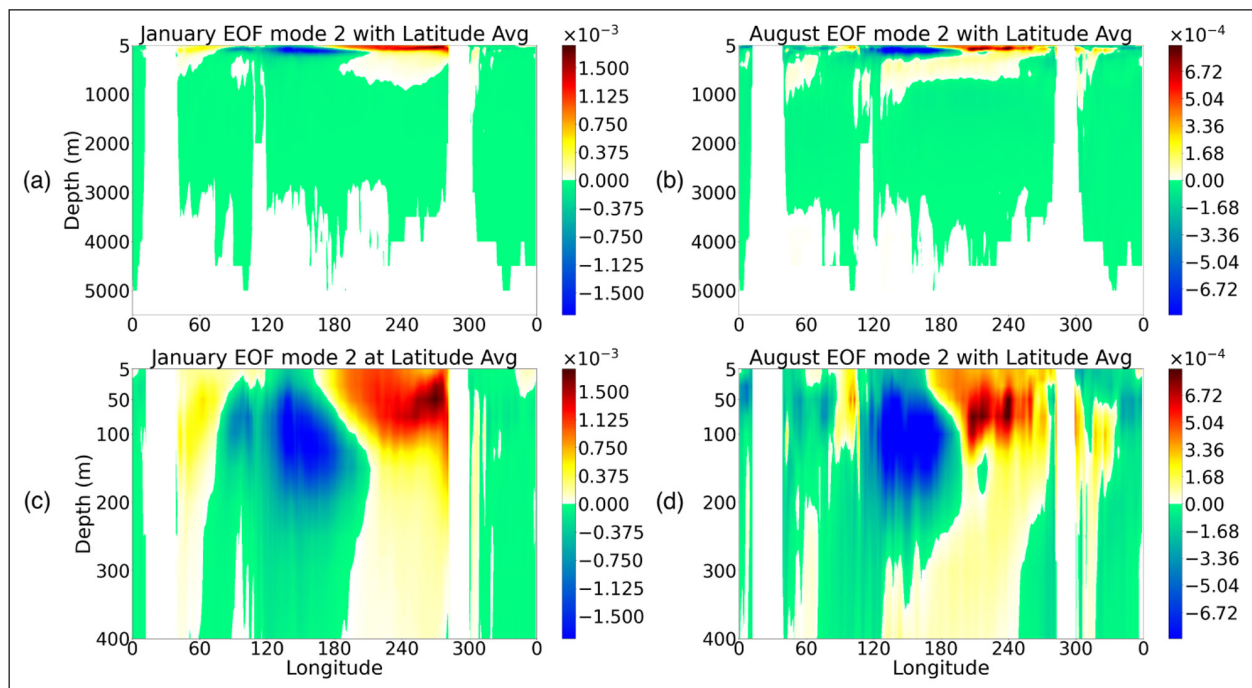

Figure 11 A zonal cross-sectional map based on the average from latitude 10°N to 10°S. The average and figure are to better show a robust 3D ENSO pattern with not only the surface warm region but also a depth structure.

ends at approximately 150 meters while August ENSO's warm pattern ends at around 120 meters (see Figure 11(c) and (d)).

4.4 PCS, TREND AND DE-TRENDED

This sub-section describes the temporal patterns of the 3D EOF-PC analysis. Principal components (PCs) are the eigenvectors of the temporal covariance matrix and are also the column vectors of the orthogonal V matrix. PCs are on a time scale and show temporal patterns.

PC1 shown in Figure 12(a) and (b) show an upward trend of temperature anomalies. Then what would be the cause of this upward trend? Does the association of EOF1 with equatorial upwelling imply that the upwelling contains some heat from global warming? If it does contains some heat from global warming, how much? The detailed mechanisms of heat transfer and distribution are still to be investigated, even if the association between upwelling and global warming is true. Some literature has already appeared to address

these questions. For example, Song and Colberg (2011) explored the existence of warming in the deep ocean; Wang et al. (2022) studied some specifics of global warming on equatorial upwelling. Studies like these may serve as the basis for further quantitative investigations on whether or how upwelling and global warming are correlated.

PC2 of January is supposed to be a time indicator of ENSO occurrence. This interpretation works well for the SST analysis on ENSO. When regarding 3D ocean, we may need to consider the time delay of heat transfer in the top 1500 meters. The El Niño pattern in the August EOF2 has already implied this time delay (see Figure 9). It is unknown how long it takes for the heat accumulation in

the eastern tropical Pacific from the surface to 150 meters and for the heat loss in the western tropical Pacific in the deep layers from about 50 to 250 meters. Figure 12(c) shows that the minimum of January PC2 is about −0.32 at January 1990. According to NOAA CPC (2001) Oceanic Niño Index (ONI) based on SST, the northern hemispheric winter of 1989–1990 was a neutral year, while 1988–1989 winter was a La Niña year, 1990–1991 winter was neutral too, and 1991–1992 winter was an El Niño year.

The second minimum value of January PC2 (shown in Figure 12(c)) is about 0.30 occurred on January 1975, which was La Niña month. The maximum January PC2 is about 0.31 occurred in January 2003. The northern hemispheric winter of 2002–2003 had an El Niño, while
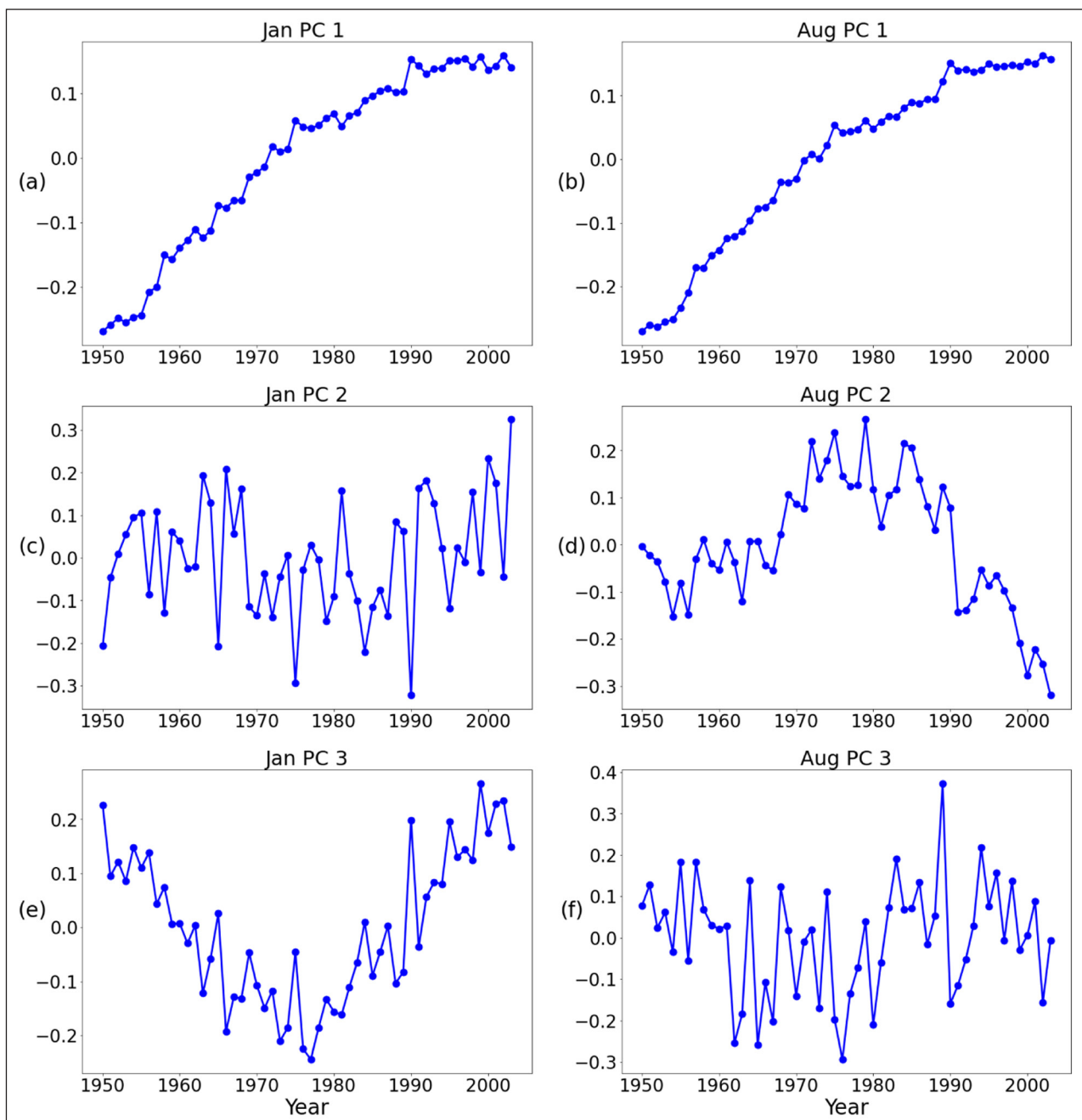

Figure 12 Principal components PC1, PC2, and PC3 of January and August from the weighted anomalies of the NASA JPL OGCM output from 1950 to 2003.

the previous winter was neutral. The above analysis of the maxima and minima implies that PC2 from the 3D covariance matrix, although associated with El Niño, is not a direct indicator of the surface El Niño time. Further analysis may be needed to interpret PC2 and its association with El Niño.

A similar analysis may be done for the maxima and minima of August PC2. Figure 12(d) shows that the global maximum of the August PC2 is about 0.25 occurred in August 1979, which was a neutral month. The global minimum of the August PC2 is about –0.32 occurred in August 2003, the same year, i.e., 2003, as the global maximum of the January PC2. As discussed earlier, the northern hemispheric winter of 2002–2003 was in El Niño condition according to SST, but the summer of 2003 was in neutral condition.

The January PC3 shown in Figure 12(e) has a V-shape, going down from the 1950s to the 1970s, and then going up from the 1970s to the 1990s. This shape agrees with that of the Atlantic Multi-decadal Oscillation (AMO) index time series, and also that of the spatial average time series of the northern hemispheric average temperature, cumulative Southern Oscillation Index (SOI), and the cumulative weighted SOI1 (CSOI1) (see Figs. 4.2 and 4.3 of Shen and Somerville (2019)). We limit ourselves in claiming similar temporal patterns among these parameters. Whether the similarity is supported by climate dynamics may be worth further investigation.

August's PC3 shown in Figure 12(f) has its global maximum in 1989 and several local minima in the 1960s and 1970s. This temporal pattern appears to agree with that of the North Atlantic Oscillation (NAO) (Hurrell (1995)). Again, we present here only the similarity of the temporal patterns. The climate dynamic support for this similarity needs further investigation.

The trend in PC1 motivated us to explore the EOF-PC analysis for the de-trended anomalies. This was done by subtracting the predicted values of the simple linear regression from the anomalies. The same volume weight is applied to the de-trended anomalies as described in the methodology section. Finally, the temporal covariance matrix can be calculated from the weighted de-trended anomalies to obtain EOFs, PCs, and eigenvalues.

By removing the trend we hope to gain insight as to why there is such a large jump of λ_1 between December and January shown in Figure 3. Figure 13 displays the resulting first eigenvalues $\lambda_1^{(d)}$ of the temporal covariance matrix from the de-trended anomalies for each month from January to December.
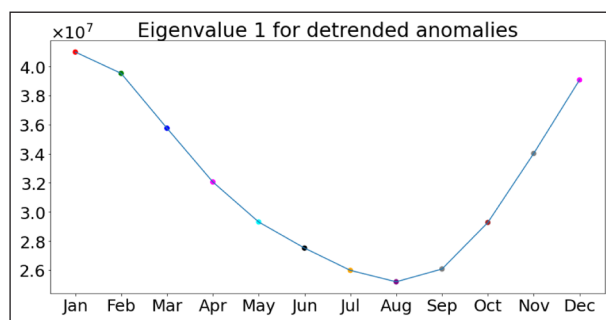

Figure 13 The first eigenvalue $\lambda_1^{(d)}$ for the each month computed from the de-trended anomalies.
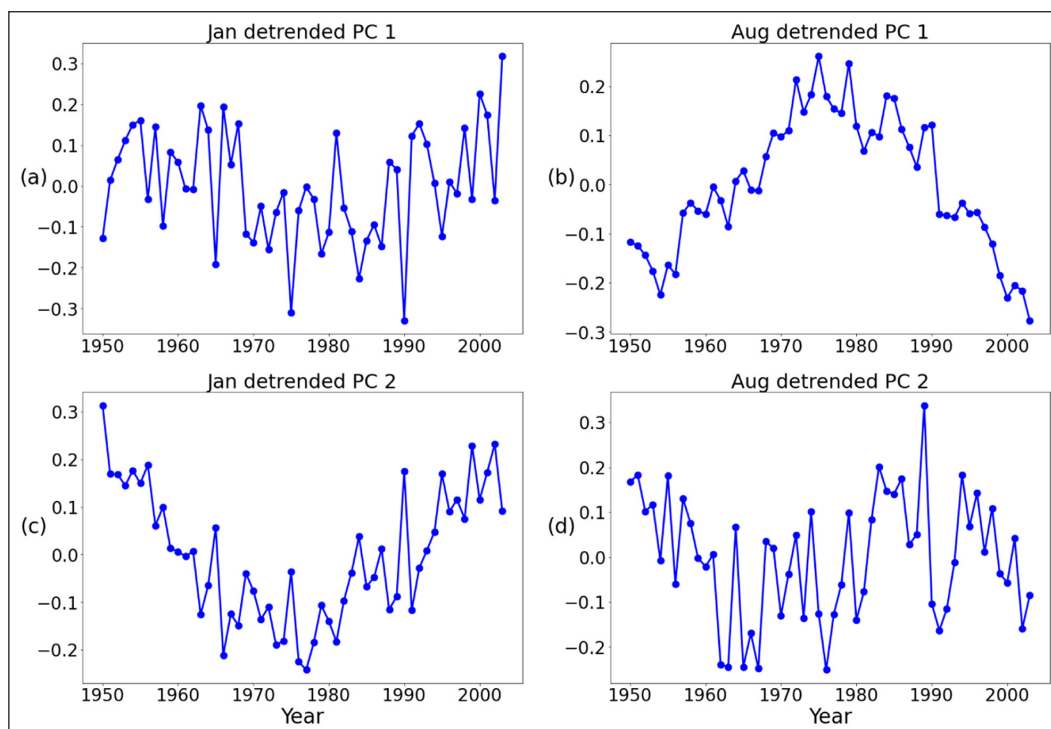

Figure 14 Principal components of January and August based on the de-trended anomalies.

The shape of $\lambda_1^{(d)}$ for the de-trended anomalies is almost the same as that of λ_2 before the de-trending shown in Figure 8, although the values of $\lambda_1^{(d)}$ and λ_2 are not exactly the same. The large gap between December to January in λ_1 is gone in $\lambda_1^{(d)}$ (see Figures 3 and 13 respectively), implying the trend may have caused the gap. On that account the temporal trend in the 12-month time period from January to December is reflected in λ_1, corresponding to the equatorial upwelling mode.

In removing the linear trend the ENSO mode becomes the first mode. Comparing the PCs before de-trending the anomalies (Figure 12) and after de-trending the anomalies (Figure 14), it is clear that there is a mode shift: De-trended PC1 corresponds to PC2 before de-trend, and de-trended PC2 corresponds to PC3 before de-trend. This mode shift due to de-trending often occurs in EOF analysis (see Chapters 9 and 10 of Shen and Somerville (2019)).

4.5 COMPARE THE 3D EOF1 TO THE 2D EOF1

As a way to compare results with other publications and reinforce the value of computing 3D EOFs, we computed 2D EOFs. It is important to note that 2D EOFs are orthonormal on each layer, in contrast, 3D EOFs are orthonormal only for the entire ocean domain of all layers, and not orthonormal on each layer. Figure 15 shows the 2D EOF1 for the top layer based on the anomalies, not the de-trended anomalies, from the NASA JPL OGCM.

To reiterate the 2D EOF calculation was only applied to the surface layer so ocean dynamics interconnection with other layers is not considered. Consequently, ENSO, not the equatorial upwelling, appears in the first mode of 2D EOFs, similar to Shen et al. (2017), Meyers et al. (1999), and Mu et al. (2018). The spatial pattern of Figure 15's 2D EOF1 is very similar to that of the 3D EOF2 at the top layer shown in Figure 9(a). However, the data for Figure 9(a) is not normalized to one and is not orthogonal to the 3D EOF2 at the same layer.
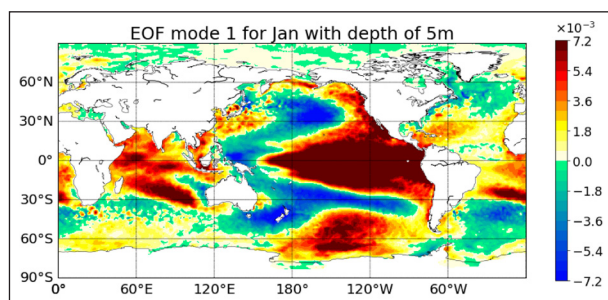

Figure 15 2D EOF1 for the top layer based on the area-weighted anomalies from the NASA JPL OGCM. ENSO signal is shown in this EOF1 for the surface layer. The 2D EOFs are computed layer by layer.

5 CONCLUSIONS AND DISCUSSION

The main purpose of this paper is to show the method of computing 3-dimensional EOFs and interpret the

physical meanings of the first two modes. To compute these EOFs, we used the temporal covariance of the volume-weighted anomalies based on the NASA JPL OGCM's monthly output of 1.0 latitude-longitude degree, 33 layers from surface to 5,500 meters depth, and a time span from 1950 to 2003. The eigenvectors of the temporal covariance are PCs. The normalized projections of the volume-weighted anomalies for the PCs form the geometric EOFs. Removing the volume factors we obtain the physical EOFs, which are displayed in our EOF figures.

Our method to compute the January PCs starts from the formulation of the $N \times Y$ space-time data matrix with the first dimension being the total number of grid boxes and the latter being the total number of years. Anomalies are multiplied by the volume weights of each depth, resulting in the weighted anomalies. After removing land/NaN data, the new weighted anomaly matrix transpose multiplied by itself forms the temporal covariance matrix.

We also computed the EOFs and PCs using SVD for the volume-weighted anomalies. Although SVD is capable of handling our data at the 1.0-degree resolution, it cannot handle the higher resolution at 1/4 degree due to computer memory resource limitations. Thus, our temporal covariance approach is useful and will be used later on to deal with 1/4-degree data or higher resolutions. Our application of SVD in this paper is for the purpose of verifying our temporal covariance approach.

NASA JPLs OGCM was used before in Shen et al. (2017) where EOFs were computed layer by layer, i.e., the two-dimensional EOF, or 2D EOF. The paper shows the ENSO signal in the first one or two EOFs, and it did not reveal the equatorial upwelling mode. This was due to the nature of a 2D calculation, which does not take into account the deep ocean dynamics interconnected among different layers and instead treats each layer in an isolated fashion. Our 3D EOFs show covariances across depth layers. Consequently, equatorial upwelling appears as 3D EOF1 and explains a third of the data's variance. In contrast, the ENSO mode, or 3D EOF2, only explains 10% of the variance.

By observing the covariance among the temperatures of different layers, we were not only able to identify different ocean dynamics than the previous 2D calculations, but also we were able to pinpoint how deep each characteristic spans. Equatorial upwelling can be seen from 100 to 400 meters. ENSO is observed mostly in the upper layers until 150 meters, and at 100 meters we can observe a large cool water patch in the western tropical Pacific ocean.

Although through the 3D EOFs and PCs, we were able to identify some meaningful ocean dynamic patterns both in the deep ocean and surface, there is still much to explore in the resulting dynamics and the effect of de-trending. Further exploration of the ocean dynamics may be made to show how the 3D EOFs are related to climate oscillation patterns, such as AMO and NAO.

Our EOFs can serve not only as a means to understand dynamics but also help in reconstructing historical climate data (Shen et al. (2017)). The 3D nature of the EOFs allows the covariance properties between layers to be used in the data reconstruction. The covariance between points of different depths are expected to add skills to reconstruction. For example, the 3D in-situ observational data can likely support more 3D EOF modes than the 2D EOF modes in the sub-surface of the ocean where in-situ observations are scarce. Can we extend the reliable reconstruction of historical data to a longer history compared with the conventional 2D reconstruction? Can the historical in situ data support a set of 3D EOFs with a finer spatial resolution? More questions like these can be asked about the 3D EOFs with regard to their applications, computations, climate dynamics, and mathematical theory.

DATA ACCESSIBILITY STATEMENT

The datasets generated during and/or analyzed during the current study are available in the "Calculate 3D EOFs" repository, https://doi.org/10.5281/zenodo.7042027.

The original NASA JPL OGCM data are available at JPL through a co-author Y. Tony Song yuhe.t.song.dr@jpl.nasa.gov.

CODE AVAILABILITY

The computer code for plotting all the figures in this paper is freely available and can be downloaded from https://doi.org/10.5281/zenodo.7042027. This GitHub link also provides the Python code for the computing for the entire paper.

ACKNOWLEDGEMENTS

FUNDING INFORMATION

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

All authors contributed to the study's conception and design. Material preparation, data collection, and analysis were performed by all authors DL, TB, YTS, TMS, and SS. All code to compute and visualize 3D EOFs was written by DL and TB. The first draft of the manuscript was written by DL and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

AUTHOR AFFILIATIONS

Danielle Lafarga orcid.org/0009-0006-9695-0650
Computational Science Research Center, San Diego State, University, San Diego, CA 92182, USA; NOAA EPP MSI Center for Earth System Sciences and Remote, Sensing Technologies, and Computational Science Research, Center, San Diego State University, San Diego, CA 92182, USA

Thomas Bui
Big Data Analytics Program, San Diego State University, San Diego, CA 92182, USA; NOAA EPP MSI Center for Earth System Sciences and Remote, Sensing Technologies, and Computational Science Research, Center, San Diego State University, San Diego, CA 92182, USA

Y. Tony Song orcid.org/0000-0001-6903-0804
Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109, USA

Thomas M. Smith orcid.org/0000-0001-7469-7849
NOAA/STAR/SCSB and CISESS/ESSIC, University of Maryland, 5825 University Research CT, College Park, MD 20740, USA

Samuel S. P. Shen orcid.org/0000-0002-4535-4684
Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92182, USA

REFERENCES

Acar, E and **Yener, B.** 2009. "Unsupervised Multiway Data Analysis: A Literature Survey." *IEEE Transactions on Knowledge and Data Engineering*, 21(1): 6–20. DOI: https://doi.org/10.1109/TKDE.2008.112

American Meteorological Society. 2012. "Equatorial Upwelling. Glossary of Meteorology." https://glossary.

ametsoc.org/wiki/Equatorial_upwelling#:~:text=The%20 rising%20of%20water%20along,either%20side%20of%20 the%20equator. Accessed 1 May 2022.

Gittens, A, Devarakonda, A, Racah, E, Ringenburg, M, Gerhardt, L, Kottalam, J, Liu, J, Maschhoff, K, Canon, S, Chhugani, J, et al. 2016. "Matrix factorizations at scale: A comparison of scientific data analytics in Spark and C+ MPI using three case studies." In *2016 IEEE International Conference on Big Data (Big Data),* 204–213. IEEE. DOI: https://doi.org/10.1109/BigData.2016.7840606

Haidvogel, DB, Arango, HG, Hedstrom, K, Beckmann, A, Malanotte-Rizzoli, P and **Shchepetkin, AF.** 2000. "Model evaluation experiments in the North Atlantic Basin: simulations in nonlinear terrainfollowing coordinates." *Dynamics of Atmospheres and Oceans*, 32(3): 239–281. DOI: https://doi.org/10.1016/S0377-0265(00)00049-X

Houser, C, Smith, A, Wernette, P and **Lehner, J.** 2020. "Spatial Frequency Analysis and Information Synthesis for Understanding Coastal Barriers." In *Reference Module in Earth Systems and Environmental Sciences.* Elsevier. DOI: https://doi.org/10.1016/B978-0-12-818234-5.00023-7

Huang, RX and **Jin, X.** 2002. "Sea Surface Elevation and Bottom Pressure Anomalies due to Thermohaline Forcing. Part I: Isolated Perturbations." *Journal of Physical Oceanography – J PHYS OCEANOGR*, 32. DOI: https://doi. org/10.1175/1520-0485(2002)032<2131:SSEABP>2.0. CO;2

Hurrell, JW. 1995. "Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation." *Science*, 269(5224): 676–679. DOI: https://doi.org/10.1126/ science.269.5224.676

Khatiwala, SP, Primeau, FW and **Holzer, M.** 2012. "Ventilation of the deep ocean constrained with tracer observations and implications for radiocarbon estimates of ideal mean age." *Earth and Planetary Science Letters*, 325–326, 116–125. DOI: https://doi.org/10.1016/j.epsl.2012.01.038

Koelling, J, Atamanchu, D, Karstensen, J, Hand**mann, P** and **Wallace, DWR.** 2022. "Oxygen export to the deep ocean following Labrador Sea Water formation." *Biogeosciences*, 19(2): 437–454. DOI: https://doi.org/10.5194/bg-19-437-2022

Liang, S, Li, X and **Wang, J.** 2012. "High-level Land Product Integration Methods." In *Advanced Remote Sensing*, 667–690. Boston: Academic Press. DOI: https://doi.org/10.1016/ B978-0-12-385954-9.00022-8

Marshall, J and **Schott, F.** 1999. "Open-ocean convection: Observations, theory, and models." *Reviews of Geophysics*, 37(1): 1–64. https://arxiv.org/abs/https://agupubs. onlinelibrary.wiley.com/doi/pdf/10.1029/98RG02739. DOI: https://doi.org/10.1029/98RG02739

Meyers, SD, O'Brien, JJ and **Thelin, E.** 1999. "Reconstruction of Monthly SST in the Tropical Pacific Ocean during 18681993 Using Adaptive Climate Basis Functions." *Monthly Weather Review*, 127(7): 1599–1612. DOI: https:// doi.org/10.1175/1520-0493(1999)127<1599:ROMSIT>2.0 .CO;2

Moon, J-H, Song, YT, Bromirski, PD and **Miller, AJ.** 2013. "Multidecadal regional sea level shifts in the Pacific over 19582008." *Journal of Geophysical Research: Oceans*, 118(12): 7024–7035. https://arxiv.org/abs/https://agupubs. onlinelibrary.wiley.com/doi/pdf/10.1002/2013JC009297. DOI: https://doi.org/10.1002/2013JC009297

Mu, D, Yan, H and **Feng, W.** 2018. "Assessment of Sea Level Variability Derived by EOF Reconstruction." *Geophysical Journal International*, 214(1): 79–87. https:// arxiv.org/abs/https://academic.oup.com/gji/article-pdf/214/1/79/27718510/ggy126.pdf. DOI: https://doi. org/10.1093/gji/ggy126

National Center for Atmospheric Research Staff. (eds.) 2013. "The Climate Data Guide: Empirical Orthogonal Function(EOF) Analysis and Rotated EOF Analysis." Accessed Jan 20, 2022. https://climatedataguide.ucar.edu/ climate-data-tools-and-analysis/empirical-orthogonal-function-eof-analysis-and-rotated-eof-analysis.

Naveira Garabato, AC, MacGilchrist, GA, Brown, PJ, Evans, DG, Meijers, AJS and **Zika, JD.** 2017. "High-latitude ocean ventilation and its role in Earth's climate transitions." *Philos. Trans. A Math. Phys. Eng. Sci.*, 375(2102): 20160324. DOI: https://doi.org/10.1098/rsta.2016.0324

NOAA CPC. 2001. "Cold and Warm Episodes by Season. NOAA/National Weather Service." Accessed May 15, 2022. https://origin.cpc.ncep.noaa.gov/products/analysis_ monitoring/ensostuff/ONI_v5.php.

NOAA CPC. 2005. "El Nio/Southern Oscillation (ENSO). NOAA/National Weather Service." Accessed May 15, 2022. https://www.cpc.ncep.noaa.gov/products/precip/CWlink/ MJO/enso.shtml.

North, GR, Bell, TL, Cahalan, RF and **Moeng, FJ.** 1982. "Sampling errors in the estimation of empirical orthogonal functions." *Monthly weather review*, 110(7): 699–706. DOI: https://doi.org/10.1175/1520-0493(1982)110<0699:SEITE O>2.0.CO;2

Roemmich, D and **Gilson, J.** 2011. "The global ocean imprint of ENSO." *Geophysical Research Letters*, 38(13). DOI: https:// doi.org/10.1029/2011GL047992

Sabatier, R and **Traissac, P.** 1994. "The ACT (STATIS method)." *Comput. Statist. Data Anal*, 18: 97–119. DOI: https://doi. org/10.1016/0167-9473(94)90134-1

Shchepetkin, AF and **McWilliams, JC.** 2003. "A method for computing horizontal pressure-gradient force in an oceanic model with a nonaligned vertical coordinate." *Journal of Geo-physical Research: Oceans*, 108(C3). https://arxiv.org/abs/https://agupubs.onlinelibrary.wiley. com/doi/pdf/10.1029/2001JC001047. DOI: https://doi. org/10.1029/2001JC001047

Shen, SSP, Behm, GP, Song, YT and **Qu, T.** 2017. "A Dynamically Consistent Reconstruction of Ocean Temperature." *Journal of Atmospheric and Oceanic Technology*, 34(5): 1061–1082. DOI: https://doi.org/10.1175/JTECH-D-16-0133.1

Shen, SSP and **Somerville, R.** 2019. "Climate Mathematics: Theory and Applications." Cambridge University Press. DOI: https://doi.org/10.1017/9781108693882

Song, T, Gross R, Wang, X and **Zlotnicki, V.** 2010. "A Non-Boussinesq Terrain-Following OGCM for Oceanographic and Geodetic Applications." *Advances in Geosciences, Volume 18: Ocean Science (OS)*, 18. DOI: https://doi.org/10.1142/9789812838148

Song, Y and **Haidvogel, D.** 1994. "A Semi-implicit Ocean Circulation Model Using a Generalized Topography-Following Coordinate System." *Journal of Computational Physics*, 115(1): 228–244. DOI: https://doi.org/10.1006/jcph.1994.1189

Song, YT and **Colberg, F.** 2011. "Deep ocean warming assessed from altimeters, Gravity Recovery and Climate Experiment, in situ measurements, and a non-Boussinesq ocean general circulation model." *Journal of Geophysical Research: Oceans*, 116(C2). https://arxiv.org/abs/https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2010JC006601. DOI: https://doi.org/10.1029/2010JC006601

Song, YT and **Hou, TY.** 2006. "Parametric vertical coordinate formulation for multiscale, Boussinesq, and non-Boussinesq ocean modeling." *Ocean Modelling* 11(3): 298–332. DOI: https://doi.org/10.1016/j.ocemod.2005.01.001

Stanimirova, I, Walczak, B, Massart, DL, Simeonov, V, Saby, CA and **Crescenzo, ED.** 2004. "STATIS, a three-way method for data analysis. Application to environmental data." *Chemometrics and intelligent laboratory systems*, 73(2): 219–233. DOI: https://doi.org/10.1016/j.chemolab.2004.03.005

Vallejo-Arboleda, A, Vicente-Villardón, JL and **Galindo-Villardón, MP.** 2007. "Canonical STATIS: Biplot analysis of multi-table group structured data based on STATIS-ACT methodology." *Computational statistics & data analysis*, 51(9): 4193–4205. DOI: https://doi.org/10.1016/j.csda.2006.04.032

Wang, L-C, Dao, TL and **Yu, J-Y.** 2022. "Continued weakening of the equatorial Pacific upwelling annual cycle in CMIP5 future projections." *Scientific Reports*, 12(1): 15595. DOI: https://doi.org/10.1038/s41598-022-19874-2

Weisberg, RH and **Qiao, L.** 2000. "Equatorial Upwelling in the Central Pacific Estimated from Moored Velocity Profilers." *Journal of Physical Oceanography*, 30(1): 105–124. DOI: https://doi.org/10.1175/1520-0485(2000)030<0105:EUITCP>2.0.CO;2

Wyrtki, K and **Eldyn, G.** 1982. "Equatorial Upwelling Events in the Central Pacific." *Journal of Physical Oceanography*, 12(9): 984–988. DOI: https://doi.org/10.1175/1520-0485(1982)012<0984:EUEITC>2.0.CO;2

Yashayaev, I and **Loder, JW.** 2017. "Further intensification of deep convection in the Labrador Sea in 2016." *Geophysical Research Letters*, 44(3): 1429–1438. https://arxiv.org/abs/https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016GL071668. DOI: https://doi.org/10.1002/2016GL071668

Zhang, Z and **Moore, JC.** 2015. "Empirical Orthogonal Functions." In *Mathematical and Physical Fundamentals of Climate Change*, 161–197. Boston: Elsevier. DOI: https://doi.org/10.1016/B978-0-12-800066-3.00006-1

Tellus A: Dynamic Meteorology and Oceanography is a peer-reviewed open access journal published by Stockholm University Press.