



Can we use sea surface temperature and productivity proxy records to reconstruct Ekman Upwelling?

Anson Cheung¹, Baylor Fox-Kemper¹, and Timothy Herbert¹

¹Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, RI 02912, USA

Correspondence: Anson H. Cheung (anson_cheung@brown.edu)

Abstract. Marine sediments have greatly improved our understanding of the climate system, but their interpretation often assumes that certain climate mechanisms operate consistently over all timescales of interest and that variability at one or few sample sites is representative of an oceanographic province. In this study, we test these assumptions using modern observations in an idealized manner mimicking paleo-reconstruction to investigate whether sea surface temperature and productivity proxy records in the Southern California Current System can be used to reconstruct Ekman upwelling. The method uses Extended Empirical Orthogonal Function (EEOF) analysis of covariation of alongshore windstress, chlorophyll and sea surface temperature as measured by satellites from 2002 to 2009. We find that EEOF1 does not reflect an Ekman upwelling pattern, but instead much broader California Current processes. EEOF2 and 3 reflect upwelling patterns, but these patterns are timescale dependent and are regional. Thus, the skill of using one site to reconstruct the large scale dominant patterns is spatially dependent. Lastly, we show that using multiple sites and/or multiple variables generally improve field reconstruction. These results together suggest caution is needed when attempting to extrapolate mechanisms that may be important on seasonal time scales (e.g. Ekman upwelling) to deeper time, but also the advantage of having multiple proxy records.

1 Introduction

The climate system varies across multiple timescales and is driven by both stochastic processes and deterministic forcings (Huybers and Curry, 2006). Paleoclimate records help us understand mechanisms of climate variability and change over long timescales by extending instrumental records beyond the historical period. Numerous studies have used paleoclimate records to understand climate system responses to different external forcings (e.g. Shakun et al., 2012), have put recent climate change into a long term context (e.g. Abram et al., 2016; PAGES2k Consortium, 2013), and have helped benchmark climate models (e.g. Harrison et al., 2015).

Marine sediment is one of the most widely used archives for paleoclimate studies. Using marine sediments for paleoclimate inference usually involves multiple steps, where one first measures multiple sensors, frequently proxies for sea surface temperature (SST) and productivity, from a single site. Then, one compares them with other nearby local records, hemispheric reconstructions, forcing reconstructions. Lastly, one applies modern large scale climatology to explain changes observed in paleoclimate records (e.g. Abram et al., 2016; Goni et al., 2006; Leduc et al., 2010; MARGO, 2009; McGregor et al., 2007; Vargas et al., 2007). While these comparisons have improved our understanding about paleoclimate significantly, uncertain-



ties and oversimplifications often may result in overly broad interpretations and assertions. Notably, this approach typically assumes that (1) certain climate mechanisms always operate over the past at all timescales of interest, and (2) large scale phenomena can be linked to variability at one or a few sample sites (i.e., a paleoclimate record location). In actuality, some

This paper illustrates an approach to test commonly asserted interpretations of SST and productivity proxy records by using
5 observational data to analyze a region where a known mechanism drives a large fraction of the variability, and with well preserved high resolution sedimentary records – the southern California region, an example of an Eastern Boundary Upwelling Systems (EBUS). There are strong scientific and societal interests to understand EBUS because physical and biogeochemical changes in these regions are known to have significant impacts on regional climate (Snyder et al., 2003; Ravelo et al., 2004; Jacox et al., 2014) and global fishery industry (Ryther, 1969; Pauly and Christensen, 1995; Ware and Thomson, 2005). Un-
10 fortunately, it remains uncertain how EBUS will change on decadal to centennial timescales in the future (Bakun et al., 2015; Garcia-Reyes et al., 2015, and references therein). Nevertheless, underlying sediments in these regions often accumulate rapidly and contain a wealth of paleoclimate information, in particular organic biomarkers and associated proxies. Thus, this has allowed high resolution (subdecadal timescale) and high quality paleoclimate reconstructions along many EBUS, which provide additional constraints on past and future changes of EBUS (e.g. Leduc et al., 2010; McGregor et al., 2007; Vargas et al., 2007).

Variability in SST and productivity reconstructions along EBUS are often regarded as a response to Ekman pumping (e.g. Leduc et al., 2010; McGregor et al., 2007; Vargas et al., 2007; MARGO, 2009). However, many other processes are also at play
15 in EBUS and can drive SST and productivity changes (e.g. eddies, zonal advection, surface heat flux variations, changes in nutrient sources and concentration forced by subsurface processes, and large-scale climate variability that affects the stratification) (Di Lorenzo et al., 2005; Chhak and Di Lorenzo, 2007; Gruber et al., 2011; Jacox et al., 2016; Rykaczewski and Dunne,
20 2010; Xiu et al., 2018). Depending on spatial and temporal timescales, these processes can overwhelm the Ekman signal in SST and productivity changes recorded by proxy records.

Here we use high resolution modern observations available during the satellite era, to probe the spatial and temporal influence of Ekman pumping on environmental parameters of interest (e.g. SST and productivity). We apply the Extended Empirical
25 Orthogonal Function (EEOF) approach (Chen and Harr, 1993) to analyze covariation between sea surface temperature, productivity, and alongshore wind stress in the Southern California Current System using high resolution satellite data. We test the hypotheses that (1) the dominant covarying EEOF pattern resembles region-wide Ekman upwelling, (2) Ekman upwelling patterns, and thus the wind-stress magnitude, can be recovered using time-averaged proxies, and (3) large-scale changes are not the dominant drivers of variability at a single paleoclimate site. We also assess the benefits of using multiple proxy records from multiple sites to better understand climate variability of the past in EBUS regions.

30 2 California Current System

The availability of numerous high resolution spatiotemporal data (e.g. repeated hydrography, gliders, satellites) and advances in modeling have allowed us to better understand variability of the California Current System (CCS) on multiple timescales. The CCS is made up of the California Current, California Undercurrent, and upwelling zones, which interact with a variety of local



topographic features and estuaries. On first order, the CCS, as a whole, is driven by large-scale climate forcing. Changes in atmospheric pressure systems (subtropical high, Aleutian low) alter wind strength and direction, which in turn affects currents' direction, strength and upwelling variability. The stratification in the region is set by large-scale features and forcing of the North Pacific. Variations in topographic features, wind forcing, freshwater inputs, and submesoscale-mesoscale features across spatial scales also play important roles in determining spatiotemporal characteristics of the CCS. Lynn and Simpson (1987); Checkley Jr and Barth (2009); Capet et al. (2008) provide overviews on dynamics of CCS and drivers of SST, chlorophyll, and wind forcing variability.

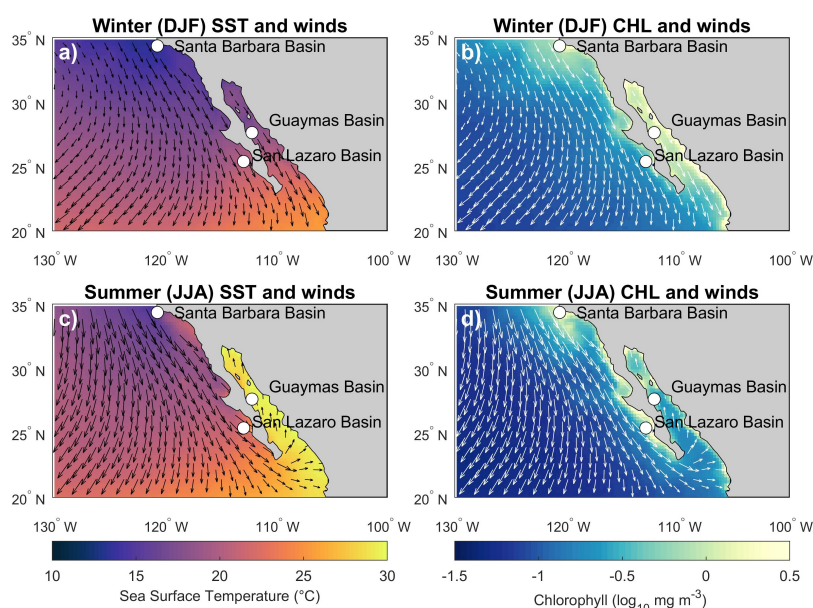


Figure 1. a) Winter (December, January, February) sea surface temperature average and wind pattern; b) Winter chlorophyll monthly average and wind pattern; c) Summer (June, July, August) sea surface temperature average and wind pattern; d) Summer chlorophyll monthly average and wind pattern. The basins highlighted are where high-resolution (subdecadal) sediment cores were previously retrieved and analyzed.

The optimal marine sediments to reconstruct subdecadal climate variability require high sedimentation rate with minimal bioturbation, and hence anoxic depositional environments. Along the CCS, these conditions mostly occur south of 24°N with exceptions of silled basins (e.g. Santa Barbara Basin) (van Geen et al., 2003). As a result, previous high resolution (subdecadal) paleoclimate studies were mostly done in the Southern part of the CCS (SCCS; Fig. 1) (e.g. Goni et al., 2006; Abella-Gutiérrez and Herguera, 2016; Zhao et al., 2000).

This study made use of high spatiotemporal resolution estimates of sea surface temperature (SST), chlorophyll-a (CHL) and alongshore wind stress (TAU) from satellite measurements to assess the role of Ekman pumping in driving SST and productivity changes along SCCS. We used Extended Empirical Orthogonal Function to assess the covariation between these variables because they are expected to be correlated spatially and temporally if Ekman theory is indeed the primary mechanism driving changes in the region. In addition, we assessed the effects of time averaging and spatial subsampling on the ability to



recover dominant patterns within the spatial window analyzed. Such assessment allows us to determine the fidelity of using proxies, which are time averaged and undersampled spatially, to understand Ekman pumping in SCCS. Details of the data and method used can be found in sections 2.1 and 2.2.

2.1 Data

5 We used sea surface temperature (SST) from GOES, chlorophyll-a (CHL) from MODIS, and alongshore wind stress (TAU) observations from QuikSCAT that span from June 2002 to November 2009. Although CHL does not equate precisely to primary productivity, and also differs from productivity inferred from proxy records, CHL provides a first order estimate of productivity (Henson et al., 2010). All data were derived and available from National Aeronautics and Space Administration Jet Propulsion Laboratory PO.DAAC and ocean color data server.

10 For TAU calculation, we used the descending pass of level 3 gridded Jet Propulsion Laboratory v2 QuikSCAT surface wind observations (SeaPAC, 2006). The QuikSCAT satellite is equipped with the SeaWinds scatterometer, a microwave radar that measures ocean radar backscatter over a cross section, which varies with satellite parameters and surface geometry (Freilich et al., 1994; Chelton and Freilich, 2005). Surface wind vectors can be estimated using model functions to estimate the relationship between wind and radar backscatter over the cross section. Level 3 data were derived using the Direction Interval Retrieval
15 with Threshold Nudging wind vector solutions based on Level 2B data, which used the QSCAT-1B geophysical model function (Perry, 2001). Level 3 QuikSCAT data provides $0.25^\circ \times 0.25^\circ$ spatial resolution on a daily timescale. The QuikSCAT accuracy is about 0.75 ms^{-1} in the along-wind component and about 1.5 ms^{-1} in the crosswind component (Chelton and Freilich, 2005).

We utilized SST observations from Geostationary Environmental Satellites (GOES). GOES satellites provide near-time SST measurements along the west coast of North America. We used level 3 gridded GOES 6km near time real SST daily data
20 after May 12, 2003 (NOAA/NESDIS, 2003b) and averaged hourly SST data to daily mean resolution prior to May 12, 2003 (NOAA/NESDIS, 2003a). Level 3 GOES SST data provides $0.05^\circ \times 0.05^\circ$ spatial resolution with better than 1K SST accuracy (Wick et al., 2002).

For CHL concentrations, we used ocean color from the Moderate Resolution Imaging Spectroradiometer on the Aqua satellite (MODIS-Aqua) (Hu et al., 2012). MODIS-Aqua is sun synchronous and measures 36 spectral bands. We used level
25 3 standard mapped image CHL measurements from MODIS-Aqua v2014.0 (O.B.G.P., 2015). Level 3 CHL data provides $0.041^\circ \times 0.041^\circ$ spatial resolution on near daily timescale with an accuracy of approximately $\pm 35\%$ (Dall'Olmo et al., 2005).

2.2 Method

2.2.1 Observation Pre-processing

To allow comparison between SST, CHL, and TAU, CHL and SST were regridded to $0.25^\circ \times 0.25^\circ$ spatial resolution. This
30 was done by bounding the datasets to $15^\circ - 45^\circ N, 130^\circ - 100^\circ W$, then calculating the area weighted CHL and SST value over each new grid cell. We further restricted our latitudinal extent to $15^\circ - 35^\circ N$ to make the analysis more computationally



efficient. Repeated analysis using different spatial domains (case 1: only east of $125^{\circ}W$; case 2: only north of $20^{\circ}N$) suggests our conclusions are insensitive to the spatial extent selected for analysis (not shown).

Since the primary interest is Ekman driven upwelling along the coast, we computed the TAU by using (1):

$$\tau = \rho C_D U |\mathbf{U}| \quad (1)$$

5 where τ = alongshore wind stress, ρ = air density, C_D = drag coefficient, U = wind speed, $|\mathbf{U}|$ = alongshore wind vector. $|\mathbf{U}|$ was calculated by summing the alongshore component of zonal and meridional wind vectors such that $-|\mathbf{U}|$ and $|\mathbf{U}|$ represent equatorward and poleward wind stress respectively. We used constant values for the coefficients, where $\rho = 1.2 \text{ kg/m}^3$ and $C_D = 1.2 \times 10^{-3}$ (Large and Pond, 1981).

Linear interpolation of all of the near-daily datasets temporally ensured uniform daily sampling rate data at each grid cell.
10 The logarithm of CHL data was taken after regridding but before EEOF analysis because CHL exhibits a nearly log-normal distribution (Campbell, 1995). Before EEOF analysis, each variable was normalized by dividing the dataset by its domain-wide and all-time standard deviation, which makes the anomaly variations in each variable comparable to each other in terms of occurrence likelihood (assuming approximately Gaussian distributions).

To follow the logic of analyzing fields that would resemble proxy records, no removal of mean or climatological states or
15 seasonality from the satellite records was performed. Thus, the EEOF analysis is performed on the total fields, rather than the anomaly fields.

2.2.2 Extended Empirical Orthogonal Function

Extended Empirical Orthogonal Function (EEOF) decomposition analysis was used to extract dominant patterns with covariation in SST, CHL, and TAU. EEOF is a variant of Empirical Orthogonal Function (EOF) analysis, a method that extracts
20 coherent, orthogonal patterns by optimizing variance into multiple orthogonal functions in time and space. Multiple variants of the EOF exist, which all involve taking into account temporal correlations of a variable or correlations between variables (e.g. Bretherton et al., 1992; Hannachi et al., 2007, and reference therein). Examining multiple time snapshots as a single field allows EOF-based analysis to extract propagating patterns (e.g. Chen and Harr, 1993) and covarying patterns (Kutzbach, 1967). Figures 2 and 3 show examples of three different EOF based methods that are fundamental to the analysis herein (EOF, EEOF
25 (temporal correlation), EEOF (multiple variables)).

The EEOF method used in this study involved extracting dominant covarying patterns by taking into account both temporal correlation within the same variable (symmetric lead-lag relationships) and correlation between variables. We employed the singular value decomposition (SVD) method (Bretherton et al., 1992) to decompose the covarying pattern of SST, CHL, and TAU into the relevant EEOF objects.

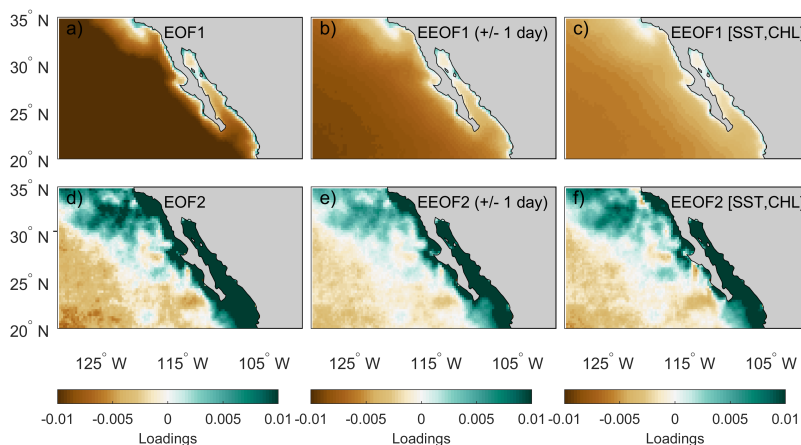


Figure 2. Example of decomposing sea surface temperature into different modes by using an (a, d) empirical orthogonal function, (b, e) an extended empirical orthogonal function with 1 day lead and lag, and (c, f) an extended empirical orthogonal function with chlorophyll included.

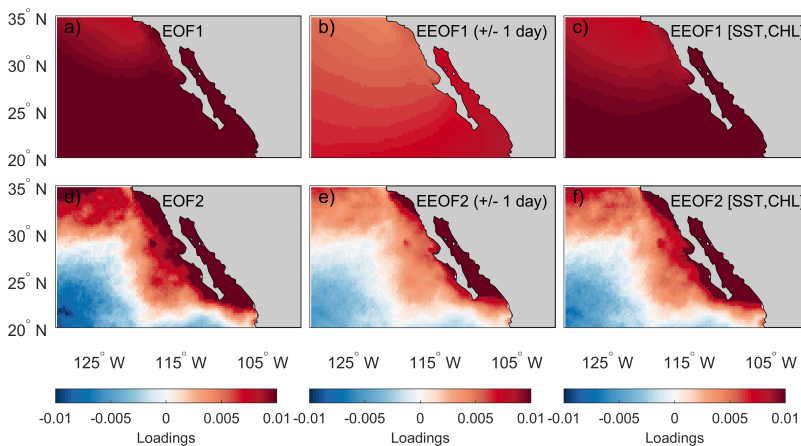


Figure 3. Example of decomposing chlorophyll into different modes by using an (a, d) empirical orthogonal function, (b, e) an extended empirical orthogonal function with 1 day lead and lag, and (c, f) an extended empirical orthogonal function with sea surface temperature included.

To consider time correlation of a variable X for EEOF analysis, we form the following data matrix:

$$X = \begin{pmatrix} x_{1,1} & \cdots & x_{1,j} & x_{1+k,1} & \cdots & x_{1+k,j} & x_{1+2k,1} & \cdots & x_{1+2k,j} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{m-2k,1} & \cdots & x_{m-2k,j} & x_{m-2k+1,1} & \cdots & x_{m-2k+1,j} & x_{m,1} & \cdots & x_{m,j} \end{pmatrix} \quad (2)$$



Where $x_{t,i}$ is a data point at a certain time snapshot t and space gridpoint i , $t = 1, 2, \dots, m$, $i = 1, 2, \dots, j$, $k =$ time unit of lead and lag included, $m =$ temporal length of the dataset, and $j =$ total spatial grid points covered. Thus, X is the concatenation of multiple reproductions of $x_{t,i}$, with each column featuring x evaluated at sequential times, and each row representing every spatial value of x , concatenated with spatial maps that are displaced in time to provide lead and lag information. Similarly, the data matrix, M , with three variables can be written as follow:

$$M = \left(SST \quad | \quad CHL \quad | \quad TAU \right) \quad (3)$$

Where SST, CHL, and TAU are submatrices with structure similar to matrix X . Note that each row of M is a complete spatiotemporal set of each variable, including every spatial location and lead and lag times for each variable, so that M is effectively the concatenation of three X matrices, one for each variable. Then, using SVD, we can decompose (3) into:

$$M = USV^T \quad (4)$$

where $U =$ a matrix of left orthogonal, singular vectors as columns with temporal information of the M matrix (Principal Components (PCs)), $S =$ singular values, and $V =$ a matrix of right orthogonal, singular vectors as columns with spatial information (Extended Empirical Orthogonal Functions (EEOFs)) of the M matrix. Note that the SVD method arrives at a basis of eigenvectors of the covariance matrices $M^T M$, i.e., $M^T M V = S^2 V$, and $M M^T$, i.e., $M M^T U = S^2 U$, so this approach is equivalent (although slightly different algorithmically) to generating EEOFs by eigenvalue decomposition.

Since proxy records reflect time averaged environmental information (usually monthly or longer), daily satellite information for analysis does not accurately depict the temporal smoothing characteristics in proxies. Hence, we performed EEOF analysis independently on daily data after averaging it into 30 days (\sim monthly), and 365 days (\sim annual) with non-overlapping means. The relatively short span of satellite observations does not allow us to extend our analysis to longer time periods that might also be of interest.

2.2.3 Determining the significance of modes and lead-lag

Instead of using singular values to determine the significance of each mode, we selected the number of modes and lead-lags to retain by evaluating the skill to reconstruct TAU. This approach was motivated by the interest of this study to detect Ekman upwelling, which involves covariation of SST, CHL, and TAU, and our inability to reconstruct TAU directly using proxies. Reconstruction of TAU (TAU_{rec}) was carried out as follow:

$$M_0 = \left(SST \quad | \quad CHL \quad | \quad TAU(0) \right) \quad (5)$$

where $TAU(0) =$ the columns for TAU in the original data matrix were replaced with zeros. Then,

$$M_{rec} = r M_0 V_n V_n^T \quad (6)$$



$$M_{rec} = \left(SST_{rec} \mid CHL_{rec} \mid TAU_{rec} \right) \quad (7)$$

where r = rescaling factor calculated by $\frac{var(SST|CHL)}{var(SST_{rec}|CHL_{rec})}$, V_n = spatial information obtained from decomposing M (Eq. 4) with n numbers of mode retained, where $n=1...5$. Note that as n is much smaller than the rank of M_{rec} , $V_n V_n^T$ is not the identity matrix, but is better thought of as the projection of M_0 onto the leading modes of M . As zero wind stress is inconsistent with any of the modes V_n , multiplying M_0 by this factor adds TAU variability back into the zeroed values that is more consistent with the observed SST and CHL , which is M_{rec} .

We used Root Mean Square Error (RMSE) as a metric to measure agreement between reconstructed TAU and actual TAU:

$$RMSE = \sqrt{\overline{(TAU_{rec} - TAU)^2}} \quad (8)$$

where $\overline{[\cdot]}$ represents mean of the data.

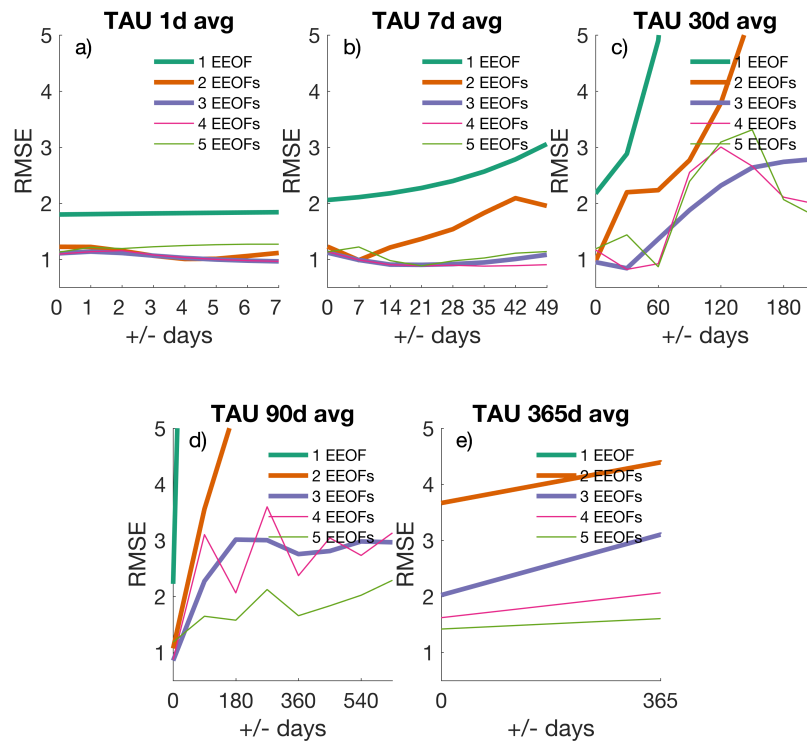


Figure 4. Root Mean Square Error of reconstructed wind stress with respect to actual wind stress using a) daily data; b) 7-day averaged data; c) 30-day averaged data; d) 90-day averaged data; e) 365-day averaged data.

10 Our analysis shows that reconstruction using three modes with no lead-lag information included provides the most stable result in predicting TAU from SST and CHL regardless of averaging timescale (Fig. 4). This result, and similar results of



convergence accuracy by adding more modes, suggest that the first three modes ($n = 3$) are reliable in this and other analyses, which will be used for the remainder of this paper.

2.2.4 Reconstruction of Principal Components

We determined how well a single proxy from one location can represent a large-scale covarying pattern by reconstructing the temporal variation extracted from EEOF analysis. We focused on three sites with previously published high-resolution paleoclimate records – Santa Barbara Basin, San Lazaro Basin, and Guaymas Basin – and two environmental variables, SST and productivity (Goni et al., 2006; Abella–Gutiérrez and Herguera, 2016; Zhao et al., 2000). This was achieved by only retaining the target time series from the location in M_{tar} :

$$M_{tar} = \begin{pmatrix} 0 & \cdots & \cdots & tar_{1,j} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & tar_{m,j} & \cdots & 0 \end{pmatrix} \quad (9)$$

Then we follow (10):

$$U_{rec} = M_{tar}(S_3V_3^T)^{-1} \quad (10)$$

where U_{rec} = reconstructed PCs, S_3 and V_3 were derived from (4) with $n = 3$ modes retained for analysis. In this scenario, only the parts of V_n associated with the target location were retained for reconstruction. The pseudo-inversion of the matrix $S_3V_3^T$ was done using Moore-Penrose pseudo-inverse, which amounts to inverting only the non-singular degrees of freedom, while zeroing out the remaining modes.

3 Results and Discussion

3.1 Does the dominant covarying pattern reflect Ekman upwelling?

EEOF analysis on daily resolution data displays spatial patterns that are distinct from what would be expected from Ekman upwelling. Instead of strong cross-shore gradients, TAU and CHL display a weak cross-shore gradient, and SST exhibits a meridional gradient pattern (Fig. 5). These patterns remain dominant when 30-day and 365-day averaged data were used.

The fact that EEOF1 does not resemble Ekman upwelling pattern has two major implications. First, this implies that wind stress is not the only forcing that drives CHL and SST changes along EBUS. Previous studies have reported different mechanisms that could control changes in CHL or SST along EBUS on various timescales. For instance, changes in subsurface nutrient concentration and sources have shown to alter primary productivity (Chhak and Di Lorenzo, 2007; Rykaczewski and Dunne, 2010) whereas surface heat flux has shown to exert dominant control on sea surface temperature in the California Current System (Di Lorenzo et al., 2005). Our study confirms these results and further iterates the importance of considering



different factors that could affect CHL and SST along EBUS, which are often used as indicators for changes in Ekman-driven upwelling. Second, paleoclimate reconstructions in the SCCS will be unlikely to reflect Ekman upwelling, in contrast to the common paradigm in the field, and couplings observed between proxy reconstructions of e.g. SST and productivity likely capture other processes.

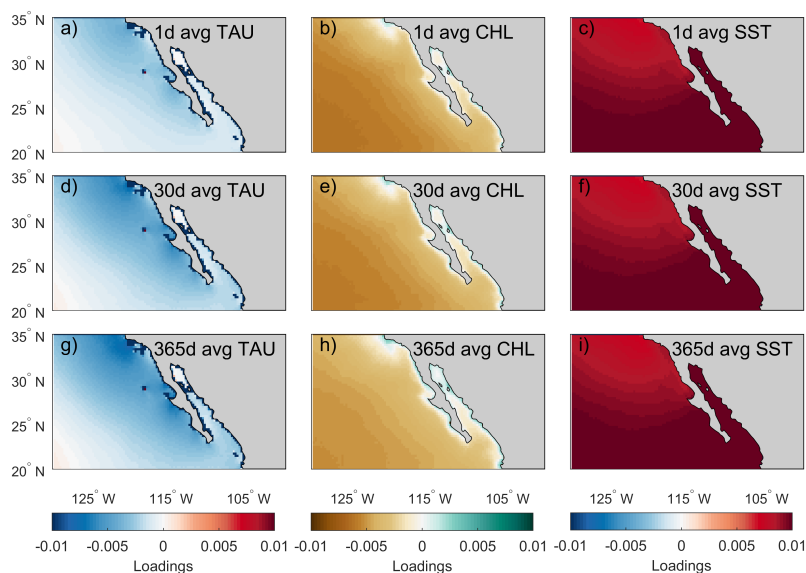


Figure 5. EEOF1 spatial pattern of TAU, CHL, and SST using a–c) daily; d–f) 30 day averaged; g–i) 365 day averaged data.

5 3.2 Can time-averaged proxies be used to reconstruct Ekman upwelling?

Even though the dominant covarying pattern does not reflect Ekman upwelling, the EEOF method allows us to decompose multiple covarying patterns for analysis. Our results suggest that EEOF2 and EEOF3 resemble Ekman upwelling pattern on daily timescales, but they reflect upwelling at different locations (Figs. 6 – 7). Specifically, EEOF2 depicts upwelling conditions off Baja California whereas EEOF3 reflects upwelling or other rapid change in conditions at Sea of Cortez. This presents an opportunity to understand whether time averaged proxies can be used to reconstruct Ekman upwelling, given optimal site selection.

Visual comparison of EEOF2 and EEOF3 patterns across different averaging windows suggest these patterns change with respect to the averaging window. For both EEOFs, their patterns resemble Ekman upwelling when data with daily resolution is used. These Ekman-upwelling-like patterns disappear when 365 day averaged data is used instead and only spatially incoherent structures are retained (bottom rows of Figs. 6 – 7). The disappearance of an Ekman upwelling pattern suggests that either Ekman upwelling is a subannual process and/or that this process is not a dominant feature on an annual timescale. While it is not possible to diagnose the underlying cause using our method, these results imply that marine sedimentary records, which generally integrate over the annual cycle, cannot capture Ekman upwelling variations in this region. Furthermore, these results



highlight the importance of considering what timescales are reflected in the proxy record. On the assumption that some proxies are seasonally biased (e.g. "integrated production temperature" applied to the interpretation of alkenone paleotemperature estimates by (Conte et al., 1992)) we add a sine weighting function (maximum in March and minimum in Sept) to the 30 day averaged dataset and reanalyze the resulting EEOF pattern. We find that the pattern is similar to the one without weighting (not shown). This suggests that the seasonal cycle does not dominate the resulting EEOF patterns over this spatial and temporal domain.

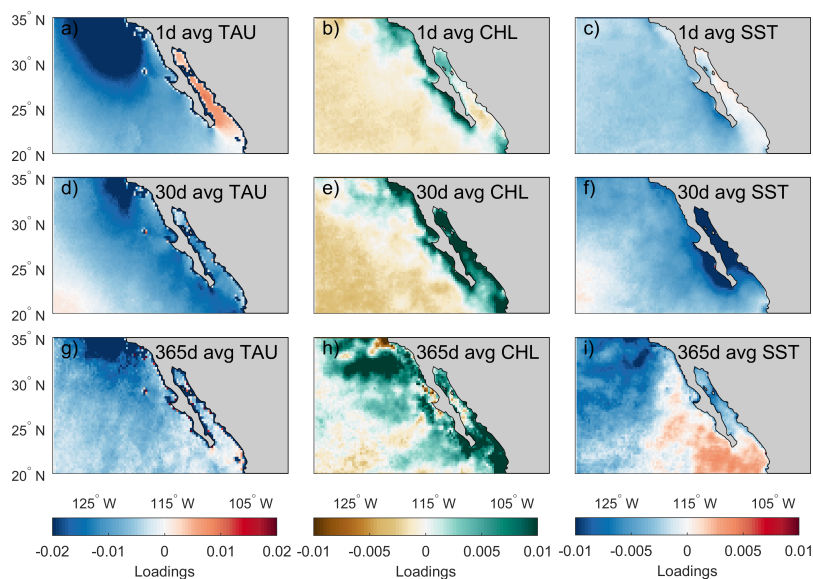


Figure 6. EEOF2 spatial pattern of TAU, CHL, and SST using a–c) daily; d–f) 30 day averaged; g–i) 365 day averaged data.

3.3 Are there benefits to analyzing records from multiple sites?

Since an upwelling pattern is only observed in analysis using daily and 30 day averaged data, we focus on assessing the potential benefits to analyzing records from multiple sites on 30 day (~ monthly) data. We acknowledge that most sedimentary records integrate over annual cycle. However, since we cannot recover upwelling pattern in the first three modes when using 365 day averaged data, we here consider an idealized situation instead, where proxy records integrate climate information on ~ monthly timescale.

With only a single proxy type measurement from one site, one can only assume it reflects the dominant large-scale circulation pattern of that area (represented by EEOF1/PC1 in this case). However, comparisons between PC1 and reconstructed PC1s based on a variable from one site shows that the ability to recover the dominant pattern depends on the location and variable (Fig. 8). This varying relationship suggests small-scale processes can drive variability at a proxy site, which can lead to behavior that is different from large-scale circulation. Therefore, caution is needed when trying to extrapolate variability in a single proxy record from one paleoclimate site to infer large-scale circulation changes. Nevertheless, in the absence of additional

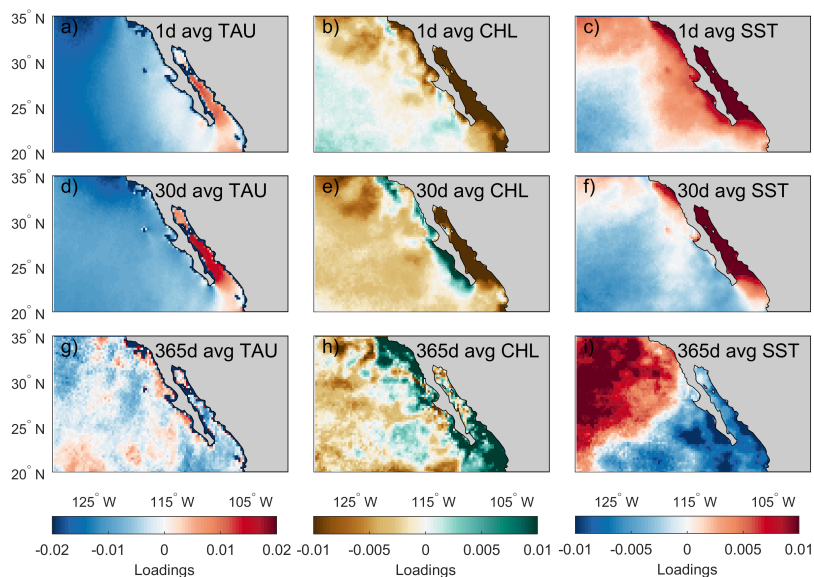


Figure 7. EEOF3 spatial pattern of TAU, CHL, and SST using a–c) daily; d–f) 30 day averaged; g–i) 365 day averaged data.

sites available to recover sediment cores, we find that measuring multiple variables often lead to better constraint of large scale climate variability (Fig. 9a).

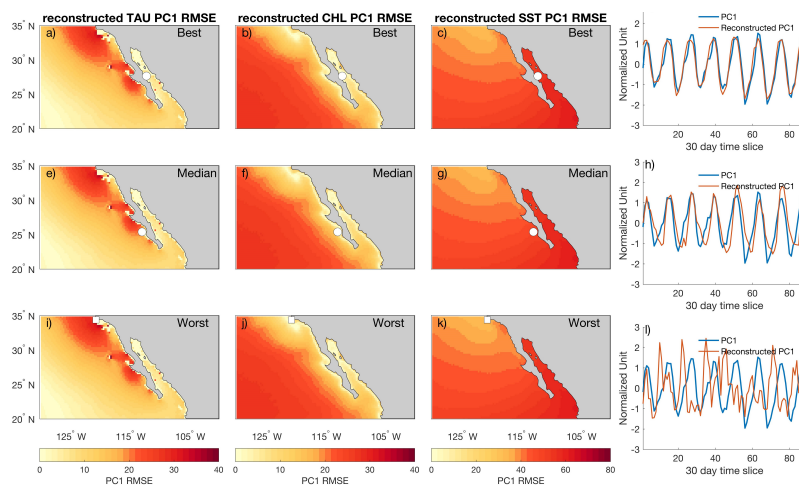


Figure 8. (a–d) Best, (e–h) Median, and (i–l) Worst PC1 reconstruction spatial RMSE and timeseries using only 1 variable from 1 site. White marker indicates the site used in that reconstruction, with circle indicating SST and square indicating CHL.

Multiple drilling expeditions in SCCS have recovered cores from different locations, which allows us to determine whether there are benefits to analyzing records from multiple sites. With multiple sites available, we can potentially reconstruct different patterns of large scale variability (Figs. 5–7). In the case of 30 day averaged data, a multiple site-based reconstruction allows



us to reconstruct spatiotemporal patterns that are associated with Ekman-driven upwelling (Fig. 9). There is also a tendency of increasing reconstruction skill when more sites and proxies are used. Therefore, there is a potential to recover multiple covarying patterns that are driven by different dynamics.

Adding reconstruction sites and variables analyzed can also potentially improve the ability to reconstruct spatiotemporal variability in the spatial domain analyzed. This has been shown in other pseudoproxy experiments that concern hemispheric reconstruction (e.g. Wang et al., 2014). Although our reconstruction technique is rather simple compared to commonly used climate field reconstruction techniques in pseudoproxy experiments and other reconstructions (e.g. Wang et al., 2014), we show that similar results emerge, where increasing number of sites and/or variables can help better reconstruct full field data that contains multiple variables (Fig. 10; Eq. 3). Therefore, these results together argue for the notion of using multiple sites and proxies for paleoclimate reconstruction.

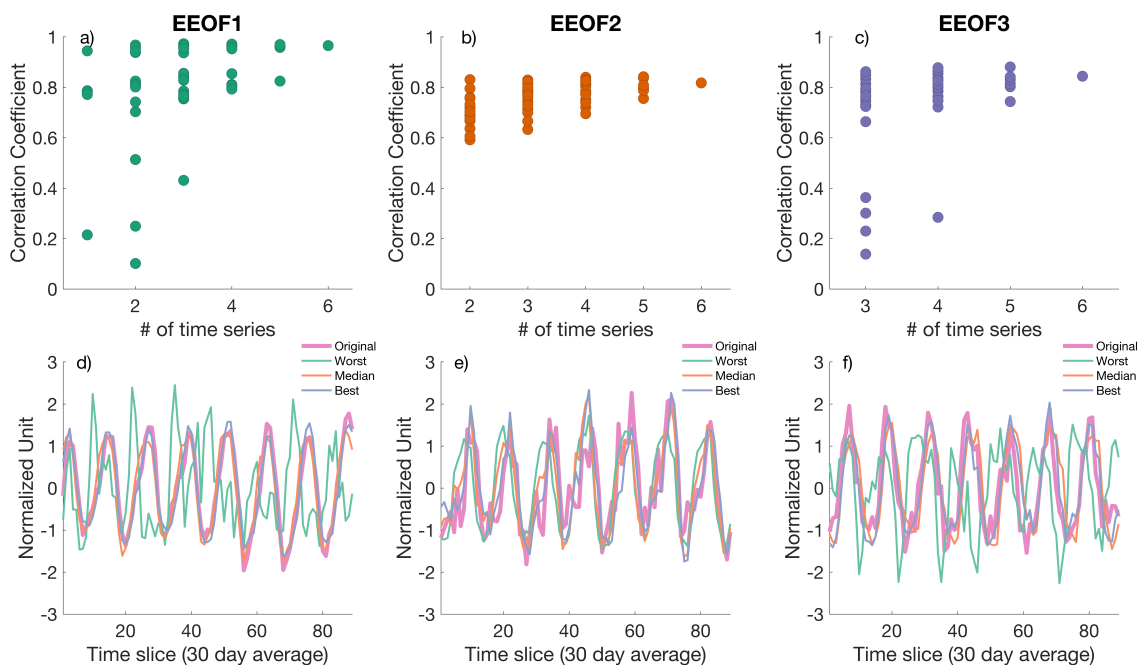


Figure 9. Correlation coefficient between reconstructed and actual a) EEOF1, b) EEOF2, c) EEOF3 temporal pattern using 30 day averaged data with varying numbers of time series from the target basins. Also shown are best, median, worst and original normalized temporal patterns of d) EEOF1, e) EEOF2, and f) EEOF3.

3.4 Implications

While this study only focuses on the case of Ekman upwelling in SCCS, it has general implications for paleoclimate studies. First, our analysis provides empirical evidence that it is important to consider the spatial representativeness of a proxy record. This calls for careful interpretation in each proxy record developed in order to avoid over simplification and over interpretation

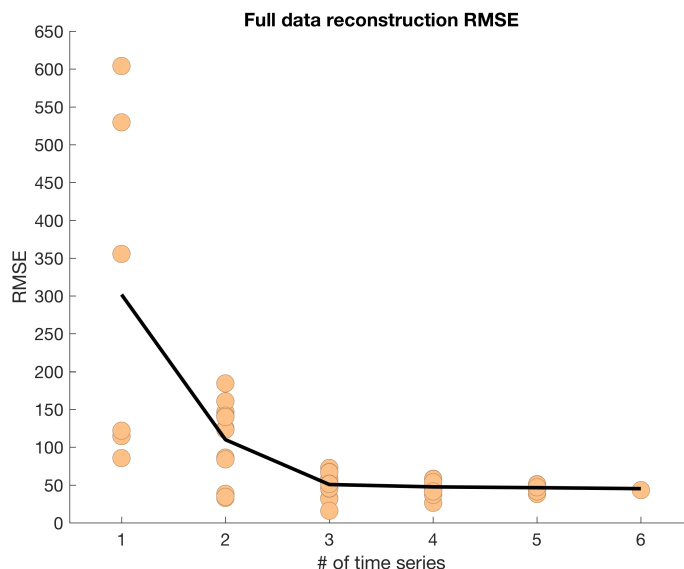


Figure 10. Full data reconstruction RMSE using different numbers of timeseries as input for reconstruction

of the climate system. Second, we demonstrate that depending on time-averaging and timescale of interest, mechanisms such as Ekman upwelling might/might not be an important process that drives variability in proxy records. Therefore, it is also important to understand whether the proxies applied and the record are able to resolve timescales where the mechanism of interest dominates (e.g. El Nino Southern Oscillation on interannual timescale). Third, we show that analyzing different proxy records from multiple sites can help us reconstruct multiple covarying patterns and improve climate field reconstruction. Last, we propose and demonstrate a multivariate method that allows us to test the assumptions regarding spatial and temporal sampling. We expect that this method can be easily applied to other regions also to provide a first order constraint on how the proxy records can be interpreted.

3.5 Limitations

- 10 There are multiple limitations that have to be taken into account when applying results from this analysis to a paleoclimate context. Firstly, our analysis is only based on 7 years of instrumental data. It is possible that the patterns established in this study are only applicable to the years analyzed due to potential nonstationary covarying relationships between the variables analyzed. Furthermore, the short length of the instrumental records does not allow us to assess the impacts of basin-scale low-frequency climate variability.
- 15 Secondly, our analysis assumes that signals from proxy records can capture surface ocean conditions perfectly and are free from other noise. This assumption is certainly violated, with multiple studies pointing to different sources of uncertainties in sedimentary records (e.g. Dolman and Laepple, 2018). Nevertheless, our analysis provides an idealized scenario to understand



assumptions associated with spatial and temporal sampling, and marks an important step to better interpreting paleoclimate records.

Thirdly, we assume that each statistical mode retrieved in this study is tied to a dynamical mechanism. However, previous studies have cautioned against such interpretations (e.g. Hannachi et al., 2007). Nevertheless, our study does not aim to diagnose Ekman upwelling processes but simply aims to determine whether it is possible to recover Ekman upwelling related patterns in proxy records. Hence, we argue the distinction between a dynamical mode and statistical mode does not undermine our results.

Lastly, our multiple record analysis assumes proxy records contain perfect age models. In most cases, this assumption is also invalid. It is inevitable that each sedimentary record contains absolute age uncertainties. Therefore, using marine sedimentary records for a multi-site proxy reconstruction with a high-temporal resolution is more challenging and might yield a different conclusion than ours.

4 Conclusions

This study aimed to evaluate assumptions commonly made in paleoclimate studies – (1) certain mechanism operates in the past on all timescales of interest, and (2) large-scale phenomena can explain the most variance in a small location (i.e. a paleoclimate site). We tested these assumptions by focusing on the Southern California Current System and used observational records to understand whether it is possible to reconstruct Ekman Upwelling using multiple sedimentary records. We introduced an Extended Empirical Orthogonal Function framework and applied it to satellite records to make inferences about paleoclimate records. Our results indicate the dominant TAU, CHL, SST covarying pattern does not resemble Ekman upwelling. In addition, the relationship between these variables appears to depend on timescales and spatial scales. A positive result is that our analysis suggests that a few sediment sites can monitor large scale fields associated with the Southern California Current. Lastly, we highlight the potential benefits of using multiple proxy records to understand different large scale covarying patterns. Our study suggests that instrumental records are helpful for testing assumptions in paleoclimatology, and the associated spatial and temporal-scale extrapolations made based on paleoclimate reconstructions. Testing these assumptions might help us better interpret proxy records and understand past climate changes.

Author contributions. AHC, BFK, TDH conceived the study. AHC and BFK designed the analysis. AHC carried out the analysis and wrote the manuscript with inputs from BFK and TDH.

Competing interests. The authors declare no conflict of interest

Acknowledgements. The Geostationary Environmental Satellites SST, QuikSCAT surface wind data were obtained from the NASA EOSDIS Physical Oceanography Distributed Active Archive Center (PO.DAAC) at the Jet Propulsion Laboratory, Pasadena, CA. MODIS chlorophyll-



a data were obtained from NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. We thank R. Vachula, N. Richter, S. Garelick for helpful comments, suggestions and discussions. AHC was supported by the Brown University Presidential Fellowship. BFK was supported by ONR N00014-17-1-2393.



References

- Abella-Gutiérrez, J. and Herguera, J. C.: Sensitivity of carbon paleoproductivity in the Southern California Current System on different time scales for the last 2ka, *Paleoceanography*, 31, 953–970, <https://doi.org/10.1002/2015PA002872>, 2016.
- Abram, N. J., McGregor, H. V., Tierney, J. E., Evans, M. N., McKay, N. P., and Kaufman, D. S.: Early onset of industrial-era warming across the oceans and continents, *Nature*, 536, 411–418, <https://doi.org/10.1038/nature19082>, 2016.
- Bakun, A., Black, B. A., Bograd, S. J., Garcia-Reyes, M., Miller, A. J., Rykaczewski, R. R., and Sydeman, W. J.: Anticipated Effects of Climate Change on Coastal Upwelling Ecosystems, *Curr. Clim. Change. Rep.*, 1, 85–93, <https://doi.org/10.1007/s40641-015-0008-4>, 2015.
- Bretherton, C., Smith, C., and Wallace, J.: An Intercomparison of Methods for Finding Coupled Patterns in Climate Data, *J. Climate*, 5, 541–560, 1992.
- Campbell, J. W.: The lognormal distribution as a model for bio-optical variability in the sea, *J Geophys Res Oceans*, 100, 13 237–13 254, 1995.
- Capet, X., McWilliams, J. C., Molemaker, M. J., and Shchepetkin, A.: Mesoscale to submesoscale transition in the California Current System. Part I: Flow structure, eddy flux, and observational tests, *Journal of physical oceanography*, 38, 29–43, 2008.
- Checkley Jr, D. M. and Barth, J. A.: Patterns and processes in the California Current System, *Progress in Oceanography*, 83, 49–64, <https://doi.org/10.1016/j.pocean.2009.07.028>, 2009.
- Chelton, D. and Freilich, M.: Scatterometer-Based Assessment of 10-m Wind Analyses from the Operational ECMWF and NCEP Numerical Weather Prediction Models, *Mon. Wea. Rev.*, 133, 409–429, <https://doi.org/10.1175/MWR-2861.1>, 2005.
- Chen, J.-M. and Harr, P. A.: Interpretation of Extended Empirical Orthogonal Function (EEOF) Analysis, *Mon. Wea. Rev.*, 121, 2631–2636, 1993.
- Chhak, K. and Di Lorenzo, E.: Decadal variations in the California Current upwelling cells, *Geophys. Res. Lett.*, 34, L14 604, <https://doi.org/10.1029/2007GL030203>, 2007.
- Conte, M. H., Eglinton, G., and Madureira, L. A.: Long-chain alkenones and alkyl alkenoates as palaeotemperature indicators: their production, flux and early sedimentary diagenesis in the Eastern North Atlantic, *Organic Geochemistry*, 19, 287 – 298, [https://doi.org/https://doi.org/10.1016/0146-6380\(92\)90044-X](https://doi.org/https://doi.org/10.1016/0146-6380(92)90044-X), 1992.
- Dall’Olmo, G., Gitelson, A. A., Rundquist, D. C., Leavitt, B., Barrow, T., and Holz, J. C.: Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands, *Remote Sensing of Environment*, 96, 176–187, 2005.
- Di Lorenzo, E., Miller, A. J., Schneider, N., and McWilliams, J. C.: The Warming of the California Current System: Dynamics and Ecosystem Implications, *Journal of Physical Oceanography*, 35, 336–362, <https://doi.org/10.1175/JPO-2690.1>, 2005.
- Dolman, A. M. and Laepple, T.: Sedproxy: a forward model for sediment-archived climate proxies, *Climate of the Past*, 14, 1851–1868, <https://doi.org/10.5194/cp-14-1851-2018>, 2018.
- Freilich, M. H., Long, D. G., and Spencer, M. W.: SeaWinds: a scanning scatterometer for ADEOS-II-science overview, *Proceedings of IGARSS ’94 - 1994 IEEE International Geoscience and Remote Sensing Symposium*, Pasadena, CA, USA, 2, 960–963, 1994.
- Garcia-Reyes, M., Sydeman, W. J., Schoeman, D. S., Rykaczewski, R. R., Black, B. A., Smit, A. J., and Bograd, S. J.: Under Pressure: Climate Change, Upwelling, and Eastern Boundary Upwelling Ecosystems, *Front. Mar. Sci.*, 2, 109, <https://doi.org/10.3389/fmars.2015.00109>, 2015.



- Goni, M. A., Thunell, R. C., Woodworth, M. P., and Müller-Karger, F. E.: Changes in wind-driven upwelling during the last three centuries: Interocean teleconnections, *Geophys. Res. Lett.*, 33, L15 604, <https://doi.org/10.1029/2006GL026415>, 2006.
- Gruber, N., Lachkar, Z., Frenzel, H., Marchesiello, P., Münnich, M., McWilliams, J. C., Nagai, T., and Plattner, G.-K.: Eddy-induced reduction of biological production in eastern boundary upwelling systems, *Nat. Geosci.*, 4, 787–792, <https://doi.org/10.1038/NGEO1273>, 2011.
- Hannachi, A., Jolliffe, I. T., and Stephenson, D. B.: Empirical orthogonal functions and related techniques in atmospheric science: A review, *Int. J. Climatol.*, 27, 1119–1152, <https://doi.org/10.1002/joc.1499>, 2007.
- Harrison, S. P., Bartlein, P. J., Izumi, K., Li, G., Annan, J., Hargreaves, J., Braconnot, P., and Kageyama, M.: Evaluation of CMIP5 palaeo-simulations to improve climate projections, *Nat. Clim. Change*, 5, 735–743, <https://doi.org/10.1038/NCLIMATE2649>, 2015.
- Henson, S., Sarmiento, J., Dunne, J., Bopp, L., Lima, I., Doney, S., John, J., and Beaulieu, C.: Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity, *Biogeosciences*, 7, 621–640, 2010.
- Hu, C., Lee, Z., and Franz, B.: Chlorophyll-a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference, *J Geophys Res Oceans*, 117, <https://doi.org/10.1029/2011JC007395>, 2012.
- Huybers, P. and Curry, W.: Links between annual, Milankovitch and continuum temperature variability, *Nature*, 441, 329–332, <https://doi.org/10.1038/nature04745>, 2006.
- Jacox, M., Moore, A., Edwards, C., and Fiechter, J.: Spatially resolved upwelling in the California Current System and its connections to climate variability, *Geophysical Research Letters*, 41, 3189–3196, 2014.
- Jacox, M. G., Hazen, E. L., and Bograd, S. J.: Optimal Environmental Conditions and Anomalous Ecosystem Responses: Constraining Bottom-up Controls of Phytoplankton Biomass in the California Current System, *Sci. Rep.*, 6, 27 612, <https://doi.org/10.1038/srep27612>, 2016.
- Kutzbach, J.: Empirical Eigenvectors of Sea-Level Pressure, Surface Temperature and Precipitation Complexes over North America, *J. Appl. Meteor.*, 6, 791–802, 1967.
- Large, W. and Pond, S.: Open Ocean Momentum Flux Measurements in Moderate to Strong Winds, *J. Phys. Oceanogr.*, 11, 324–336, [https://doi.org/https://doi.org/10.1175/1520-0485\(1981\)011<0324:OOMFMI>2.0.CO;2](https://doi.org/https://doi.org/10.1175/1520-0485(1981)011<0324:OOMFMI>2.0.CO;2), 1981.
- Leduc, G., Herbert, C. T., Blanz, T., Martinez, P., and Schneider, R.: Contrasting evolution of sea surface temperature in the Benguela upwelling system under natural and anthropogenic climate forcings, *Geophys. Res. Lett.*, 37, L20 705, <https://doi.org/10.1029/2010GL044353>, 2010.
- Lynn, R. J. and Simpson, J. J.: The California Current system: The seasonal variability of its physical characteristics, *J. Geophys. Res.*, 92, 12 947–12 966, <https://doi.org/10.1029/JC092iC12p12947>, 1987.
- MARGO: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum, *Nat. Geosci.*, 2, 127–132, <https://doi.org/10.1038/NGEO411>, 2009.
- McGregor, H. V., Dima, M., Fischer, H. W., and Mulitza, S.: Rapid 20th-Century Increase in Coastal Upwelling off Northwest Africa, *Science*, 315, 637–639, <https://doi.org/10.1126/science.1134839>, 2007.
- NOAA/NESDIS: GOES Level 3 6km Near Real Time SST 1 Hour. Ver. 1. PO.DAAC, CA, USA., Dataset accessed [2018-06-27] at <http://dx.doi.org/10.5067/GOES3-1HOUR>, 2003a.
- NOAA/NESDIS: GOES Level 3 6km Near Real Time SST 24 Hour. Ver. 1. PO.DAAC, CA, USA., Dataset accessed [2018-06-27] at <http://dx.doi.org/10.5067/GOES3-24HOR>, 2003b.



- O.B.G.P.: MODIS Aqua Level 3 Global Daily Mapped 4 km Chlorophyll a v2014.0. Ver. 2014.0. PO.DAAC, CA, USA., Dataset accessed [2018-07-26], 2015.
- PAGES2k Consortium: Continental-scale temperature variability during the past two millennia, *Nat. Geosci.*, 6, 339—346, <https://doi.org/10.1038/NGEO1797>, 2013.
- 5 Pauly, D. and Christensen, V.: Primary production required to sustain global fisheries, *Nature*, 374, 255—257, 1995.
- Perry, K. L.: SeaWinds on QuikSCAT Level 3 Daily, Gridded Ocean Wind Vectors (JPL SeaWinds Project), Version 1.1, JPL Document D-20335, Jet Propulsion Laboratory, Pasadena, CA, 2001.
- Ravelo, A. C., Andreasen, D. H., Lyle, M., Lyle, A. O., and Wara, M. W.: Regional climate shifts caused by gradual global cooling in the Pliocene epoch, *Nature*, 429, 263, 2004.
- 10 Rykaczewski, R. R. and Dunne, J. P.: Enhanced nutrient supply to the California Current Ecosystem with global warming and increased stratification in an earth system model, *Geophys. Res. Lett.*, 37, L21 606, <https://doi.org/10.1029/2010GL045019>, 2010.
- Ryther, J. H.: Photosynthesis and Fish Production in the Sea, *Science*, 166, 72–76, <https://doi.org/10.1126/science.166.3901.72>, 1969.
- SeaPAC: SeaWinds on QuikSCAT Level 3 Daily Gridded Ocean Wind Vectors (JPL Version 2). Ver. 2. PO.DAAC, CA, USA, Dataset accessed [2018-07-27] at <http://dx.doi.org/10.5067/QSXXX-L3002>, 2006.
- 15 Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, *Nature*, 484, 49—54, <https://doi.org/10.1038/nature10915>, 2012.
- Snyder, M. A., Sloan, L. C., Diffenbaugh, N. S., and Bell, J. L.: Future climate change and upwelling in the California Current, *Geophysical Research Letters*, 30, 2003.
- 20 van Geen, A., Zheng, Y., Bernhard, J. M., Cannariato, K. G., Carriquiry, J., Dean, W. E., Eakins, B. W., Ortiz, J. D., and Pike, J.: On the preservation of laminated sediments along the western margin of North America, *Paleoceanography*, 18, 1098, <https://doi.org/10.1029/2003PA000911>, 2003.
- Vargas, G., Pantoja, S., Rutllant, J. A., Lange, C. B., and Ortlieb, L.: Enhancement of coastal upwelling and interdecadal ENSO-like variability in the Peru-Chile Current since late 19th century, *Geophys. Res. Lett.*, 34, L13 607, <https://doi.org/10.1029/2006GL028812>, 2007.
- 25 Wang, J., Emile-Geay, J., Guillot, D., Smerdon, J. E., and Rajaratnam, B.: Evaluating climate field reconstruction techniques using improved emulations of real-world conditions, *Climate of the Past*, 10, 1–19, <https://doi.org/10.5194/cp-10-1-2014>, 2014.
- Ware, D. M. and Thomson, R. E.: Bottom-Up Ecosystem Trophic Dynamics Determine Fish Production in the Northeast Pacific, *Science*, 308, 1280–1284, <https://doi.org/10.1126/science.1109049>, 2005.
- Wick, G. A., Bates, J. J., and Scott, D. J.: Satellite and skin-layer effects on the accuracy of sea surface temperature measurements from the GOES satellites, *Journal of Atmospheric and Oceanic Technology*, 19, 1834–1848, 2002.
- 30 Xiu, P., Chai, F., Curchitser, E. N., and Castruccio, F. S.: Future changes in coastal upwelling ecosystems with global warming: The case of the California Current System, *Sci. Rep.*, 8, 2866, <https://doi.org/10.1038/s41598-018-21247-7>, 2018.
- Zhao, M., Eglinton, G., Read, G., and Schimmelmann, A.: An alkenone (U37K') quasi-annual sea surface temperature record (AD 1440 to 1940) using varved sediments from the Santa Barbara Basin, *Organic Geochemistry*, 31, 903–917, [https://doi.org/10.1016/S0146-6380\(00\)00034-6](https://doi.org/10.1016/S0146-6380(00)00034-6), 2000.