Challenges and Alternatives to Empirical Orthogonal Functions for Earth System Data

Christine A. Shields[1], Derek DeSantis[2], Alexandra Jonko[3], Jiwoo Lee[4], Julie M. Caron[1], Adam S. Phillips[1], John Fasullo[1]

[1] NSF National Center for Atmospheric Research, Boulder, CO 80302, USA

[2] Computer, Computational & Statistical Science Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

[3] Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

[4] Atmospheric, Earth, and Energy Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

Corresponding author: shields@ucar.edu

Abstract

Empirical Orthogonal Functions (EOFs) applied to gridded Earth system data enables users to diagnose modes of variability with relative ease. Yet, many challenges to interpretation exist such that they must be used with awareness and intention when applied to gridded climate data, especially with large ensembles. Utilizing data from two different Earth system modelling large ensemble frameworks, the Energy Exoscale Earth System Model (E3SM) and the Community Earth System Model (CESM), as well as reanalysis data, common EOF pitfalls are summarized and discussed. Challenges include erroneous mode swapping, sign flipping, and the temporal variability of the centers of action. For modes of variability with similar contribution to variance, mode swapping is not uncommon. Sign flipping can occur with almost any mode where the pattern is correct, but the sign is arbitrary. Although the variability of the center of action is not necessarily problematic, it potentially complicates interpretation over multi-century timescales. A wide variety of alternative methods to EOFs exist, but fitness-for-purpose must be evaluated. Additionally, illustrations of alternative methods and examples of proper use are provided. Alternative methods fit into three categories: EOF variants, linear methods, and multilinear methods are provided.

Key Words/Phrases

Climate data, Earth system models, EOFs, PCA, Modes of variability

Introduction

Principal Component Analysis (PCA) is a widely used method for assessing gridded climate datasets. Whether the data takes the form of reanalyses, observational products, or model output, understanding Earth’s climate and its variability often involves quantifying patterns of variability, disentangling them from external forcing such as anthropogenic climate change, or simplifying predictions with the use of linear inverse modelling[1]. PCA is a commonly used approach which decomposes data into (1) spatial patterns, referred to as Empirical Orthogonal Functions (EOFs) and (2) a time-varying principal components (PC), representing the variation of the amplitudes of the EOFs over time. Although patterns produced from geophysical data can be interpreted as physical variance patterns occurring naturally in the Earth system, the data structures themselves are distinct from real physical processes. Patterns are defined to be orthogonal and sorted by contributions to total variance, in a specified domain, temporally across the dataset. EOF modes are uncorrelated with each other, in that the different patterns of variance may not necessarily be related to each other[2,3,4,5,6]. For example, a very common internal mode of climate variability in the Northern Hemisphere is the Pacific North American Pattern (PNA), a storm track pattern. In observations, the PNA is the leading EOF of interannual sea -level pressure (PSL) anomalies. The second leading mode is the NPO (North Pacific Oscillation), an orthogonal pattern oriented as a meridional dipole in contrast to the PNA, whose structure is primarily zonal. Reproducing these patterns in observations is influenced by parametric choices such as time period and domain bounds. Reproducing patterns in ensembles of climate models can be complicated further by orthogonal modes with similar variance across the dataset, such that the ordering of the modes is inconsistent across individual simulations. This is called mode-swapping. It has been documented in a number of papers[5,7,8,9,] and can have serious implications for interpretability of the dominant mode in both models and observations. The purpose of this article is to, in one place, examine common pitfalls in PCA analysis when applying EOFs and outline approaches for dealing with them. Our intent is *not* to describe dominant modes of variability, nor to provide a deep-dive into PCA mathematics, but rather to utilize the novel opportunity provided by large and multi-model ensembles to explore statistical aspects of EOF analysis. While EOF analysis can also be applied to spatially non-gridded data, here, we limit our illustrations to spatially gridded climate data. We acknowledge that some of these pitfalls are simply natural characteristics of the method, rather than problems per se, however, challenges do arise when using EOFs to evaluate physical patterns, and particularly multi-model comparisons and large ensemble studies. Throughout this article, we intertwine discussion of EOFs, their modes, and respective physical interpretations, all together, to frame PCA within the context of typical community practices. However, it is important to acknowledge again that the data structures themselves are indeed distinct from the physical patterns, and it is our interpretation of them that is the challenge. Further discussion can be found in Methods under *Techniques: Standard EOF Computation*. We follow in the spirit of earlier literature, each aimed at addressing specific aspects of the challenge [5,8]. Lee et al. (2019)[8] proposed applying common basis functions (CBFs) to EOFs as a means of using observations to mitigate for pitfalls of PCA in evaluating climate models. CBFs can correct for mode swapping, sign swapping (i.e, the arbitrarily assigned signs of EOFs that are inconsistent with observations), and potentially misinterpretation related to mode center of action variability. Although the application of CBF works well for historical simulations, it is not well-suited for cases where the patterns of modes change, such as in past or future climates. Techniques considered to be an alternative to EOFs and traditional PCA analysis include rotated EOFs (REOFs). Rotated EOFs can address issues related to the production of potentially unphysical modes and mode mixing, and have been explored in earlier work[10,11]. Even using REOFs, however, some of the classic problems, such as interpretability and geophysical pattern reliance on domain choice, still remain[12,13]. Many issues arise when analyzing climate model simulations, especially in large ensembles, which consist of multiple simulations (>20) initialized with slightly different conditions but using identical external climate forcings. Over the past decade, large ensembles have emerged as a critical and necessary tool for disentangling natural and forced climate signals[14,15], in part, by benchmarking modes of variability (MOV) internal noise[16,17,18,19]. Here, we coalesce the most common challenges found in large ensembles into one reference paper and discuss not only the solutions and potential alternatives, but also the purposes and interpretability of these methods. EOFs are powerful tools, however their limitations need to be acknowledged when applied to climate data.

Challenges with EOFs

Mode swap

Sequential EOFs for a given variable are ranked by the percentage of variance they explain. When two modes explain very similar amounts of variance, mode swapping can occur, i.e, a physical mode gets paired with a neighboring EOF rather than its canonical pattern. We define mode swapping as the misassignment of a mode of variability to an adjacent EOF. Mode swapping shows up whenever EOFs are not clearly distinguishable, for example by the North's test for Rule of Thumb for EOF significance[3] which evaluates separability by estimating the sampling uncertainty in the associated eigenvalues.

To illustrate with a large ensemble framework, consider wintertime, January-March (JFM) sea level pressure over the North Atlantic Region (20:80°N, -90:40°E). In the ERA-5 reanalysis, EOF2 (East Atlantic Pattern, EA) accounts for 14.9% of variance and EOF3 (Scandinavian pattern, SCA) for 11.5%, only a 3.4% difference. If another dataset, or model shifts the gap by ≧ 3.4%, EOF2 and EOF3 swap: the second EOF now looks SCA-like and the third EA-like. A similar swap can also happen between EOF3 and EOF4, pushing SCA into EOF4. Figure 1 shows this in two CESM2 (Community Earth System Model version 2) CMIP6 ensemble members. In r1i1p1f1, EOF2 matches the EA pattern with the larger variance (upper middle compared to upper left panels). In r3i1p1f1, the larger variance aligns with SCA rather than EA (compare upper right with lower left).

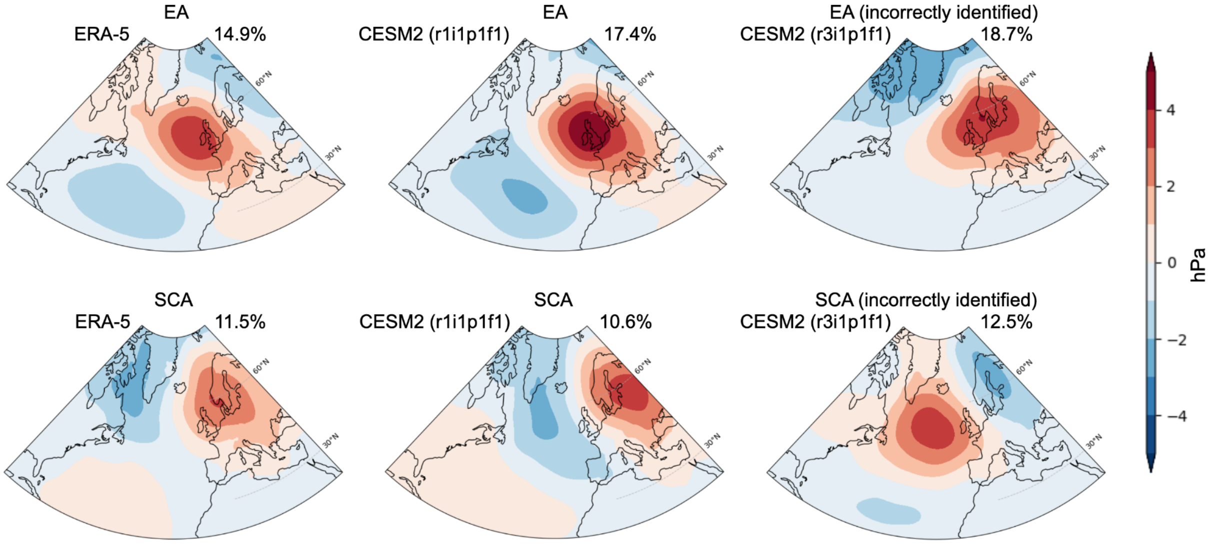


Figure 1: East Atlantic pattern (EA, row 1) and Scandinavian pattern (SCA, row 2) shown for ERA5 (first column), CESM2 ScenarioMIP ensemble member r1i1p1f1 (second column) and CESM2 ScenarioMIP ensemble member r3i1p1f1 (third column). Mode swapping is evident in the third column. EA and SCA are defined as the 2nd and 3rd EOF patterns of area-weighted PSL computed over 20:80°N, -90:40°E for JFM 1979-2022. Units are in hPa and variance explained is listed at the top right of each panel.

The same phenomenon occurs in the Southern Hemisphere with the Pacific South American patterns (PSA1 and PSA2, Figure 2). Using sea level pressure for the winter months of June through August (JJA) in the southern hemisphere, ERA5 reanalysis quantifies these modes as even more closely varying with a mode separation of only 1%. Not surprisingly, when using ensemble climate model data, mode swapping occurs frequently across ensemble members. As in Figure 1, the middle panels of Figure 2 show a CESM ensemble member (1181.010) for which modes match the expected patterns compared to ERA5, and the PSA1 and PSA2 correspond to EOF2 and EOF3, respectively. The panels on the right show a different ensemble member (1161.009) where the modes have been swapped. Here PSA1 corresponds to EOF3, and PSA2 to EOF2. Mode swapping amongst different realizations of the same model can have serious implications when diagnosing teleconnections and their respective regional weather impacts. To demonstrate potential erroneous impacts and illustrate the ubiquitousness of mode swapping in ensembles, we provide an example of mode swapping frequency between PSA1 and PSA2 in Table1 using the CMP6 archive, with and without the removal of the forced response (defined as the removal of the ensemble mean). Percentage metrics are highly dependent on sampling. For example, if the full CESM and E3SM large ensembles are used, instead of the CMIP6 archive subset, the percentage statistics change such that E3SM1,2 mode swaps with the forced response occur 88% and 86% of the time respectively, and CESM1,2 18% and 56%. With the forced response removed these metrics are 53%, 48%, 33%, 44% respectively.

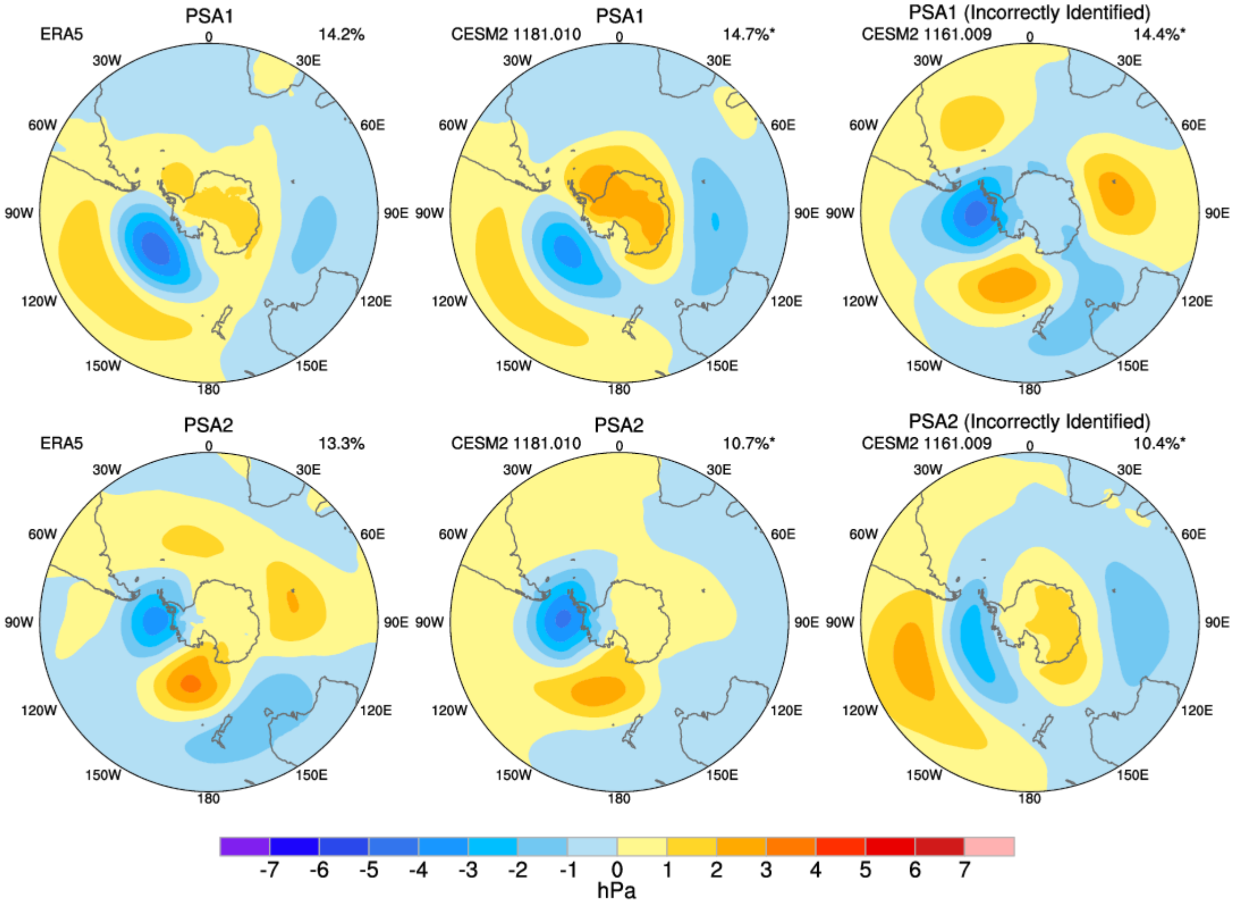
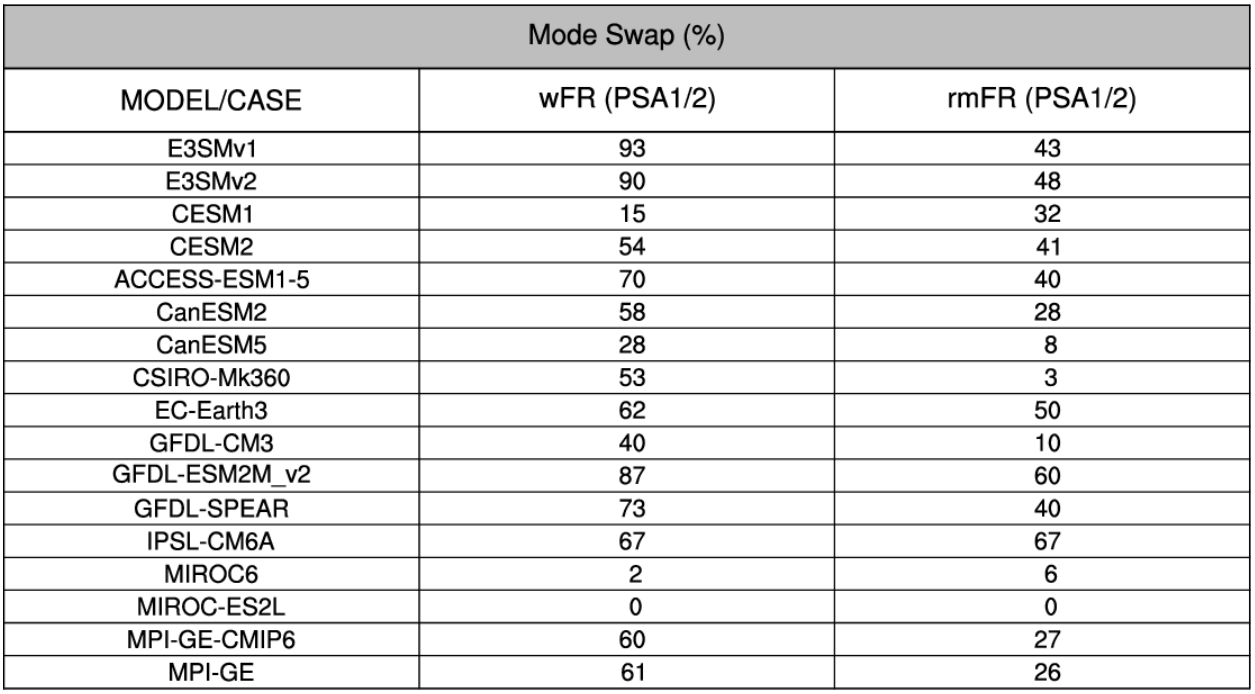


Figure 2: Pacific South American (PSA) modes 1 (row 1) and 2 (row 2) shown for ERA5 (first column), CESM2 Large Ensemble member 1181.010 (second column) and CESM2 Large Ensemble member 1161.009 (third column). Mode swapping is evident in the third column. PSA1 and PSA2 are defined as the 2nd and 3rd EOF patterns of area-weighted PSL computed over 20:90°S, 0:360°E for June-August 1950-2023. Units are in hPa and variance explained is listed at the top right of each panel. The patterns are created by regressing global PSL anomalies onto normalized PC timeseries.



Mode Swap (%)		
MODEL/CASE	wFR (PSA1/2)	rmFR (PSA1/2)
E3SMv1	93	43
E3SMv2	90	48
CESM1	15	32
CESM2	54	41
ACCESS-ESM1-5	70	40
CanESM2	58	28
CanESM5	28	8
CSIRO-Mk360	53	3
EC-Earth3	62	50
GFDL-CM3	40	10
GFDL-ESM2M_v2	87	60
GFDL-SPEAR	73	40
IPSL-CM6A	67	67
MIROC6	2	6
MIROC-ES2L	0	0
MPI-GE-CMIP6	60	27
MPI-GE	61	26

Table 1. Percentage of time mode swapping occurs between PSA1 and PSA2 for JJA (Southern Hemisphere winter) for the period of 1950-2022 using the CMIP6 archive. The left column names the model, the middle column shows statistics with the forces response included (wFR) and the right column shows statistics with the forced response removed (rmFR). The forced response is computed as the ensemble mean sea level pressure (SLP), which is removed from the SLP field prior to calculation of the PSA1 and 2 modes for each fully coupled ensemble. More detail in Methods. Note percentage metrics are highly dependent on method of sampling and should not be used as exact characteristics of modelling frameworks, rather as a demonstration of the depth of the mode swapping challenge for researchers.

Sign flip

Sign flipping is another common feature and can occur with almost any mode of variability. In EOF analysis, the sign of each EOF mode is arbitrary, due to the symmetric characteristics of eigenvalue decomposition. Flipping the sign of both the spatial pattern and its PC time series does not impact their interpretation or statistical significance. However, when comparing across models or simulations, inconsistencies in sign conventions may cause visual discrepancies and introduce challenges in statistical comparison, thus we include it here. To illustrate this point, we plot the Pacific North American (PNA) pattern (Figure 3) using the Energy Exoscale Earth System Model version 2 (E3SMv2) large ensemble as defined by the first EOF pattern of area-weighted PSL computed over 20:85°N, 120°E:120°W for June-August 1950-2022. The hemispheric pattern is created by regressing global PSL anomalies onto the normalized PC timeseries. Knowing the correct phasing (positive or negative) for the mode based on observations allows us to identify and correct sign flipping issues.

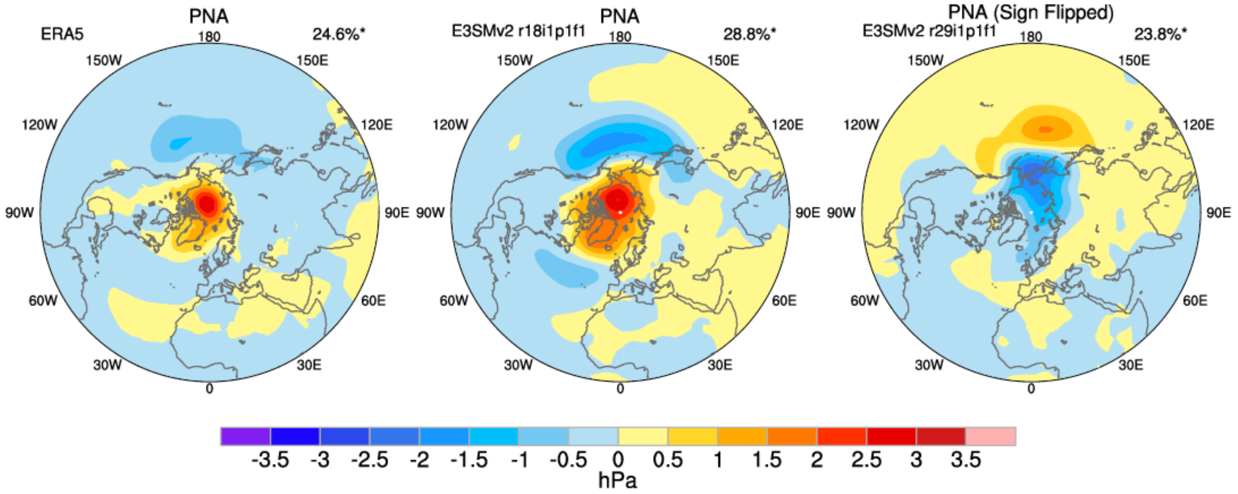


Figure 3: Pacific North American pattern shown for ERA5 (first panel), E3SMv2 member r18i1p1f1 (second panel) and E3SMv2 member r29i1p1f1 (third panel). Sign flipping is evident in the third panel. Units are in hPa and variance explained for each pattern is listed at the top right. The PNA is defined as the first EOF pattern of area-weighted PSL computed over 20:85°N, 120°E:120°W for June-August 1950-2022. Units are in hPa and variance explained is listed at the top right of each panel. The patterns are created by regressing global PSL anomalies onto normalized PC timeseries.

Center of action variability

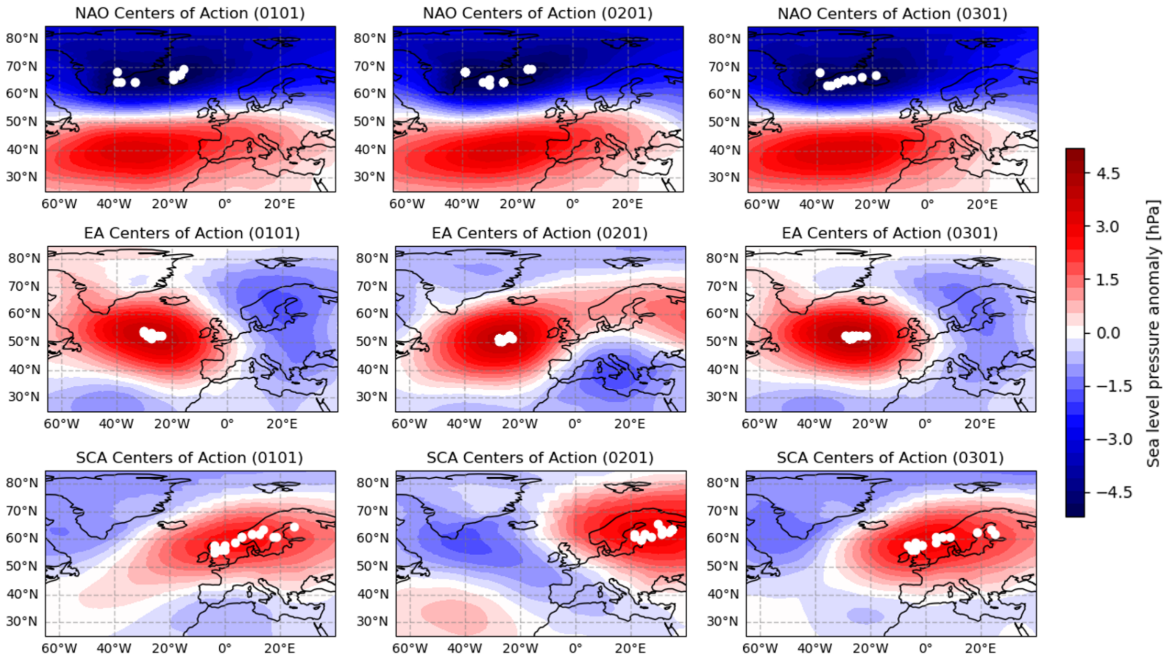


Figure 4: Center of action variability for North Atlantic Oscillation (top), East Atlantic (middle) and Scandinavian (top) patterns computed for three ensemble members of the E3SMv2 Large Ensemble: 0101 (left), 0201 (center), and 0301 (right), for a combination of historical and SSP370 simulations. The NAO is defined as the first EOF, the East Atlantic pattern as the second EOF and the Scandinavian pattern as the third EOF of area-weighted PSL computed over 20:80°N, 90°W:40°E for January-March over 100-year periods staggered by 10 years. The mean EOFs are shown, as well as the centers of action for each of the sixteen 100-year periods between 1850 and 2100, marked by white dots.

While center-of-action variability is not a flaw of PCA, it does highlight an important limitation of fixed-pattern decompositions: PCA assumes temporally invariant spatial structures. Shifts in the location of the center-of-action, defined as the geographic location of the maxima and minima associated with an EOF pattern, do occur, and can be used to evaluate multi-model, ensemble, or temporal variability. We use temporal variability to illustrate the concept. EOFs are typically calculated from a single time interval, implicitly treating mode structure as temporally invariant. However, this is not always the case, especially when EOFs are applied to data generated in a changing climate. While data can be detrended prior to computation of the EOFs, this is typically done by fitting simple trends and can miss more complex climate change signals. When climate modes exhibit shifts in their centers of action, these shifts may be represented as amplitude changes or rotations among neighboring EOFs, complicating interpretation.

To investigate temporal center-of-action stability, we calculated EOFs from successive 100-year segments of simulations spanning historical and SSP370 scenarios from 1850 to 2100 for three members of the E3SMv2 Large Ensemble (0101, 0201, and 0301). The start date of each 100-year segment is advanced by ten years between windows, resulting in sixteen time slices for each ensemble member. Figure 4 shows the centers of action for the leading three EOFs of PSL, corresponding to the North Atlantic Oscillation (NAO), EA, and SCA, respectively. The loading patterns are averaged over the sixteen time slices and shown as contours in the background. Centers of action are marked as white dots. If EOF patterns showed no temporal variability, all white dots would cluster tightly over the maximum of the average loading pattern. However, we see that they trace out a broad region of variability, especially for the NAO and SCA patterns (Figure 4). This scatter demonstrates that even the leading modes of variability exhibit significant non-stationarity over multi-century timescales. To quantify whether two spatial patterns can be considered the “same” climate mode in the presence of sampling uncertainty - a spatial analog to the North (1982)[3] eigenvalue criterion - we computed the pairwise pattern-correlation distribution across all time-slices. For a perfectly stable mode, correlations would cluster tightly near 1. While this holds for the NAO, with 88%, 88%, and 100% of r values above 0.8 for ensemble members 0101, 0201, and 0301, respectively, we find a wide, often near-uniform spread from 0 to 1 for EA and SCA. (EA: 58%, 96%, and 53%, SCA: 44%, 34%, and 30%). These results indicate that higher-order EOFs exhibit far lower robustness, consistent with rotational degeneracy and sensitivity to non-stationary forcing. Overall, this demonstrates that spatial shifts in mode structure constitute a significant, and often underappreciated source of uncertainty. When EOFs derived from historical or control simulations are projected onto future climate scenarios—especially under strong anthropogenic forcing—these shifts are likely to accelerate, potentially biasing our interpretation of circulation changes and feedbacks.

Direct adjustment methods

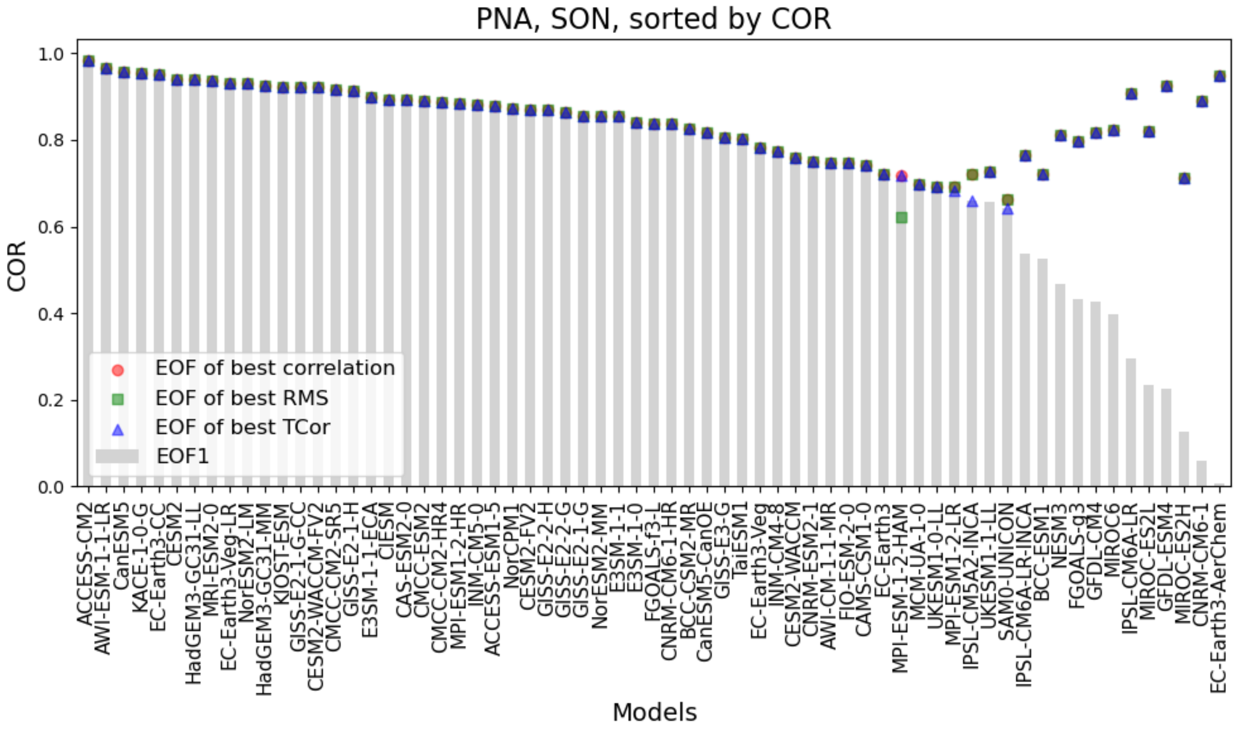
Mode swap

Mode swapping complicates both the physical interpretation and statistical comparison of EOFs, and addressing it is essential for studies of climate variability given physical patterns and data structures are used in tandem. Often, researchers resort to manual mode swapping to ensure consistency in mode identification. While this approach can be effective for small datasets, it is tedious, subjective, and impractical for large ensembles or multi-model comparisons.

Recent efforts have focused on automating the mode-swapping process, using objective criteria to match modes across datasets. One such technique involves statistical matching, where EOF modes from a target dataset (i.e., a model simulation) are compared to those from a reference dataset using similarity metrics such as spatial correlation, root-mean-square error (RMSE), or temporal correlation of the EOF PC time series to the Common Basis Functions (CBF)[8] (Table 2 and Methods) derived PC time series. Methods such as CBFs have been developed to address some of the limitations of traditional EOFs. Such similarity metrics allow the user to quantify the differences between the mode in question and a reference control. For example, each EOF mode from the target dataset (e.g., EOF 1, EOF 2, EOF 3) is compared against a specific reference mode (e.g., REF EOF 1), and the mode with the highest correlation or lowest RMSE is selected as the best match. We considered the first three EOF modes after conducting the North Test[3] to examine robustness of each order of EOF modes, where we observed a significant decrease of the percentage of explained variability for EOF modes higher than the third in most cases. In some cases, the reference PC time series used for temporal correlation may be derived from the CBF, particularly when day-to-day or year-to-year reproducibility is not expected—for instance, in simulations following the

Historical or AMIP experiment protocols of CMIP. While this approach can automatically resolve most cases of EOF mode swapping (see Fig. 5, upper panel), it is not always applicable in a consistent way —particularly when different similarity metrics identify different best-matching modes. Figure 5 (bottom panel) illustrates this for the PNA pattern during SON, which is one of the modes and seasons with the highest incidence of EOF swapping [8]. A few models display non-overlapping markers, indicating disagreement in mode identification and necessitating manual inspection and mode swapping.

CBF approaches emphasize dynamical consistency and interpretability, reducing ambiguity in mode identification by using dominant modes from reference datasets as fixed bases to detect similar patterns in model output—thereby eliminating the need for post-hoc mode swapping.



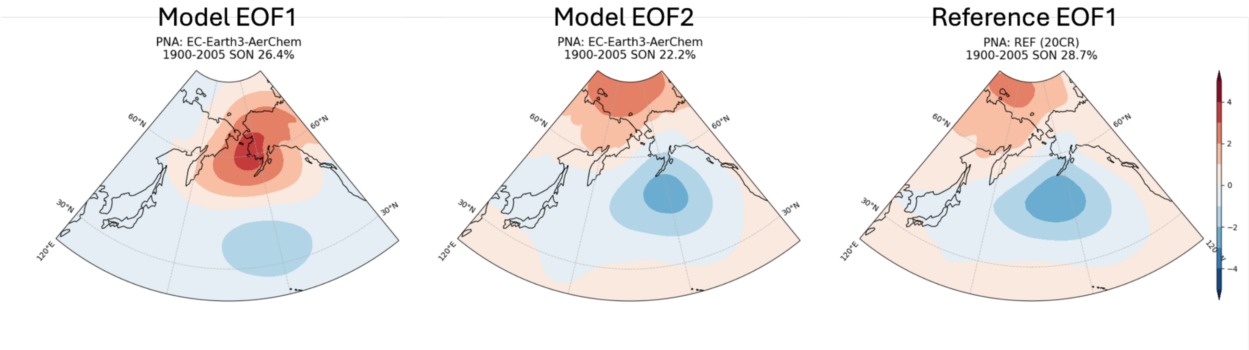


Figure 5. Mode swapping pattern correlation adjustment method illustration. (Upper panel) Pattern correlation of CMIP6 models' Pacific/North American (PNA) pattern during the September-October-November (SON) season (upper panel). The PNA pattern for each model was determined by the leading Empirical Orthogonal Function (EOF 1) of sea level pressure fields. Pattern correlation was then calculated between each model's EOF 1 spatial pattern and the 20th Century Reanalysis (20CR) PNA pattern (SON season). For each model, the pattern correlation of the leading three EOFs (EOF 1-3) was assessed against the 20CR PNA using spatial pattern correlation, Root Mean Square Error (RMSE), and temporal correlation between the EOF principal component (PC) time series and a Common Basis Function PC time series. The highest pattern correlation achieved after potentially swapping EOF modes based on these criteria is indicated by markers, while the pattern correlation of the original EOF 1 is shown as a gray bar. The increasing spread between EOF1 and the alternative EOF on the right demonstrates the difficulty in automated selection of the best-matching EOF mode for models on the right, which highlights challenges in systematically identifying robust climate patterns across different models (further details in Lee et al., 2019[8]). (Lower panel) Spatial pattern of model's EOF 1 (left), 2 (middle) and reference dataset's EOF 1 demonstrating an example EOF swap case, obtained from EC-Earth3-AerChem model (far-most right in the upper panel) and the 20CR for PNA pattern during the SON season. This figure is a modified and expanded version from Figure 6 in Lee et al., 2019[8].

Sign Flip

Evaluating the EOF sign is required to ensure consistency when evaluating or intercomparing model simulations. A common manual approach involves checking the sign of each EOF spatial pattern or PC time series against a reference and flipping it if necessary—again, a process that becomes impractical with large datasets or ensemble analyses. To streamline this, automated sign correction techniques have been applied, which typically rely on pattern correlation: the EOF from a target dataset is compared to a reference pattern, and if the correlation is positive, the sign is retained; if negative, the sign is flipped. This approach mirrors the strategy used in automated mode matching and provides a consistent, objective way to address sign ambiguity. In addition to the correlation-based technique, region-specific sign corrections provide a practical alternative and/or complementary method for ensuring EOF sign consistency, particularly for specific physical modes of variability and their spatial signatures. This method calculates the mean of the EOF pattern over a targeted geographical region and flips the sign if it deviates from the expected convention. For example, the Pacific Decadal Oscillation (PDO) often shows negative loading in the central North Pacific, while the Southern Annular Mode (SAM) is typically negative over the Southern Ocean. Applying this check across predefined regions allows consistent alignment with standard mode definitions and can be especially useful when correlation metrics yield ambiguous results.

While both mode swapping and sign flipping aim to ensure coherent EOF representation across datasets, they address distinct issues. Mode swapping corrects for *ordering ambiguities when eigenvalues are close* or degenerate, noting that in linear algebra, degeneracy occurs when an eigenvalue (the frequency) can be associated with more than one eigenvector (the mode), and also taking into consideration that EOFs represent a continuum of modes and spatial patterns[20]. Sign flipping, however, corrects for *orientation ambiguities of individual modes*. Both corrections are essential for creating comparable EOF structures that are physically interpretable across observations, models, ensemble members, or experimental protocols.

Alternative methods

The limitations and challenges associated with EOF analysis highlighted thus far arise because PCA imposes several constraints—most notably linearity, orthogonality, and an implicit assumption of normally distributed fluctuations of the climate system. These limitations can limit PCA's effectiveness in capturing the full complexity of climate dynamics and feedbacks . Climate datasets frequently exhibit nonlinear interactions, spatial and temporal dependencies, raising questions about these assumptions. Furthermore, climate datasets often contain inherent physical constraints such as non-negativity (e.g., precipitation, concentrations, volumes). Standard EOF analysis does not enforce such constraints: its linear combinations can easily reconstruct physical fields with negative values of intrinsically non-negative variables. Working with anomalies does not avoid this issue. Once the climatological mean is added back, the EOF reconstruction can still violate physical bounds. As a result, EOF-derived modes may capture mathematically valid variability patterns while nonetheless producing physically impossible states, complicating interpretation and downstream use.[21]

Motivated by these limitations, there is a need to explore alternative decomposition methods capable of addressing more general structures inherent to climate data. This section introduces several powerful approaches beyond PCA with further details in the Methods Section. We categorize these methods based on a few criteria and provide examples applied to climate questions:

- **EOF Variants:** These techniques build off the core idea of EOFs, providing slight tweaks to try and overcome one or more limitations. Examples include common basis functions, rotated EOFs, and sparse EOFs.

- **Linear Methods:** EOFs are a linear technique to define and discover modes, and as such, belong to a larger class of linear methods. Other examples include factor analysis (FA), independent component analysis, nonnegative matrix factorization (NMF), and Dynamic Mode Decomposition (DMD).
- **Multilinear Methods:** EOFs are discovered by flattening the spatial axis into a single dimension. Multilinear methods, such as the Canonical Polyadic Decomposition or the Tucker Decomposition generalize PCA to higher dimensional arrays in order to discover multidimensional modes.

Table 2 summarizes our categorization. Each approach offers unique strengths in handling interacting relationships, temporal dynamics, and spatial topology, thus providing climate researchers with versatile tools better suited for capturing and interpreting complex patterns embedded in climate datasets. Note that this categorization is far from complete. We have chosen a small, but representative list of mode extraction methods, many of which have been applied throughout Earth system science. All the methods described here arise from linear analysis - we have chosen not to include nonlinear methods such as autoencoders[22] to keep the discussion more contained. Moreover, the above methods can be combined (e.g. nonnegative Tucker Decomposition, Table 2) to add further layers of sophistication. Here, we directly compare EOFs to the methods in Table 2. In Methods, we provide an overview of the techniques and discuss which issues each method addresses.

Method	Addresses	Limitation	Category	Example Uses	Packages
EOF	Maximizes variance with orthogonal modes	Modes may mix physical signals	EOF	Gridded climate data	Many analysis software packages will have an EOF routine, but, e.g. pyEOF, CVDP, PMP.

					or DOI: 10.5334/jors.122
CBF	Shared spatial basis across datasets	May obscure dataset-specific structure	EOF Variant	Multiple models or observations in a unified spatial framework	PMP
Rotated EOF	Localized, interpretable patterns	Rotation is subjective; loses orthogonality	EOF Variant	Avoid the unphysical dipole like EOF analysis pattern	pyEOF
Sparse EOF	Enhances interpretability via sparsity	Requires tuning; sensitive to noise	EOF Variant	Fingerprinting	scikit-learn
FA	Models shared + unique variance (latent factors)	Identifiability - Different modes may explain data	Linear Method	Latent dynamics (e.g., unobserved climate drivers) are suspected to govern the observed data	scikit-learn
ICA	Finds statistically independent components	Source signals must be non-Gaussian	Linear Method	Investigate sea level pressure and water storage	scikit-learn
NMF	Nonnegative, parts-based decomposition	Non-unique; depends on initialization	Linear Method	Precipitation, cloud cover, energy fluxes; characterizing drought behavior in river basins and quantifying the sources of atmospheric particles	scikit-learn
DMD	Extracts coherent	Assumes linear	Linear Method	Traveling waves, oscillations, and	pydmd

	dynamical modes	dynamics; sensitive to noise		instabilities in geophysical flows and climate systems; LIMs	
CP Decomposition	Multivariate generalization of EOF	Non-orthogonal; difficult optimization and scaling	Multilinear Method	Application that preserves multidimensional information (e.g. e.g., spatial modes, temporal modes, ensemble modes)	tensorly
Tucker Decomposition	Multi-mode compression with core interactions	Interpretation can be nuanced; requires mixing tensor	Multilinear Method	Application that preserves multidimensional information (e.g., e.g., spatial modes, temporal modes, ensemble modes)	tensorly

Table 2. Summary of fixing or alternate techniques for EOF issues, including available software packages addressing issues, limitations, and example uses. References and detailed descriptions of specific methods and terms are found in the Methods section.

Alternative Method Comparisons

We now compare and contrast a sample of each method category (EOF variant, linear, multilinear) with EOF analysis. Examples for other methods can be found in the Appendix. The aim of this section is not to extract deep insights about the Earth system, but to illustrate how these tools can be applied in practice and to highlight their key differences. We emphasize that our goal is to demonstrate the types of additional information that can be obtained using alternative techniques. This comparison is not intended to declare any single method as superior. In fact, such a direct comparison is often inappropriate, as each method is designed to capture and extract different features of the data. Each method serves a distinct purpose and offers a unique perspective.

Only standard EOF analysis provides a clear and interpretable ordering of modes by importance, with each mode corresponding to a descending eigenvalue and associated variance explained. The CBF method (see Figure 5) inherits this property through its shared EOF basis, allowing joint variance-based ranking. Rotated EOFs, however, lose this feature: although the total variance is preserved, the variance is redistributed among rotated modes, and no canonical ordering remains. Sparse EOFs may allow modes to be ranked by post hoc variance explained, but the presence of sparsity-inducing penalties complicates the interpretation of such rankings, and ordering often depends on external criteria such as cross-validation. In FA, modes can sometimes be heuristically ranked using loadings or commonalities; however, this ranking is neither unique nor necessarily stable across model specifications. ICA offers no natural mode ordering, as its objective is statistical independence rather than variance maximization. Similarly, NMF yields additive components without orthogonality or ordering, and the contribution of each mode must be assessed in context. DMD modes can be ordered by growth or decay rate, oscillation frequency, or modal energy, depending on the application. CPD lacks a natural ordering entirely, as components are not orthogonal and no scalar criterion governs their contribution. Tucker decomposition offers a partial analogue of variance-based ordering within each mode through the singular values of the factor matrices, but the presence of a dense core tensor, which mixes contributions across modes, precludes a globally interpretable ranking of importance. Because most alternatives lack a coherent mode ordering, spatial and temporal structures are not directly comparable. Nonetheless, we highlight representative modes from a sample from each category (EOF variant, linear method, multilinear method, Table 2) based on the considerations outlined above. Examples for methods in Table 2 not shown in the main text, are in the Appendix, for reference.

For our baseline, we compare against the sea level pressure (PSL) EOF modes from a single run of the CESM2 large ensemble monthly historical output. For this EOF baseline, even though it is standard practice to remove the seasonal cycle before computing EOFs because it often dominates the variance, we have chosen to not remove the seasonal cycle, nor apply any type of temporal filtering which can drastically change EOF results. We have made this decision for illustrative purposes so that we have a consistent baseline against other methods. Note that removing the seasonal cycle helps focus the EOF decomposition on interannual and longer-term variability rather than the predictable annual oscillation. Examples of differences due to temporal filtering can be found in the Methods section under T*echniques: Standard EOF Computation*. Figure 6 contains the top four EOF modes sorted by explained variance.

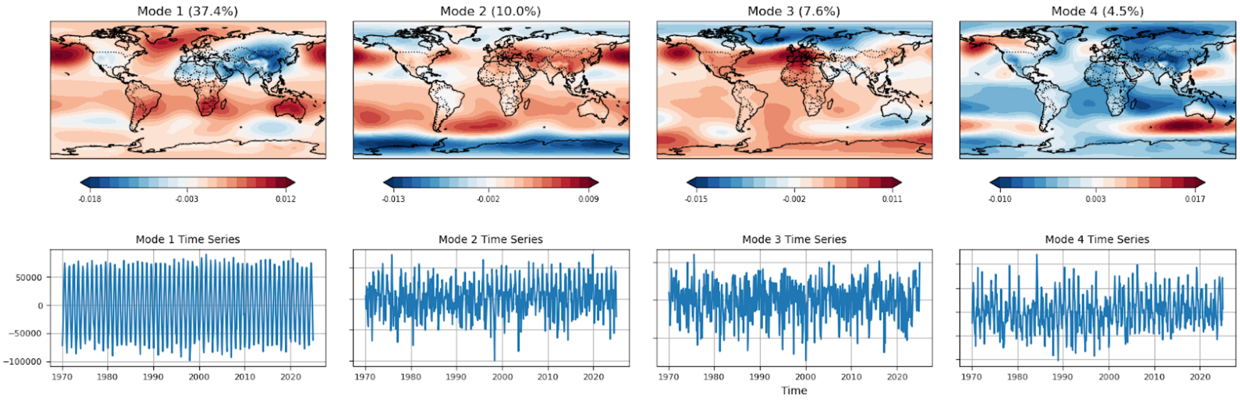


Figure 6. PSL EOFs 1,2,3,4 (with percent variance explained) and respective PC timeseries for CESM2 model large ensemble, first member. No temporal filtering is done for consistency and comparison across alternative methods. See the Methods section under *EOF: Standard Computation* for an illustration on the potentially large differences due to temporal filtering.

Rotated EOFs (EOF Variant)

The rotated EOFs produce results that are broadly similar to those from standard EOF analysis (Figure 7). The first three modes largely mirror the original EOFs, aside from sign changes and minor variations in spatial structure. The most notable deviation occurs in the fourth mode, where the spatial pattern differs indicating that rotation has reoriented the variance into a distinct structure not present in the unrotated EOFs.

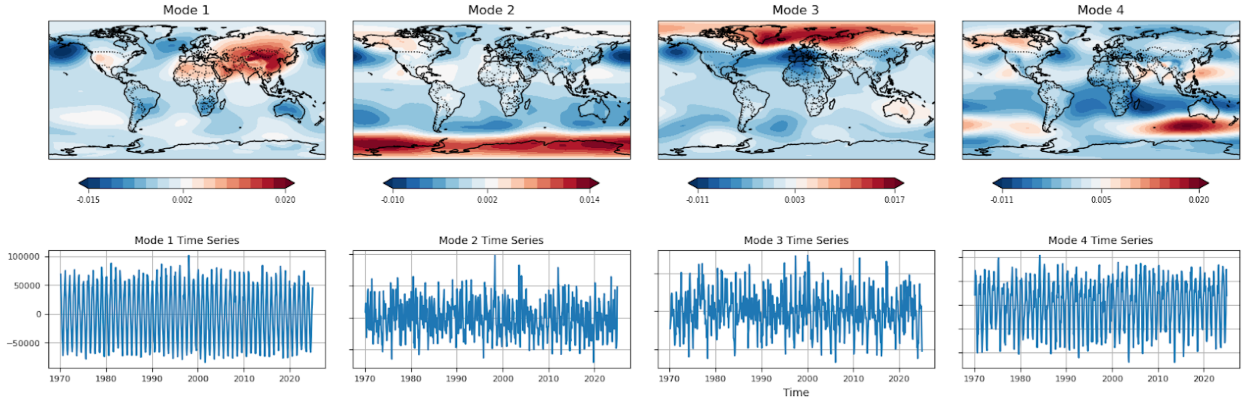


Figure 7. As in Figure 6, but an example of rotated EOF.

Factor Analysis (Linear Method)

FA results begin to diverge more noticeably from those of standard EOFs (Figure 8). A mode closely resembling the leading EOF still emerges, but subsequent FA modes appear as mixtures of multiple EOF patterns. For example, FA mode 2 resembles a combination of EOFs 1 through 3, while FA mode 3 aligns primarily with EOF 1 but includes features of EOF 4. FA mode 4 appears to blend characteristics of EOFs 2 and 3.

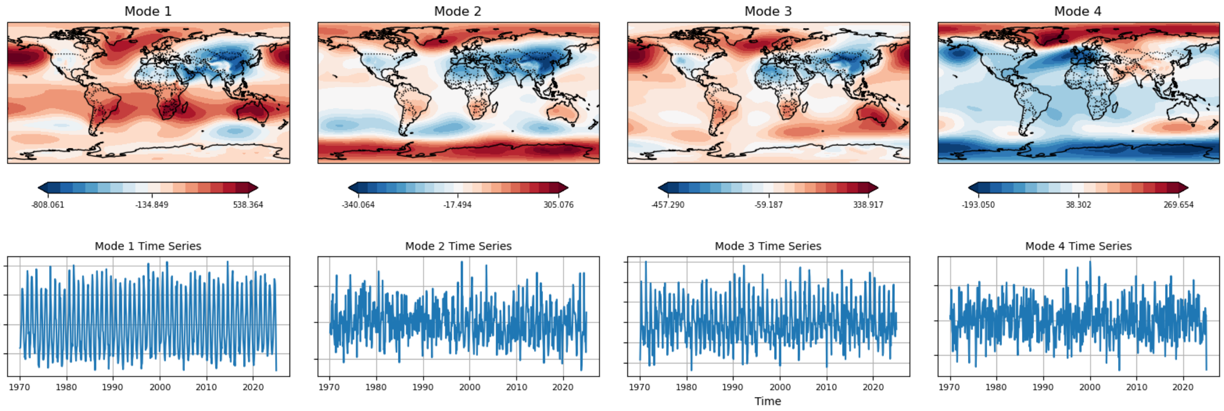


Figure 8. As in Figure 6 but using Factor Analysis.

Tucker Decomposition (Multilinear Method)

The Tucker decomposition is a multivariate generalization of the singular value decomposition (SVD) that extracts separate sets of modes for latitude, longitude, depth, etc. In practice, it works by finding a dominant subspace just like SVD but for each dimension individually which gives modes. Tucker also discovers the linear mixing of these modes in a small "core" tensor. These modes can be weighted and combined to reconstruct the full spatiotemporal field. Tucker allows independent control over the number of modes (ranks) in each dimension. For example, with a multirank of (5, 10, 10) - i.e., 5 temporal, 10 latitude, and 10 longitude modes - the core tensor has 5 temporal slices, each describing how to combine the 10 latitude and 10 longitude modes to form spatial patterns associated with a given temporal mode (Figure 9).

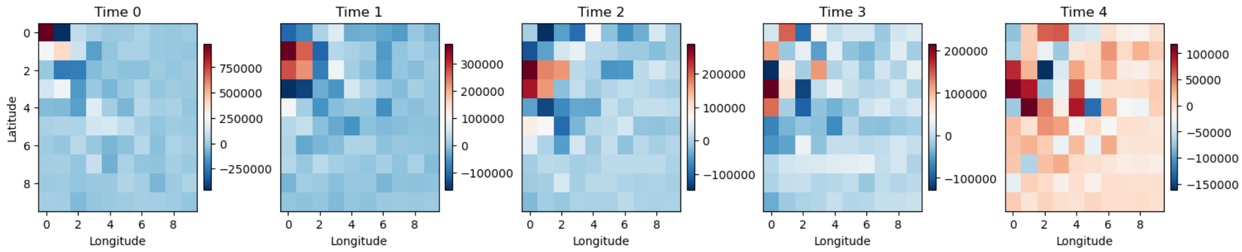


Figure 9. Tucker decomposition core tensor of PSL CESM2 large ensemble member in Figure 6, with a multirank of (5, 10, 10), with temporal modes 0 through 4, 10 latitudes (y-axis) and 10 longitudes (x-axis).

By identifying the largest weights in the core tensor, we can isolate the most dominant spatial structures associated with each temporal mode. Using this approach, we construct the leading spatial modes for the first four temporal components (Figure 10).

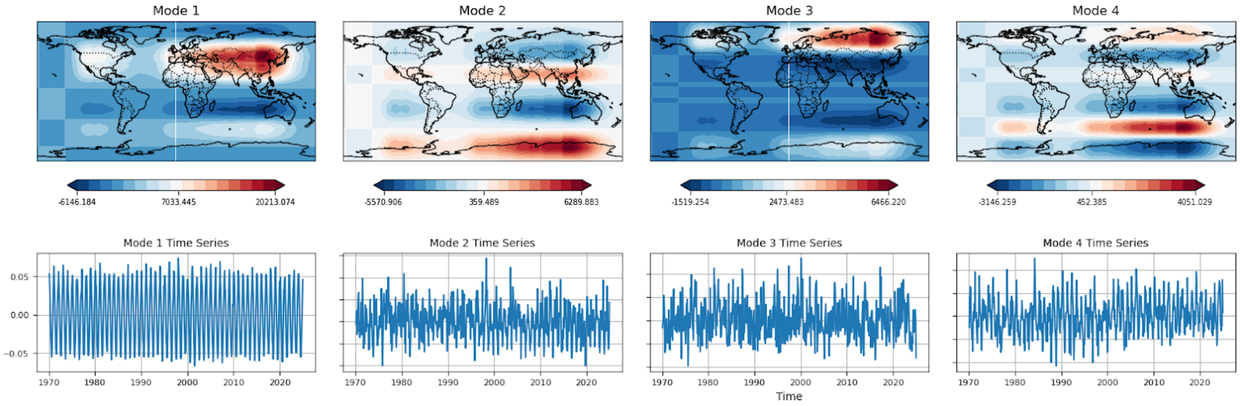


Figure 10. As in Figure 6, but for Tucker decomposition for the first four temporal components using the largest weights.

Importantly, the core tensor enables flexible linear combinations of spatial and temporal modes, offering richer analysis. This reconstruction is obtained by contracting the core tensor with the factor matrices, equivalently summing over rank indices the weighted outer products of the corresponding mode vectors, with the contraction weights given by the entries of the core tensor. This multilinear contraction preserves interactions across modes and reduces to standard matrix multiplication in the two-way (SVD) case.

When all core weights are used in combination (Figure 11), the reconstruction recovers the baseline EOF structure (up to potential sign differences), demonstrating that Tucker generalizes EOFs while retaining interpretability through its separable and multi-ranked structure.

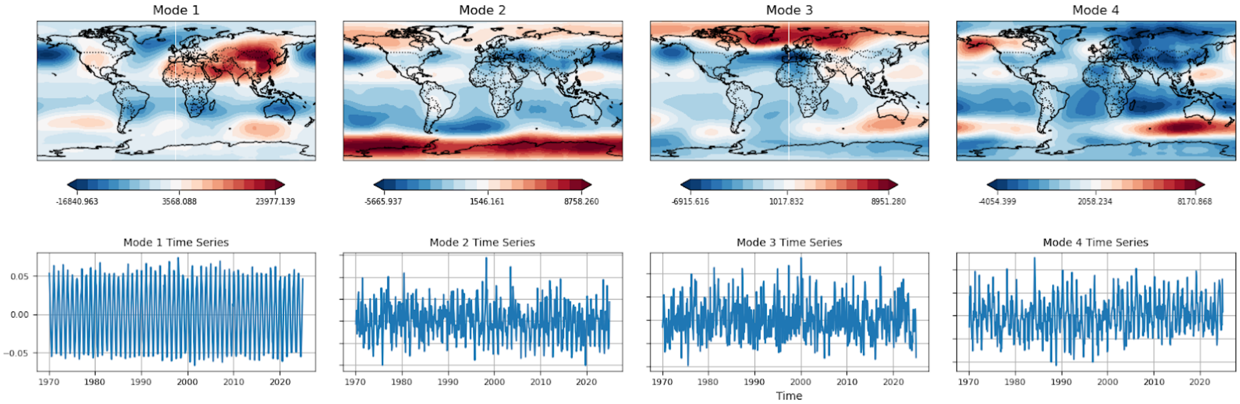


Figure 11. As in Figure 9, except using all core weights in combination.

The primary benefit of the Tucker decomposition is that it does not need to flatten data fields into vectors to perform its analysis. As such, the Tucker decomposition is naturally suited for mulit-way analysis where coupled structure and interactions are difficult to ascertain through vectorization. Examples include evolution of 3D variables such as ocean and atmosphere, cross-multivariable interactions, or multi-model ensemble outputs. Each factor matrix provides a set of dominant patterns for a single axis (e.g., space, time, variables), analogous to singular vectors of a matrix unfolding for EOFs, while the core tensor encodes how these mode-wise bases interact across dimensions. Individual Tucker components are therefore not rank-one objects. Instead, the data are reconstructed as weighted sums of outer products of mode-specific basis vectors, with the weights given by entries of the core tensor. As a result, physical or dynamical meaning does not attach to any single mode in isolation but emerges from the coupling structure in the core, which specifies which combinations of spatial, temporal, or variable patterns are active and with what strength.

Summary and Discussion

Digesting and interpreting gridded climate data is not always straightforward when computing modes of variability with statistical methods. Here, using large ensembles commonly used for climate variability diagnosis, we have attempted to summarize common challenges associated with the application of EOFs leading to misinterpretation, and have provided some examples, best practices, and alternative methods to consider. It can be tempting to assign physical meaning to a statistical mode, but intention and understanding of the physical processes must always come into play when evaluating the mathematics. Does it make physical sense? If not, we caution readers from the overinterpretation of EOFs themselves. Common challenges that manifest from EOF applications include mode swapping and sign flipping; however, they are relatively easy to handle by directly operating on the EOF itself. It is our intent to provide a quick reference guide (e.g., Table 2) detailing EOFs and alternative methods that broadly fit into three categories: EOF variants, linear methods, and multilinear methods. Any method that extracts modes of variability must be fit for purpose, always taking care to make sure modes are applicable to a specific science question and avoiding misinterpretation[21].

Not yet fully discussed, but equally important is the question of computing modes of variability under climate change, now typically addressed with the application of large ensembles[9,17]. How does one parse the variability of a system when the base state, by definition, is still changing? As the Earth system continues to remain out of equilibrium with the influx of anthropogenic greenhouse gas forcing, a key consideration is how, and if, the forced response should be removed. There is no right answer, rather, it entirely depends on the purpose and science question being asked. If one chooses to remove the forced response to elucidate a baseline, natural state, then detrending the data and using anomalies to compute the EOF, is a common approach. The difference between detrending, or not, can have significant consequences for interpretation. In our mode swapping example in Figure 2, the difference between the rate of mode swapping for PSA1 and PSA2 markedly changes depending on whether or not the forced response is removed. However, even detrending and removing the forced response will not necessarily remove any feedback that occurs due to the forcing itself. These types of questions inspired the creation of ForceSMIP, (Forced Component Estimation Statistical Methods Intercomparison Project[23]), which we encourage readers to follow for a deeper dive into forced response issues[24,8,25,26]. Finally, we acknowledge that this article does not exhaustively cover all interpretation pitfalls with EOFs or statistical methods, but we hope our overview, and comparison of methods, aid readers when choosing techniques best suited for their science.

Methods

Datasets

Two reanalyses, two large ensembles, and the CMIP6 (Coupled Model Intercomparison Project Phase 6) database are used to demonstrate EOF issues and alternatives. For reanalysis products, we sample sea level pressure (PSL) at monthly intervals for the period of 1950-2023, from the ECMWF reanalysis, version 5, (ERA5)[27], as well as the 20th Century Reanalysis (20CR)[28] from 1900- 2005. Both analyses provide pressure level data with ERA5 regridded to 0.25° horizontal resolution and 20CR regridded to 2° horizontal resolution. For large ensemble simulations suites, we utilize four sets of large ensembles from fully coupled Earth system models and include: E3SMv1 and v2 (the Department of Energy's Energy Exascale Earth System Model, versions 1 and 2), and CESM1 and CESM2 (the Community Earth System Model, versions 1 and 2)[16]. Again, we analyze PSL at monthly intervals from the historical simulations, where E3SMv1 employs 17 ensemble members, E3SMv2, 21 members, CESM1 (40 members), and CESM2 (50 members). E3SM employs the E3SM Atmosphere model (EAM) and the Model for Prediction Across Scales-Ocean: MPAS-Ocean, whereas CESM employs the Community Atmosphere model (CAM) and the Parallel Ocean Program, version 2 (POP2). Further details on versions 1 and 2 for both modelling frameworks are found in the respective modelling documentation papers[29,30,31,32]. From the

CMIP6 archive, we utilize PSL at monthly intervals for the historical simulations for the same period as the 20CR. In addition to E3SM and CESM, CMIP6 archive models[8,9] included in Table 1 include ACCESS-ESM1-5, CanESM2, CanESM5, CSIRO-Mk360, EC-Earth3, GFDL-CM3, GFDL-ESM2M_v2, GFDL-SPEAR, IPSL-CM6A, MIROC6, MIROC-ES2L, MPI-GE-CMIP6, MPI-GE.

Techniques

Here we describe all methodologies discussed including standard EOF computation, EOF alternatives, and finally techniques applied to directly adjust standard EOFs. Once again here, we emphasize that there is no best method for all applications. Rather, each different technique illuminates different features of the data, which is helpful for different types of analysis. For every method, we will describe how it is computed, and then demonstrate an application on a CESM2 LE member. The purpose is to display example outputs, rather than perform any scientific analysis. Finally, for each method we will briefly recap 1) the structure that the method isolates and 2) discuss when this method could be useful or limited.

Technique: Standard EOF Computation

To allow for complete comparison between the different methods, we will briefly recap how standard EOFs, or principal component analysis (PCA), is computed. Given a data matrix $\mathbf{X}$ of dimensions $p \times n$, where p is the number of spatial locations and n is the number of time steps, the data is typically centered by subtracting the temporal mean at each location. The covariance matrix $\mathbf{C} = (1/n) \times \mathbf{X} \times \mathbf{X}^t$ is then computed, and its eigenvalue decomposition yields orthonormal eigenvectors (the EOFs) and eigenvalues that indicate the amount of variance explained by each mode. By design, the spatial patterns (the EOFs) and their corresponding temporal coefficients (the PCs) are orthogonal. The **rank** of this decomposition refers to the number of modes retained, i.e., how many spatial–temporal patterns are needed to approximate the data well.

The explained variance is one of the strongest features of EOFs, allowing one to order the modes by importance. Further, it is often the case that a few leading modes capture the majority of the total variance. For this reason, EOF analysis can result in a parsimonious representation of the dynamics. The orthogonality further helps to isolate uncorrelated modes. However, EOF analysis can also produce patterns that are complex and challenging to interpret physically, potentially leading to misunderstanding or confusion [5]. Challenges with EOF analysis include:

- **Assumes data are best described by orthogonal spatial modes**, which may not correspond to physically meaningful structures.
- **Mixes physical patterns** when eigenvalues are nearly degenerate, leading to spatial modes that are hard to interpret.
- **Captures variance, not structure** — modes are ranked by explained variance, even if they do not align with meaningful dynamical or physical features.
- **Sensitive to sampling variability**, especially in the presence of noise or short time series (Figure 12).
- **Assumes linear correlations** — does not capture nonlinear interactions.
- **Applies only to two-dimensional (matrix) data** — requires flattening multidimensional arrays, which discards structural information (e.g., separating space and time).
- **Allows negative loadings**, which may be difficult to interpret in contexts like precipitation or energy where signals are inherently nonnegative.
- **Lacks uniqueness when eigenvalues are repeated**, making modes unstable across realizations or datasets.

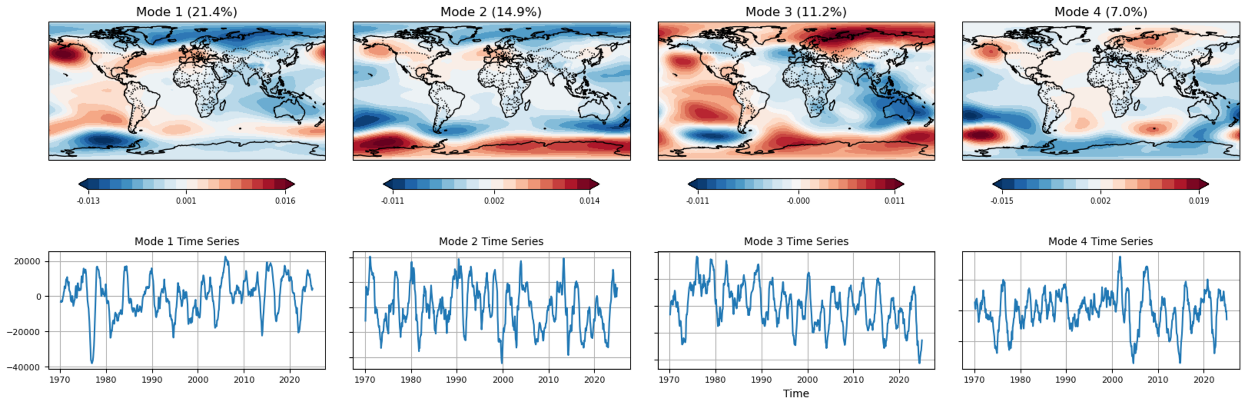


Figure 12. As in Figure 6, except using a 1-year rolling average to demonstrate the potentially large differences when temporal filtering is applied.

Extracted Structure and Applicability - Standard EOF analysis isolates orthogonal spatial modes that maximize explained variance under a linear covariance model. This makes EOFs effective at identifying the most energetic large-scale patterns and providing a compact, ordered representation of variability. The method is most useful as a baseline dimensionality-reduction and diagnostic tool, particularly for well-sampled, approximately linear systems where variance is a meaningful proxy for importance. However, because EOFs prioritize variance rather than dynamical or physical structure, they can mix distinct processes, be sensitive to sampling and preprocessing choices, and yield modes that are difficult to interpret physically. Despite these challenges, EOFs have cemented themselves as the most ubiquitous tool within Earth sciences for dimensionality reduction and mode discovery.

Techniques: EOF Variants

While standard EOF analysis provides an optimal low-rank decomposition in terms of explained variance, several variants have been developed to address its limitations or to impose additional constraints motivated by scientific interpretability. These variants often modify the underlying basis functions, adjust orthogonality constraints, or introduce sparsity-promoting priors. In this section, we briefly discuss three common approaches: EOF expansions using **Common Basis Functions**, **Rotated EOFs**, and **Sparse EOFs**.

Technique - EOF Variants: Common Basis Functions

The standard EOF analysis is designed for one sample matrix **X** of data. The Common Basis Function (CBF) approach is a multivariate generalization of EOF analysis designed to extract shared spatial patterns across multiple related datasets. Naive computations of EOFs for ensembles can be derived either from computing EOFs separately for each dataset, or by stacking the data to compute combined EOFs. In contrast to stacked EOF approaches which optimize variance only in an ensemble-averaged sense and can be dominated by high-variance members, the CBF method explicitly seeks a single set of spatial basis functions that are simultaneously important across all datasets, enabling direct and consistent comparison between them.

Suppose we have M datasets $\{\mathbf{X}^1, \mathbf{X}^2, ..., \mathbf{X}^M\}$, each of size $p \times n$, where p is the number of spatial locations and n is the number of time steps. The goal is to find a common set of orthonormal spatial basis vectors $\mathbf{B} \in \mathbb{R}^{pxr}$ (i.e., $\mathbf{B}^T\mathbf{B} = \mathbf{I}$), along with dataset-specific temporal coefficients $\mathbf{S}^m \in \mathbb{R}^{rxn}$ such that:

$$\mathbf{X}^m \approx \mathbf{B}\mathbf{S}^m \qquad \text{for m} = 1, ..., \text{M}.$$

This can be formulated as a maximum likelihood problem[33] [,34] . As a result, the extracted modes exhibit a consistent ordering of explained variance within each dataset, rather than reflecting variability dominated by a subset of ensemble members. The CBF method is particularly useful when comparing multiple models or observations in a unified spatial framework[39,35].

Extracted Structure and Applicability - The CBF approach isolates a shared set of orthonormal spatial modes that jointly maximize explained variance across multiple related datasets. This makes explicit the spatial structures that are common to all datasets while allowing their temporal evolution to differ, enabling direct cross-dataset comparison in a unified basis. CBF is particularly useful when analyzing ensembles, multi-model archives, or observational–model comparisons where structural alignment is desired. However, because the method still prioritizes variance and enforces a single global spatial basis, it can suppress dataset-specific features and remains subject to the same linearity and interpretability limitations as standard EOF analysis.

Technique - EOF Variants: Rotated EOFs

While standard EOFs are optimal in terms of explained variance and form an orthogonal basis, their spatial patterns often suffer from poor, unphysical interpretability, especially when modes have nearly equal eigenvalues and represent mixtures of physical structures. **Rotated EOFs (REOFs)** attempt to address this by applying a rotation—typically orthogonal or oblique—to a subset of the leading EOFs to produce more localized and physically interpretable patterns[36,37]. Compared to standard EOFs, REOFs trade orthogonality and variance ordering for better physical localization and interpretability, especially in the presence of degeneracies among leading modes[9].

The procedure begins with standard EOF analysis to obtain the leading r EOF loading vectors, typically denoted $\mathbf{L} \in \mathbb{R}^{p \times r}$, where p is the number of spatial points. An orthogonal matrix $\mathbf{R} \in \mathbb{R}^{r\times r}$ $(\mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I})$is then applied to yield new, ‘rotated’ loadings:

$$\mathbf{L'} = \mathbf{L}\cdot\mathbf{R},$$

where $\mathbf{R}$ is chosen to optimize some criterion. Probably the most common method is the **Varimax** objective, which maximizes the variance of squared loadings across modes[38] . The corresponding rotated principal components (temporal coefficients) are $\mathbf{S'} = \mathbf{R}^T\cdot\mathbf{S}$, where $\mathbf{S}$ contains the original principal components. This transformation preserves the total variance and subspace spanned by the original EOFs, but the rotated modes are no longer orthogonal. ROEFs yield regionalized patterns that depend on the number of retained modes. When only a few modes are used, the patterns can be highly sensitive to changes in *r*. As *r* approaches the number of spatial grid points, ROEFs tend to localize into patterns centered around individual points.

Previous studies have shown that REOF analysis is able to avoid the unphysical dipole-like EOF analysis pattern that often appears when the known dominant mode has the same sign across the domain and to simplify spatial structures while retaining the robust patterns[21,39].

Extracted Structure and Applicability - Rotated EOF analysis isolates spatial modes within the leading EOF subspace that are optimized for localization and interpretability rather than variance optimality. This rotation makes physically coherent and regionally confined patterns more explicit, particularly when standard EOFs mix structures due to near-degenerate eigenvalues. REOFs are most useful when interpretability of spatial patterns is prioritized over strict orthogonality or variance ordering, such as in regime or teleconnection analysis. However, the resulting modes depend on the number of retained EOFs and the chosen rotation criterion, and the loss of orthogonality can complicate statistical interpretation and comparison across datasets.

<u>Technique - EOF Variants: Sparse EOF</u>

Standard EOF analysis yields spatial patterns that are typically global in extent and difficult to interpret physically when the dominant modes have broad spatial support. **Sparse EOFs** aim to improve interpretability by promoting spatial localization of the modes through sparsity constraints. In this formulation, spatial patterns are encouraged to have many near-zero entries, highlighting only the most relevant regions contributing to variability.

Sparse EOFs are typically formulated as a **regularized matrix factorization** problem, where the goal is to approximate the data matrix $\mathbf{X} \in \mathbb{R}^{p \times n}$ as $\mathbf{X} \approx \mathbf{B}\,\mathbf{S}$, with $\mathbf{B} \in \mathbb{R}^{p \times r}$ representing spatial modes and $\mathbf{S} \in \mathbb{R}^{r \times n}$ the temporal coefficients, just as in standard EOFs. However, instead of requiring that the columns of **B** be orthogonal eigenvectors, the optimization adds a sparsity-inducing penalty, often the ℓ_1 norm:

$$\text{minimize} \quad \|X - B\,S\|^2\_F + \lambda \textstyle\sum_k \|b_k\|_1 \;, \quad \text{subject to} \quad \|b_k\|_2 = 1 \quad \text{for each mode } k,$$

where $\mathbf{b_k}$ is the k-th column of **B**, and a penalty term $\alpha > 0$ controls the degree of sparsity[40,41]. The penalty term controls the degree of sparsity - larger α results in more sparse EOFs. The result is a set of EOF-like modes that retain much of the explanatory power of standard EOFs while being easier to associate with physical mechanisms or localized features. Sparse PCA appears to have had limited application within the Earth sciences, though it has been applied for instance for more interpretable fingerprinting[42].

Extracted Structure and Applicability - Sparse EOF analysis isolates spatial modes that explain variability while explicitly enforcing localization through sparsity constraints. This highlights regionally confined structures and suppresses weak or diffuse loadings, improving physical interpretability relative to standard EOFs. Sparse EOFs are most useful when the underlying processes are expected to be spatially localized or when interpretability of individual regions is critical, such as in detection or fingerprinting studies. However, the results depend sensitively on the choice of sparsity penalty, sacrifice strict orthogonality, and can bias variance attribution by excluding low-amplitude but dynamically relevant structures.

Techniques: Linear Methods

While EOFs (PCA) are among the most widely used techniques for identifying dominant patterns of variability in spatiotemporal datasets, they belong to a broader class of linear dimensionality reduction methods. Several alternative techniques—including **Factor Analysis (FA), Independent Component Analysis (ICA), Nonnegative Matrix Factorization (NMF),** and **Dynamic Mode Decomposition (DMD)** —offer different decompositional frameworks based on distinct statistical or structural assumptions. These methods share the goal of representing high-dimensional data using a low-rank approximation, but they differ in how the components are derived, constrained, and interpreted. In the following subsections, we compare each of these methods to EOFs, focusing on their objectives, assumptions, and implications for interpretability in the context of geophysical data.

Technique - Linear Methods: Factor Analysis

Factor analysis (FA) is a linear dimensionality reduction technique closely related to EOFs, but with a distinct modeling philosophy. While EOFs seek orthogonal directions that maximize explained variance, FA assumes that the observed variables are driven by a smaller number of **latent (unobserved) factors**, plus unique noise specific to each observed variable[43,37]. The observed data $\mathbf{X} \in \mathbb{R}^{n \times p}$ is modeled as:

$$\mathbf{X} = \mathbf{F}\Lambda^{T} + \mathbf{E},$$

where:

- $\mathbf{F} \in \mathbb{R}^{n \times r}$ contains the scores of *r* unobserved (latent) factors,

- $\Lambda \in \mathbb{R}^{p \times r}$ is the loading matrix linking the latent factors to the observed variables,

- $\mathbf{E} \in \mathbb{R}^{n \times p}$ is the idiosyncratic Gaussian noise, assumed to be uncorrelated across variables and with a diagonal covariance matrix.

The assumption of FA is that the high-dimensional data arises from a low-dimensional set of underlying processes, possibly confounding or not directly observable, such as circulation regimes or hidden drivers of variability in the climate system. The noise term accounts for individual variability that is not shared between variables. Unlike EOFs, which assume all variance is shared and seek orthogonal modes, FA models both shared and unique variance separately. This distinction allows FA to represent more flexible and realistic relationships among variables by permitting non-orthogonal loadings, which is especially useful for modeling correlated processes. While EOFs decompose the total variance in the data, FA focuses specifically on capturing the covariance structure, often through maximum likelihood estimation. Additionally, FA can more effectively account for measurement noise or unresolved structure, making it particularly useful in settings where some factors are believed to be hidden or confounded. Factor analysis is a common data analysis technique used throughout the social sciences, with limited applications in the Earth sciences[44]. FA is conceptually appealing in settings where latent dynamics (e.g., unobserved climate drivers) are suspected to govern the observed data.

Extracted Structure and Applicability - Factor analysis isolates latent factors that explain the shared covariance among observed variables while explicitly modeling variable-specific noise. This makes it well suited for identifying hidden or confounding processes that drive correlated variability, rather than simply maximizing total variance. FA is most useful when measurement noise or unresolved structure is believed to be substantial and when the goal is to infer underlying drivers rather than simply producing orthogonal modes. However, the method relies on strong distributional and noise assumptions, does not enforce spatial structure, and can be sensitive to model specification and identifiability choices.

Technique - Linear Methods: Independent Component Analysis

EOFs have as their primary objective, the successive maximization of variance of the modes. The orthogonality and uncorrelatedness come as byproducts of this goal. By contrast, **Independent Component Analysis (ICA)** has statistical independence as its main aim[45]. This makes ICA especially well-suited for extracting physically distinct processes when data are non-Gaussian. For Gaussian variables, uncorrelatedness and independence coincide, and ICA reduces to EOF/PCA. However, many climate and geophysical datasets are non-Gaussian, motivating the use of ICA as a generalization of EOFs in this setting[37] .

Conceptually, ICA shares a modeling philosophy with FA: they are both latent variable models, meaning that they both assume that observed variables are generated by a smaller number of latent (hidden) components. The difference lies in the objective. FA explains the covariance structure of the observed variables using fewer latent variables (factors). The emphasis is on modeling correlations and accounting for noise. ICA seeks to find latent variables that are statistically independent and assumes that the observed variables are mixtures of these independent sources. The goal is source separation. In linear ICA, the model is:

$$\mathbf{x} = \boldsymbol{\Lambda}\,\mathbf{f}, \quad \text{where } \mathbf{f} \text{ is a vector of statistically independent components,}$$

and $\boldsymbol{\Lambda}$ is a full-rank *mixing matrix*. The observed data vector $\mathbf{x}$ is a linear combination of these latent sources. The goal is to estimate an unmixing matrix W such that:

$$\hat{\mathbf{s}} = \mathbf{W}\,\mathbf{x}$$

recovers the independent components $\hat{\mathbf{s}} \approx \mathbf{f}$. T. The optimal W is obtained by minimizing a cost function measuring statistical dependence, typically based on non-Gaussianity, mutual information, or entropy[36] . This objective contrasts with the variance-maximizing goal of EOFs or the likelihood-based objective of factor analysis. Unlike EOFs, ICA does not impose orthogonality or rank ordering, and the decomposition is not unique: the independent components are identifiable only up to permutation and scaling. ICA has been used to investigate sea level pressure and water storage[46,47].

Extracted Structure and Applicability - Independent Component Analysis isolates latent components that are statistically independent rather than merely uncorrelated or variance-optimal. This emphasizes source separation and can reveal physically distinct processes in non-Gaussian data that are mixed in standard EOF representations. ICA is most useful when independence is a meaningful assumption, such as when multiple underlying processes contribute additively and exhibit non-Gaussian statistics.

However, the decomposition is not unique, lacks natural mode ordering, and can be sensitive to preprocessing and noise, which complicates interpretation and comparison across datasets.

Technique - Linear Methods: Nonnegative Matrix Factorization

Nonnegative Matrix Factorization (NMF) is a linear dimensionality reduction technique that differs from EOFs in one major fundamental way: it imposes **nonnegativity constraints** on both spatial and temporal components. This makes NMF especially useful when the data are naturally nonnegative (e.g., precipitation, cloud cover, energy fluxes), and when interpretability is enhanced by additive, parts-based representations[46]. Given a nonnegative data matrix $\mathbf{X} \in \mathbb{R}^{+p\times n}$, NMF approximates $\mathbf{X}$ as the product of two low-rank nonnegative matrices:

$$\mathbf{X} \approx \mathbf{W}\,\mathbf{H},$$

where:

- $\mathbf{W} \in \mathbb{R}^{+p\times r}$ contains the spatial patterns (basis vectors),

- $\mathbf{H} \in \mathbb{R}^{+r\times n}$ contains the temporal activations or coefficients.

Unlike EOFs, which yield orthogonal modes with both positive and negative loadings, NMF provides strictly additive components, allowing for a “parts-based” decomposition. This can yield spatial structures that are localized and physically interpretable, as each field is reconstructed as a nonnegative linear combination of a few basic building blocks. In contrast to EOFs, which can produce modes with large canceling positive and negative values, NMF enforces nonnegativity, leading to more interpretable and parts-based representations. While EOFs offer an optimal decomposition in terms of variance explained, NMF prioritizes interpretability and sparsity at the cost of optimality. Unlike EOFs, NMF modes are not constrained to be orthogonal or ordered by explained variance, which does bring about ambiguity of importance for the underlying structures. Further, NMF lacks a unique solution and is sensitive to initialization and algorithmic choices, in contrast to the closed-form solution of EOFs obtained via SVD. In Earth Sciences, NMF has been applied to problems such as identifying characterizing drought behavior in river basins and quantifying the sources of atmospheric particles[49,50].

Extracted Structure and Applicability - Nonnegative Matrix Factorization isolates additive spatial–temporal components under explicit nonnegativity constraints, yielding parts-based representations of the data. This highlights localized, physically interpretable structures when variables are inherently nonnegative and when cancellation between positive and negative loadings is undesirable. NMF is most useful for exploratory analysis and feature discovery in settings where interpretability and compositional structure are prioritized over variance optimality. However, the decomposition is not unique, lacks natural mode ordering, and is sensitive to initialization and algorithmic choices, which can limit robustness and reproducibility.

Technique - Linear Methods: Dynamic Mode Decomposition

EOFs identify spatial patterns that capture maximal variance in the data, without directly considering how those patterns evolve over time, i.e., the PCs are selected to be orthogonal. By contrast, **Dynamic Mode Decomposition (DMD)** is a data-driven technique designed to extract spatiotemporal patterns that evolve according to approximate linear dynamics. DMD is based on the idea that the temporal evolution of the system can be approximated by a linear operator, which maps each snapshot of the system state to the next. The resulting decomposition yields spatial patterns, each associated with a fixed temporal frequency and a corresponding growth or decay rate. This makes DMD particularly well-suited for identifying coherent structures such as traveling waves, oscillations, and instabilities in geophysical flows and climate systems[51,52]. There are many different algorithms that fall under the DMD category; however, in its most basic form, DMD constructs the least-squares linear operator that advances snapshots forward in time and directly eigendecomposes this operator to obtain modes and their associated growth or decay rates[53]. This is very similar to the Linear Inverse Modeling (LIM), and Principal Oscillation Pattern / Principal Interaction Pattern (POP/PIP) methods used throughout the climate literature.

As clarified by Tu et al. (2013) [52], DMD, LIM and POP/PIP methods can all be understood as estimating essentially the same finite-dimensional linear evolution operator from time-lagged data, differing primarily in formulation, assumptions, and historical context[52]. In LIM, this operator is estimated by fitting a linear stochastic dynamical system whose deterministic component governs the mean evolution of the state and whose eigenvalues characterize dominant decay and oscillatory time scales. POP/PIP analysis similarly focuses on the eigenmodes of a linear propagator inferred from lag-covariance statistics, with an emphasis on identifying dynamically relevant oscillatory and interacting patterns. These approaches are mathematically equivalent under conditions of linear dynamics, sufficiently rich state representations, evenly sampled time series, and when the same lagged least-squares estimator is used for the propagator. DMD can be viewed as the most direct generalization of this idea, as it explicitly forms the empirical linear map without imposing an explicit stochastic model. Moreover, DMD admits a broad range of generalizations - including extended DMD, kernel DMD, total-least-squares DMD, optimized DMD, and control-aware variants - that relax linearity, incorporate noise and inputs, or approximate Koopman operators on richer observable spaces, making it a more flexible framework in practice. For this reason, we focus on DMD over the LIM or POP analysis as alternative methods. However, to keep the analysis simple here, we will restrict ourselves to basic DMD.

Mathematically, DMD takes a sequence of state vectors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x\_k}$, and seeks a matrix $\mathbf{A}$ such that

$$\mathbf{x}_{j+1} \approx \mathbf{A}\,\mathbf{x}_j$$

for each time slice j. Equivalently, the state vectors are arranged into two data matrices: $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k]$ and $\mathbf{X}' = [\mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_{k+1}]$. The matrix $\mathbf{A}$ is then estimated by solving the linear system $\mathbf{X}' \approx \mathbf{A}\,\mathbf{X}$, often using a low-rank approximation via the SVD. The eigenvalues and eigenvectors of the matrix A (or its projection onto a reduced subspace) reveal the dynamic modes and their temporal behavior. The dynamics of each mode is controlled by powers of the (complex) eigenvalue. Each mode evolves exponentially in time, either oscillating, growing, or decaying, and the full spatiotemporal dynamics of the system are represented as a linear combination of these modes.

Each eigenvalue λ_k governs the temporal evolution of the associated mode k'th. Concretely, this is determined as follows. First, the amplitudes are determined by solving a least squares problem that projects the initial state onto the DMD modes. This finds the optimal coefficients to reconstruct the starting condition as a linear combination of the modes. Writing $\lambda_k = r_k e^{i\theta k}$, the magnitude r_k determines exponential growth or decay, while the phase θ_k sets an oscillation with frequency. Using this, a continuous-time generator $\mu_k = 1/(\Delta t) * \log(\lambda_k) = \sigma_k + i\omega_k$ is estimated so that the k'th mode evolves as $e^{\sigma_k t} \cos(\omega_k t)$. The product of each mode's amplitude and its exponential evolution gives the temporal behavior of that individual mode, and these are combined together to reconstruct the full dynamics across all time steps.

Whereas EOFs prioritize variance maximization and produce orthogonal spatial patterns, DMD focuses on uncovering temporal dynamics and typically results in non-orthogonal modes. Moreover, DMD does not rank modes by explained variance, but rather by dynamical significance, such as dominant frequencies or timescales. In this sense, DMD can be viewed as complementary to EOF analysis. While EOFs are optimal for compressing information, DMD is better suited for analyzing and predicting time-evolving structures, especially in systems governed by approximately linear dynamics. Under many circumstances, DMD is equivalent to the well known method in the climate community known as **Linear Inverse Modeling (LIM)** [52]. The applications of LIM, and therefore DMD are quite vast across the Earth sciences, and it would be challenging to succinctly summarize its applications. We also note that DMD belongs to a broader class of algorithms, each designed to address different limitations of the traditional DMD framework, and ranging from mode collapse and scalability to improved representation of the underlying dynamics[54]. Given the limitations of this manuscript, we will not go into further descriptions of these techniques.

Extracted Structure and Applicability - Dynamic Mode Decomposition isolates spatiotemporal modes associated with approximately linear evolution, each characterized by a fixed frequency and growth or decay rate. This makes coherent oscillatory behavior, traveling structures, and dominant timescales explicit in both space and time, rather than prioritizing variance alone. DMD is most useful when the dynamics can be well approximated by linear evolution over the sampling window, and when temporal coherence or prediction is of primary interest. However, the method can be sensitive to noise, sampling, and rank truncation, produces non-orthogonal modes, and may struggle to represent strongly nonlinear or transient dynamics without extensions.

Techniques: Multi-Linear Methods

EOF analysis is traditionally applied to two-dimensional data matrices, such as a space × time dataset. However, many modern datasets are **multidimensional** (or *multilinear*): For example: latitude × longitude × time; or: model ensemble × latitude × longitude × time. Flattening these into a 2D matrix for EOF analysis can destroy important structure — such as the difference between the two spatial dimensions or between models and physical space.

Multilinear methods generalize the logic of EOFs to **tensors** — higher-order arrays — in ways that preserve the multidimensional structure of the data. Rather than vectorizing or slicing the data and losing directional specificity, multilinear decompositions seek to jointly decompose the full tensor along all its modes. This enables mode-specific dimensionality reduction, where distinct sets of latent factors are extracted along each axis (e.g., spatial modes, temporal modes, ensemble modes).

There are many different types of tensor decompositions, each with their own strengths and weaknesses. Here we will cover two methods: **The Polyadic Decomposition**, also called

CANDECOMP/PARAFAC[55,56] which generalizes a rank decomposition of a matrix, and the Tucker Decomposition[57], which generalizes PCA through a smaller core tensor.

Technique - Multilinear Methods: Polyadic Decomposition

Canonical Polyadic Decomposition (CPD) — also called CANDECOMP/PARAFAC — is a generalization of EOF analysis to higher-dimensional arrays, called tensors. If we combine the singular values with the PCs from EOFs, we can write

$$X \approx \sum \mathbf{u}_r \mathbf{v}_r^T = \sum \mathbf{u}_r \otimes \mathbf{v}_r$$

Here, the $\otimes$ symbol represents the **outer product** — a generalization of multiplication between vectors to form higher-dimensional arrays. Each term in the above sum has rank-1, and taking R singular vectors is a rank-R decomposition of **X**. The outer product can be performed with more vectors to create higher dimensional arrays (tensors). CP does the same with rank-1 tensors in three or more dimensions.

Just as EOF expresses a 2D matrix as a sum of outer products of spatial and temporal modes, CP expresses a 3D (or 4D, 5D, etc.) data tensor as a sum of rank-1 tensors, each constructed from one vector in each dimension:

$$X \approx \sum \mathbf{u}_r \otimes \mathbf{v}_r \otimes \mathbf{w}_r$$

This means that the data value at position *(i, j, k)* — say, (latitude i, longitude j, time k) — is approximated as a sum of products of values from: a spatial mode along dimension 1 ($\mathbf{u}_r$), a spatial mode along dimension 2 ($\mathbf{v}_r$), and a temporal mode ($\mathbf{w}_r$). Each term in the sum is like a triple-mode EOF component, one per mode of the data tensor. The number of such terms, R, is called the tensor rank, and it plays the same role as the number of retained EOFs: higher *R* captures more structure at the cost of increased complexity[58]. In general, tensors can have as many dimensions as needed, allowing for complex interactions and large compression.

The CP decomposition expresses a tensor as a sum of **R** separable components—one for each mode. This often leads to more parsimonious representations, particularly when the true structure of the data is separable across dimensions. Unlike EOFs, which discard tensor structure through matricization, CP preserves the full multidimensional nature of the data. While EOF modes are orthogonal by construction, CP components are not, yet the CP decomposition can still be unique under mild conditions, up to scaling and permutation[59]. Moreover, CP yields mode-specific components (e.g., one temporal and two spatial), which can result in better physical interpretability.

Extracted Structure and Applicability - Canonical Polyadic Decomposition isolates separable, rank-1 components across each mode of a multidimensional data tensor, yielding coupled spatial and temporal factors without flattening the data. This preserves the inherent multilinear structure and makes interactions across dimensions explicit, often resulting in parsimonious and physically interpretable representations. CP is most useful when the underlying processes are approximately separable across dimensions and when maintaining tensor structure is essential. However, the decomposition can be sensitive to noise and rank selection, lacks orthogonality, and may struggle to represent strongly non-separable or highly complex dynamics without increasing rank substantially.

Technique - Multilinear Methods: Tucker Decomposition

The **Tucker decomposition** generalizes EOF analysis to higher-order data arrays, or tensors, by allowing a low-rank projection along each dimension of the data simultaneously. Given a three-dimensional data tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}$^{I × J × K}, such as latitude × longitude × time, Tucker approximates the data as:

$$\boldsymbol{\mathcal{X}} \approx \boldsymbol{\mathcal{G}} \times_1 \mathrm{A} \times_2 \mathrm{B} \times_3 \mathrm{C}$$

Here, **A**, **B**, and **C** are factor matrices containing the principal components along the first, second, and third modes, respectively—analogous to the left and right singular matrices **U** and **V** matrices in SVD. The core tensor $\boldsymbol{\mathcal{G}} \in \mathbb{R}$^{$R_1 \times R_2 \times R_3$} captures the interaction strengths between components from each mode. The mode-n product $\times_n$ represents a generalization of matrix multiplication to tensors, allowing us to project $\boldsymbol{\mathcal{X}}$ along each axis independently[58].

Where EOFs decompose a matrix into rank-one outer products of vectors, Tucker decomposes a tensor into a multilinear product of basis vectors and a dense core. The EOF decomposition can be viewed as a special case of Tucker where the core tensor is diagonal and only two modes are present. Tucker relaxes the orthogonality and diagonal structure, allowing for richer cross-mode interactions. The number of components can be controlled independently for each axis, enabling anisotropic compression and a more tailored approximation of the data.

Extracted Structure and Applicability - Tucker decomposition isolates low-dimensional subspaces along each mode of a data tensor while explicitly modeling their interactions through a dense core tensor. This makes cross-mode coupling and anisotropic structure visible, allowing different levels of compression along space, time, or other dimensions. Tucker is most useful when variability is not well captured by separable rank-1 components and when different modes exhibit distinct intrinsic dimensionalities, or when the user needs to investigate more nuanced interactions between dimensions. However, the core tensor can be difficult to interpret physically, the decomposition is not unique, and the increased flexibility can lead to overfitting or reduced interpretability without careful rank selection.

Techniques: Direct Adjustment Methods

Methods that directly adjust sign flipping are automatically applied within Climate Variability Diagnostics Package, version 6, (CVDPv6)[17] and PCMDI Metrics Package (PMP)[19] packages. In the CVDPv6, the sign at a point in the center of action is used to determine whether the sign is consistent or not, whereas for the PMP, a pattern correlation is used. PMP identifies the need of sign flipping by spatial pattern correlation to the EOF pattern obtained from the reference dataset and flips the sign of the model's pattern when the pattern correlation is negative. Additional diagnoses were applied for some variability modes by assessing sign of area-averaged EOF pattern over certain geographical regions.

Mode swapping, however, is not automatically adjusted within the packages and is incumbent upon the user to adjust manually. For example, one could use the statistics for determining mode swapping described in section *Challenges with EOFs/Mode Swap*, and then manually change the mode numbers assigned to the modes of interest (i.e. PSA1, PSA2). Table 1 mode swapping statistics were computed as follows: We determine the absolute value of the pattern correlation between ERA5 PSA1 and a given model and member's PSA1 and PSA2. Then, if the absolute value of the pattern correlation between ERA5 PSA1 and the model's PSA2 is greater than the absolute value of the pattern correlation between ERA5 PSA1 and model PSA1, we conclude that the modes are swapped. The forced response is computed as the ensemble mean sea level pressure (SLP), which is removed from the SLP field prior to calculation of the PSA1 and 2 modes. For observations, we use the 30-year running mean to define the forced response, and that is also removed from PSL prior to calculation of the modes. Because mode swapping is an artifact of statistical methods, frequency metrics (Table 1) can easily change due to sampling. Examples of sampling choices include (but not limited to) detrending method, pattern correlation methodology, time period used, ensemble set, number of ensemble members, etc.

Appendix

For completeness and to provide illustrations of alternative methods discussed, we provide comparisons with Figure 6's standard monthly PSL EOF, but for methods not shown in the *Alternative Method Comparisons* section.

Sparse EOFs

Sparse EOFs were implemented using scikit-learn's MiniBatchSparsePCA algorithm. We present results for varying values of the sparsity parameter α, with larger values enforcing greater sparsity. Specifically, we examine α = 1000, 7000, and 10000.

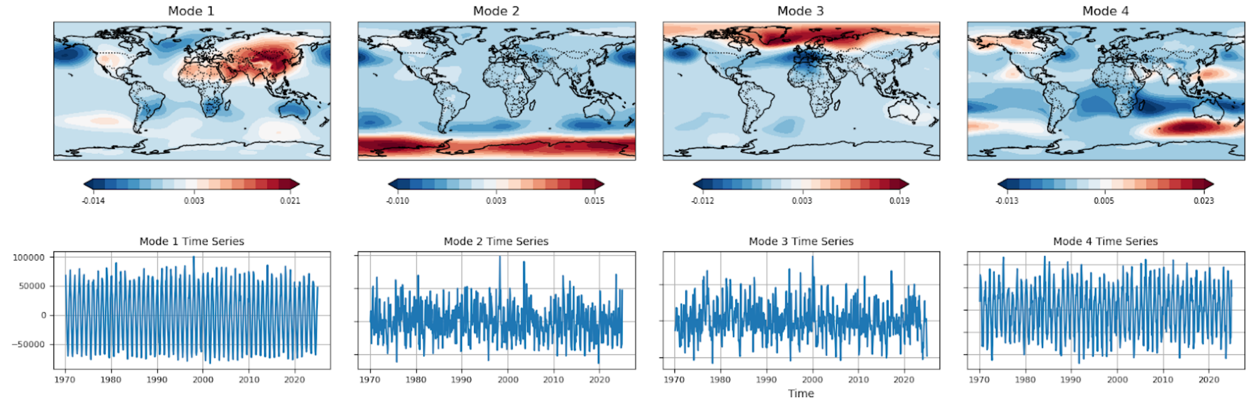


Figure A1. As in Figure 6, but an example of Sparse EOF, where α = 1000.

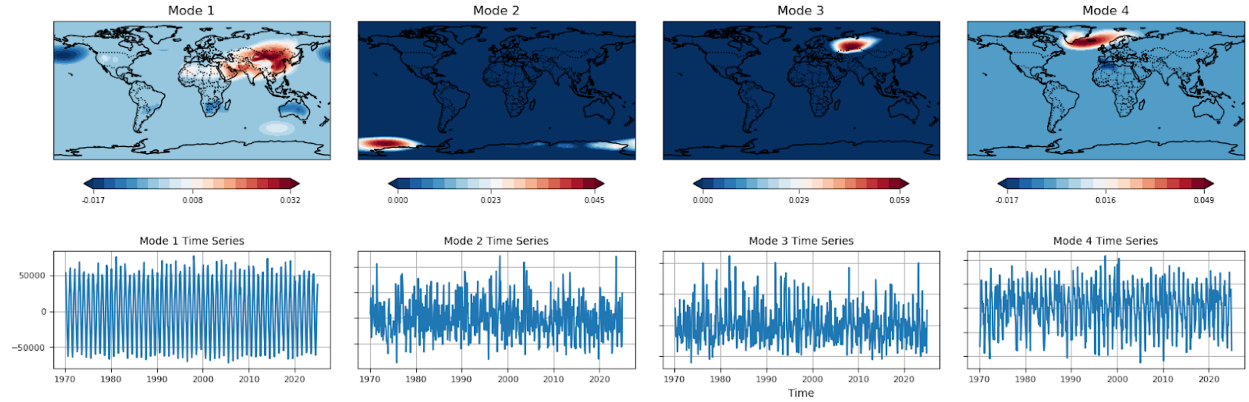


Figure A2. As in Figure A1, but an example of Sparse EOF, where α = 7000.

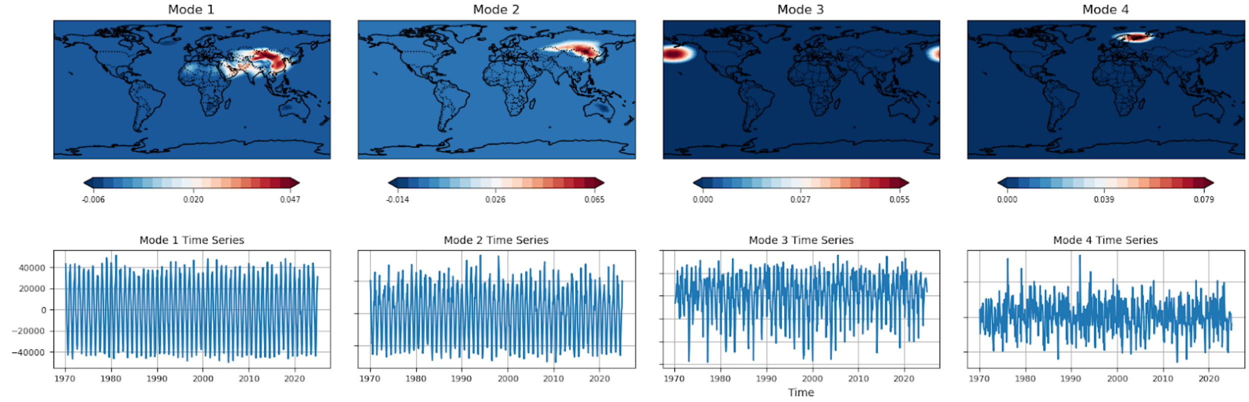


Figure A3. As in Figure A1, but an example of Sparse EOF, where α = 10000.

For lower α, the spatial patterns closely resemble those from standard EOFs, with notable differences emerging primarily near regions of high variance, such as over Russia. As α increases, the modes become increasingly localized, concentrating around the primary centers of activity for each pattern. This spatial separation is reflected in the associated principal component time series, which begin to exhibit clearer seasonal structure, similar to that of the leading EOF mode.

Independent Component Analysis

In ICA, the first mode appears quite similar to the first EOF. This means that the direction of maximal variance happens to also be the direction chosen to maximize independence. The second ICA mode is a large southern global teleconnection with nodes of importance in the Pacific and Atlantic. This appears to be a combination of the second and fourth EOFs. The third and fourth ICA modes grab the Southern and Northern Hemispheric patterns, analogous with the second and third EOFs.

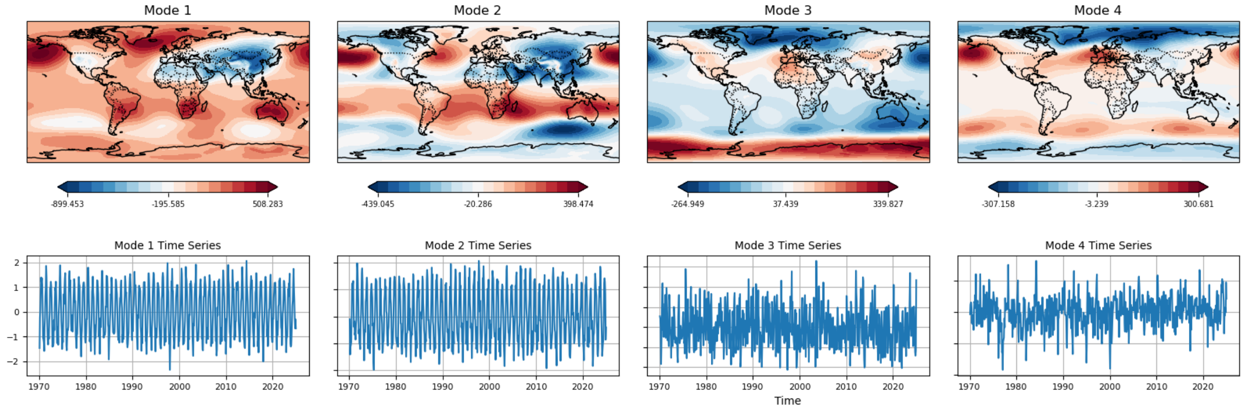


Figure A4. As in Figure 6, but for Independent Component Analysis

Nonnegative Matrix Factorization

In NMF, both the spatial and temporal modes are constrained to be nonnegative. The first two modes resemble those from the leading EOF, capturing the prominent Africa–South Asia signal. Subsequent modes highlight other localized features, such as activity near Alaska and the North Atlantic. A key distinction of NMF, relative to other methods, is that the polar regions consistently remain near zero across all modes. This arises from the additive nature of NMF: unlike EOFs or other linear methods, it cannot balance large positive contributions with offsetting negative values. As a result, the decomposition emphasizes localized, purely additive structures, leading to inherently different spatial representations.

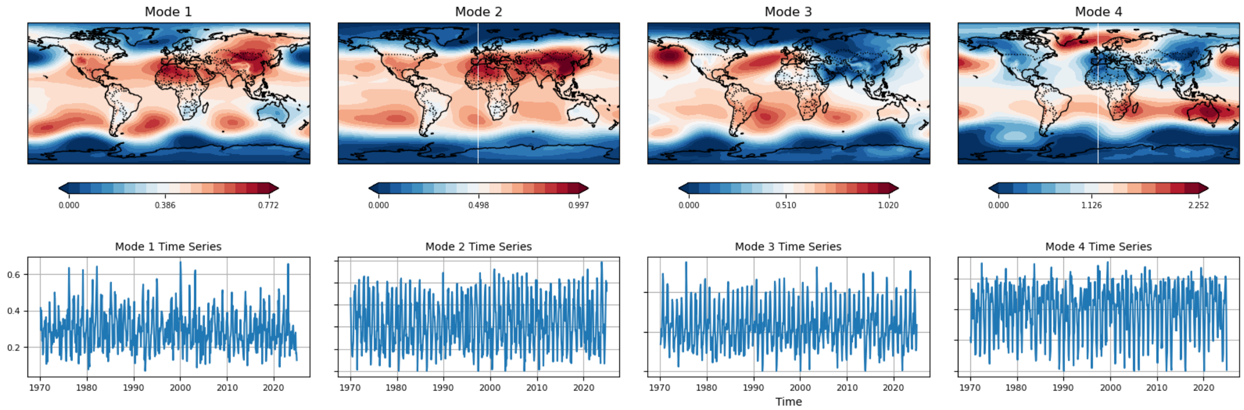


Figure A5. As in Figure 6, but for Nonnegative Matrix Factorization.

Dynamic Mode Decomposition

In the basic DMD algorithm, both spatiotemporal modes and their associated eigenvalues are extracted, with both quantities generally complex-valued. The temporal evolution of each mode is governed by powers of its corresponding eigenvalue, enabling fast and interpretable reconstruction of dynamics. Figure A6 is a plot of the leading eigenvalues, sorted by magnitude.

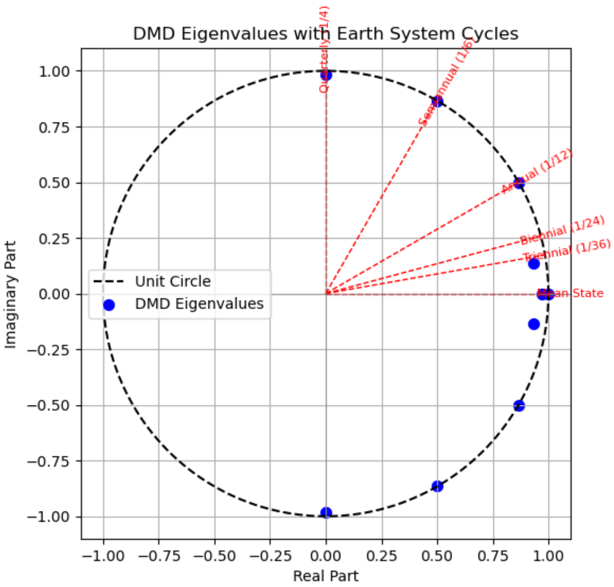


Figure A6. Leading eigenvalue with Earth System Cycles (mean state, triennial, biennial, annual, semiannual, quarterly) sorted by magnitude. Blue circles represent DMD eigenvalues and the black dashed line is a unit circle for reference.

To visualize the spatial patterns and dynamics, it is common practice to take the real parts of the modes and time series. Since DMD eigenvalues often come in complex-conjugate pairs, care must be taken to avoid redundancy by selecting only one representative from each pair. Shown here are the first four distinct modes in Figure A7.

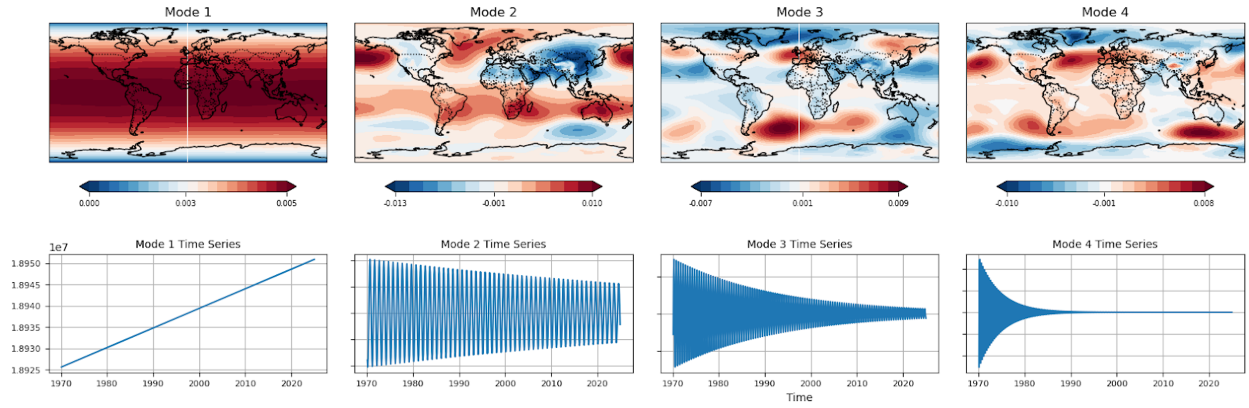


Figure A7: As in Figure 6, but for dynamic mode decomposition.

From the structure of the eigenvalues, we identify these modes as corresponding to the mean, annual, semiannual, and quarterly cycles. The magnitude of each eigenvalue determines its long-term behavior: eigenvalues with magnitude one produce persistent oscillations; those less than one decay over time; and those greater than one amplify. In this case, the mean mode has a magnitude greater than one, suggesting a long-term warming trend. Indeed, the largest eigenvalue is $\lambda = 1.0000008+0i$ which is why we see 1) no oscillations and 2) slight growth. Decaying modes identified by DMD may reflect genuine physical transients or damping processes, but they can also arise as an artifact of extrapolating eigenvalues with magnitude less than one beyond the training interval. For example, a decaying seasonal mode might suggest a reduction in seasonal variability; however, standard DMD cannot substantiate this interpretation because it assumes time-independent growth rates over the entire analysis period. Windowed approaches or more advanced variants of DMD are required to distinguish true temporal evolution from such methodological artifacts.

Canonical Polyadic Decomposition

In the CPD, we extract R modes along the latitude, longitude, and temporal dimensions, where R is the chosen decomposition rank. The spatial modes are formed by the outer product of the corresponding latitude and longitude factors, which we display below. The resulting spatial and temporal structures are highly sensitive to the choice of R. For low rank (e.g., R=4), the leading CPD modes resemble those from EOF analysis, with some notable distinctions. Specifically, the first CPD mode closely matches the leading EOF, while the others reflect the dataset's latitudinal structure, analogous to the second and third EOFs.

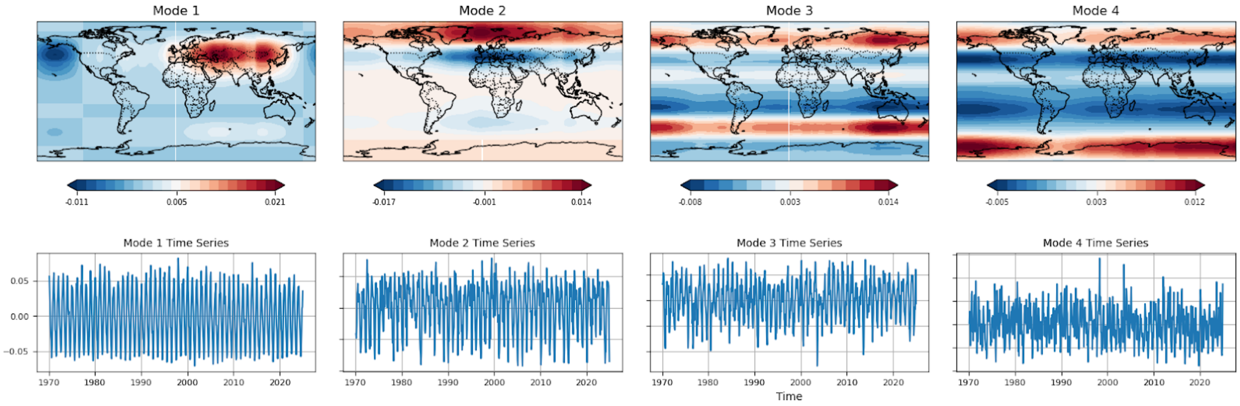


Figure A8. As in Figure 6, but for canonical polyadic decomposition using a low rand of R=4.

When the rank is increased (e.g., R=12), the decomposition reveals more complex structures. These higher-rank modes tend to exhibit dipole-like or blocky spatial patterns not seen in the lower-rank or EOF results. As CPD lacks a natural ordering of modes, we present a representative subset of these patterns in Figure A9.

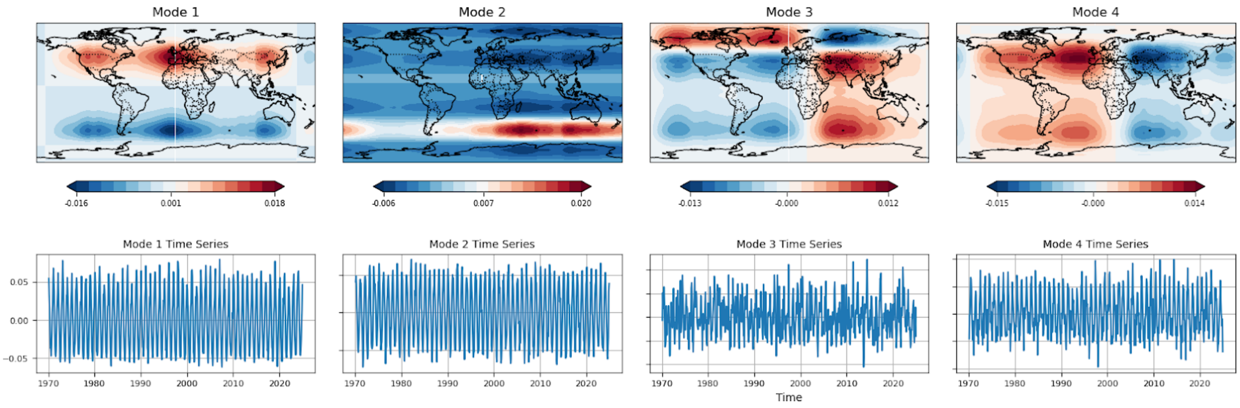


Figure A9. As in Figure 6, but for canonical polyadic decomposition using a high rank of R=12.

References

1. Penland, C., and T. Magorian, 1993: Prediction of Niño 3 Sea Surface Temperatures Using Linear Inverse Modeling. J. Climate, 6, 1067–1076, https://doi-org.cuucar.idm.oclc.org/10.1175/1520-0442(1993)006<1067:PONSST>2.0.CO;2.
2. Lorenz EN (1956) Empirical orthogonal functions and statistical weather prediction. Statistical Forecasting Project report 1, MIT Department of Meteorology.
3. North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling Errors in the Estimation of Empirical Orthogonal Functions. Mon. Wea. Rev., 110, 699–706, https://doi-org.cuucar.idm.oclc.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.C
4. von Storch H, Zwiers FW (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge (ISBN: 0 521 45071 3).
5. Hannachi A, Jolliffe IT, Stephenson DB (2007) Empirical orthogonal functions and related techniques in atmospheric science: a review. Int J Climatol 27:1119–1152.
6. Monahan AH, Fyfe JC, Ambaum MH, Stephenson DB, North GR (2009) Empirical orthogonal functions: the medium is the message. J Clim 22:6501–6514.
7. Chen, Z., B. Gan, L. Wu, and F. Jia, 2018: Pacific–North American teleconnection and North Pacific Oscillation: Historical simulation and future projection in CMIP5 models. Climate Dyn., 50, 4379–4403, https://doi.org/10.1007/s00382-017-3881-9.
8. Lee, J., Sperber, K.R., Gleckler, P.J. et al. Quantifying the agreement between observed and simulated extratropical modes of interannual variability. Clim Dyn 52, 4057–4089 (2019). https://doi-org.cuucar.idm.oclc.org/10.1007/s00382-018-4355-4.
9. Fasullo, J. T., A. S. Phillips, and C. Deser, 2020: Evaluation of Leading Modes of Climate Variability in the CMIP Archives. J. Climate, 33, 5527–5545, https://doi.org/10.1175/JCLI-D-19-1024.1.

10. Lian, T., and D. Chen, 2012: An Evaluation of Rotated EOF Analysis and Its Application to Tropical Pacific SST Variability. J. Climate, 25, 5361–5373, https://doi.org/10.1175/JCLI-D-11-00663.1.
11. Chen, X., J. M. Wallace, and K. Tung, 2017: Pairwise-Rotated EOFs of Global SST. J. Climate, 30, 5473–5489, https://doi.org/10.1175/JCLI-D-16-0786.1.
12. National Center for Atmospheric Research Staff (Eds). Last modified 22 Jul 2013. "The Climate Data Guide: Empirical Orthogonal Function (EOF) Analysis and Rotated EOF Analysis." Retrieved from https://climatedataguide.ucar.edu/climate-data-tools-and-analysis/empirical-orthogonal-function-eof-analysis-and-rotated-eof-analysis.
13. Weidman, S., Kleiner, N., & Kuang, Z. (2022). A rotation procedure to improve seasonally varying empirical orthogonal function bases for MJO indices. Geophysical Research Letters, 49(15), e2022GL099998.
14. Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. Bull. Amer. Meteor. Soc., 96, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1.
15. Milinski, S., Maher, N., and Olonscheck, D.: How large does a large ensemble need to be?, Earth Syst. Dynam., 11, 885–901, https://doi.org/10.5194/esd-11-885-2020, 2020.
16. Fasullo, J. T., and Coauthors, 2024: Modes of Variability in E3SM and CESM Large Ensembles. J. Climate, 37, 2629–2653, https://doi.org/10.1175/JCLI-D-23-0454.1.
17. Phillips, A. S., C. Deser, and J. Fasullo, 2014: A New Tool for Evaluating Modes of Variability in Climate Models. EOS, 95, 453-455, doi: 10.1002/2014EO490002.
18. Maher, N., Phillips, A. S., Deser, C., Wills, R. C. J., Lehner, F., Fasullo, J., Caron, J. M., Brunner, L., and Beyerle, U.: The updated Multi-Model Large Ensemble Archive and the Climate Variability Diagnostics Package: New tools for the study of climate variability and change, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-3684, 2024.

19. Lee, J., Gleckler, P. J., Ahn, M.-S., Ordonez, A., Ullrich, P. A., Sperber, K. R., Taylor, K. E., Planton, Y. Y., Guilyardi, E., Durack, P., Bonfils, C., Zelinka, M. D., Chao, L.-W., Dong, B., Doutriaux, C., Zhang, C., Vo, T., Boutte, J., Wehner, M. F., Pendergrass, A. G., Kim, D., Xue, Z., Wittenberg, A. T., and Krasting, J.: Systematic and objective evaluation of Earth system models: PCMDI Metrics Package (PMP) version 3, Geosci. Model Dev., 17, 3919–3948, https://doi.org/10.5194/gmd-17-3919-2024, 2024.
20. Dommenget, D. Evaluating EOF modes against a stochastic null hypothesis. Clim Dyn 28, 517–531 (2007). https://doi.org/10.1007/s00382-006-0195-8.
21. Dommenget, D., & Latif, M. (2002). A cautionary note on the interpretation of EOFs. *Journal of Climate*, *15*, 216–225.
22. Kurihana, T., and Coauthors, 2024: Identifying Climate Patterns Using Clustering Autoencoder Techniques. Artif. Intell. Earth Syst., 3, e230035, https://doi.org/10.1175/AIES-D-23-0035.1.
23. Wills, R.C. J., Deser, C., McKinnon, K.A. et al. Forced Component Estimation Statistical Method Intercomparison Project (ForceSMIP). ESS Open Archive . June 16, 2025. DOI: 10.22541/essoar.175003371.14843115/v1 .
24. Zhang Y, Wallace JM, Battisti DS (1997) ENSO-like interdecadal variability: 1900–93's. J Clim 10:1004–1020.
25. Bonfils C, Santer BD (2011) Investigating the possibility of a human component in various Pacific Decadal Oscillation indices. Clim Dyn 37:1457–1468. https://doi.org/10.1007/s00382-010-0920-1.
26. Bonfils C, Santer BD, Phillips TJ, Marvel K, Ruby Leung L, Doutriaux C, Capotondi A (2015) Relative contributions of mean-state shifts and ENSO-driven variability to precipitation changes in a warming climate. J Clim 28:9997–10013. https://doi.org/10.1175/JCLI-D-15-0341.1
27. Hersbach H, Bell B, Berrisford P, et al. The ERA5 global reanalysis. Q J R Meteorol Soc. 2020; 146: 1999–2049. https://doi.org/10.1002/qj.3803

28. Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Matsui, N., Allan, R.J., Yin, X., Gleason, B.E., Vose, R.S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R.I., Grant, A.N., Groisman, P.Y., Jones, P.D., Kruk, M.C., Kruger, A.C., Marshall, G.J., Maugeri, M., Mok, H.Y., Nordli, Ø., Ross, T.F., Trigo, R.M., Wang, X.L., Woodruff, S.D. and Worley, S.J. (2011), The Twentieth Century Reanalysis Project. Q.J.R. Meteorol. Soc., 137: 1-28. https://doi.org/10.1002/qj.776
29. Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., et al. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. Journal of Advances in Modeling Earth Systems, 11, 2089–2129. https://doi.org/10.1029/2018MS001603
30. Golaz, J.-C., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., et al. (2022). The DOE E3SM Model version 2: Overview of the physical model and initial model evaluation. Journal of Advances in Modeling Earth Systems, 14, e2022MS003156. https://doi.org/10.1029/2022MS003156
31. Hurrell, J. W., and Coauthors, 2013: The Community Earth System Model: A Framework for Collaborative Research. Bull. Amer. Meteor. Soc., 94, 1339–1360, https://doi.org/10.1175/BAMS-D-12-00121.1.
32. Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). Journal of Advances in Modeling Earth Systems, 12, e2019MS001916. https://doi.org/10.1029/2019MS001916
33. Hannachi, A., Finke, K., & Trendafilov, N. (2023). Common EOFs: A tool for multi-model comparison and evaluation. *Climate Dynamics*, *60*, 1689–1703. https://doi.org/10.1007/s00382-022-06409-8.
34. Flury, B. N. (1984). Common principal components in k groups. *Journal of the American Statistical Association*, *79*, 892–898. https://doi.org/10.2307/2288721

35. Lee, J., K. Sperber, P. Gleckler, K. Taylor, and C. Bonfils, 2021: Benchmarking performance changes in the simulation of extratropical modes of variability across CMIP generations. Journal of Climate, 34, 6945–6969, doi: 10.1175/JCLI-D-20-0832.1
36. Richman, M. B. (1986). Rotation of principal components. *Journal of Climatology*, *6*(3), 293–335. https://doi.org/10.1002/joc.3370060305
37. Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). Springer.
38. Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. https://doi.org/10.1007/BF02289233
39. Cheng, X., Nitsche, G., & Wallace, J. M. (1995). Robustness of low-frequency circulation patterns derived from EOF and rotated EOF analyses. *Journal of Climate*, *8*, 1709–1713.
40. Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, *15*(2), 265–286. https://doi.org/10.1198/106186006X113430
41. Witten, D. M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*, *10*(3), 515–534. https://doi.org/10.1093/biostatistics/kxp008
42. Weylandt, M., & Swiler, L. P. (2024). Beyond PCA: Additional dimension reduction techniques to consider in the development of climate fingerprints. *Journal of Climate*, *37*(5), 1723–1735.
43. Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). University of Chicago Press.
44. Kline, P. (1994). *An easy guide to factor analysis* (1st ed.). Routledge. https://doi.org/10.4324/9781315788135
45. Hyvärinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, *13*(4–5), 411–430. https://doi.org/10.1016/S0893-6080(00)00026-5.
46. Forootan, E., Schumacher, M., Mehrnegar, N., Bezděk, A., Talpe, M. J., Farzaneh, S., ... & Shum, C. K. (2020). An iterative ICA-based reconstruction method to produce consistent time-variable total water storage fields using GRACE and Swarm satellite data. *Remote Sensing*, *12*(10), 1639.

47. Hannachi, A., Unkel, S., Trendafilov, N. T., & Jolliffe, I. T. (2009). Independent component analysis of climate data: A new look at EOF rotation. *Journal of Climate*, *22*(11), 2797–2812.
48. Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*(6755), 788–791. https://doi.org/10.1038/44565
49. Liang, C. S., Yue, D., Wu, H., Shi, J. S., & He, K. B. (2021). Source apportionment of atmospheric particle number concentrations with wide size range by nonnegative matrix factorization (NMF). *Environmental Pollution*, *289*, 117846.
50. Talsma, C. J., Bennett, K. E., & Vesselinov, V. V. (2022). Characterizing drought behavior in the Colorado River Basin using unsupervised machine learning. *Earth and Space Science*, *9*(5), e2021EA002086.
51. Schmid, P. J. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, *656*, 5–28. https://doi.org/10.1017/S0022112010001217
52. Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., & Kutz, J. N. (2014). On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, *1*(2), 391–421. https://doi.org/10.3934/jcd.2014.1.391
53. Colbrook, M. J. (2024). The multiverse of dynamic mode decomposition algorithms. In *Handbook of Numerical Analysis* (Vol. 25, pp. 127–230). Elsevier.
54. Schmid, P. J. (2022). Dynamic mode decomposition and its variants. *Annual Review of Fluid Mechanics*, *54*(1), 225–254.
55. Harshman, R.A., (1970). Foundations of the PARAFAC procedure: Models and conditions for an “explanatory" multi-modal factor analysis, UCLA Working Papers in Phonetics, 16 , pp. 1--84. Available online from http://www.psychology.uwo.ca/faculty/harshman/wpppfac0.pdf.
56. Caroll, J. D. and Chang J.J., (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of “Eckart-Young” decomposition, Psychometrika, 35, pp. 283--319, https://doi.org/10.1007/BF02310791.

57. Tucker, L.R. Some mathematical notes on three-mode factor analysis. Psychometrika 31, 279–311 (1966). https://doi.org/10.1007/BF02289464.
58. Kolda, T. G., & Bader, B. W. (2009). Tensor decompositions and applications. *SIAM Review*, *51*(3), 455–500. https://doi.org/10.1137/07070111X.
59. Kruskal, J. B. (1977). Three-way arrays: Rank and uniqueness of trilinear decompositions. *Linear Algebra and Its Applications*, *18*(2), 95–138.

Funding

CAS, JC, JF, and JL acknowledge support by the Regional and Global Model Analysis (RGMA) component of the Earth and Environmental System Modeling Program of the U.S. Department of Energy’s Office of Biological & Environmental Research (BER) under (CAS, JC, JF) Lawrence Livermore National Lab subaward DE-AC52-07NA27344, Lawrence Berkeley National Lab subaward DE-AC02-05CH11231, and Pacific Northwest National Lab subaward DE-AC05-76RL01830. The work of JL is performed under the auspices of the U.S. Department of Energy (DOE) by Lawrence Livermore National Laboratory (LLNL) under Contract No. DE-AC52-07NA27344. AJ and DD were supported by the Regional and Global Model Analysis (RGMA) component of the Earth and Environmental System Modeling (EESM) program of the U.S. Department of Energy's Office of Science, as a contribution to the HiLAT-RASM project. The efforts of Dr. Fasullo in this work were also supported by NASA Awards 80NSSC21K1191, 80NSSC17K0565, and 80NSSC22K0046. This work was supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation (NSF) under Cooperative Agreement No. 1852977. Additionally, this research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award BER-ERCAP0033110.

Author contributions

CAS contributions include research inception and design, project management and organization, writing, and editing; DD writing, alternate method expertise, figure creation; AJ research inception and design, writing, figure creation; JL writing and figure creation; JC figure creation and statistics; AP figure creation; JF editing.

Competing Interests

CAS is an Associate Editor for npj Climate and Atmospheric Sciences.

Data availability

E3SMv2 Large Ensemble, 1850-2100, https://portal.nersc.gov/archive/home/c/ccsm/www/E3SMv2/FV1/atm/proc/tseries/month_1; CESM2 Large Ensemble, 1950-2023, https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm2le.output.html; E3SMv1 ensemble data, LLNL ESGF node, https://aims2.llnl.gov/; CESM1 Large ensemble, NSF NCAR Research Data Archive, https://rda.ucar.edu/datasets/d651027/; ERA5, 1950-2023, PSL https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels-monthly-means?tab=overview; The Twentieth Century Reanalysis (20CR), https://psl.noaa.gov/data/20thC_Rean/.

Figure legends

Figure 1: East Atlantic pattern (EA, row 1) and Scandinavian pattern (SCA, row 2) shown for ERA5 (first column), CESM2 ScenarioMIP ensemble member r1i1p1f1 (second column) and CESM2 ScenarioMIP ensemble member r3i1p1f1 (third column). Mode swapping is evident in the third column. EA and SCA are defined as the 2nd and 3rd EOF patterns of area-weighted PSL computed over 20:80°N,

-90:40°E for JFM 1979-2022. Units are in hPa and variance explained is listed at the top right of each panel.

Figure 2: Pacific South American (PSA) modes 1 (row 1) and 2 (row 2) shown for ERA5 (first column), CESM2 Large Ensemble member 1181.010 (second column) and CESM2 Large Ensemble member 1161.009 (third column). Mode swapping is evident in the third column. PSA1 and PSA2 are defined as the 2nd and 3rd EOF patterns of area-weighted PSL computed over 20:90°S, 0:360°E for June-August 1950-2023. Units are in hPa and variance explained is listed at the top right of each panel. The patterns are created by regressing global PSL anomalies onto normalized PC timeseries.

Figure 3: Pacific North American pattern shown for ERA5 (first panel), E3SMv2 member r18i1p1f1 (second panel) and E3SMv2 member r29i1p1f1 (third panel). Sign flipping is evident in the third panel. Units are in hPa and variance explained for each pattern is listed at the top right. The PNA is defined as the first EOF pattern of area-weighted PSL computed over 20:85°N, 120°E:120°W for June-August 1950-2022. Units are in hPa and variance explained is listed at the top right of each panel. The patterns are created by regressing global PSL anomalies onto normalized PC timeseries.

Figure 4: Center of action variability for North Atlantic Oscillation (top), East Atlantic (middle) and Scandinavian (top) patterns computed for three ensemble members of the E3SMv2 Large Ensemble: 0101 (left), 0201 (center), and 0301 (right), for a combination of historical and SSP370 simulations. The NAO is defined as the first EOF, the East Atlantic pattern as the second EOF and the Scandinavian pattern as the third EOF of area-weighted PSL computed over 20:80°N, 90°W:40°E for January-March over 100-year periods staggered by 10 years. The mean EOFs are shown, as well as the centers of action for each of the sixteen 100-year periods between 1850 and 2100, marked by white dots.

Figure 5. Mode swapping pattern correlation adjustment method illustration. (Upper panel) Pattern correlation of CMIP6 models' Pacific/North American (PNA) pattern during the September-October-November (SON) season (upper panel). The PNA pattern for each model was determined by the leading Empirical Orthogonal Function (EOF 1) of sea level pressure fields. Pattern correlation was then calculated between each model's EOF 1 spatial pattern and the 20th Century Reanalysis (20CR) PNA

pattern (SON season). For each model, the pattern correlation of the leading three EOFs (EOF 1-3) was assessed against the 20CR PNA using spatial pattern correlation, Root Mean Square Error (RMSE), and temporal correlation between the EOF principal component (PC) time series and a Common Basis Function PC time series. The highest pattern correlation achieved after potentially swapping EOF modes based on these criteria is indicated by markers, while the pattern correlation of the original EOF 1 is shown as a gray bar. The increasing spread between EOF1 and the alternative EOF on the right demonstrates the difficulty in automated selection of the best-matching EOF mode for models on the right, which highlights challenges in systematically identifying robust climate patterns across different models (further details in Lee et al., 2019). (Lower panel) Spatial pattern of model's EOF 1 (left), 2 (middle) and reference dataset's EOF 1 demonstrating an example EOF swap case, obtained from EC-Earth3-AerChem model (far-most right in the upper panel) and the 20CR for PNA pattern during the SON season.

Figure 6. Figure 6. PSL EOFs 1,2,3,4 (with percent variance explained) and respective PC timeseries for CESM2 model large ensemble, first member. No temporal filtering is done for consistency and comparison across alternative methods. See the Methods section under *EOF: Standard Computation* for an illustration on the potentially large differences due to temporal filtering.

Figure 7. As in Figure 6, but an example of rotated EOF.

Figure 8. As in Figure 6 but using Factor Analysis.

Figure 9. Tucker decomposition core tensor of PSL CESM2 large ensemble member in Figure 6, with a multirank of (5, 10, 10), with temporal modes 0 through 4, 10 latitudes (y-axis) and 10 longitudes (x-axis).

Figure 10. As in Figure 6, but for Tucker decomposition for the first four temporal components using the largest weights.

Figure 11. As in Figure 9, except using all core weights in combination.

Figure 12. As in Figure 6, except using a 1 year rolling average to demonstrate the potentially large differences when temporal filtering is applied.