

Model Representation of ENSO & IPO: Analyzing Principal Components of Sea Surface Temperature

Alexander Fogal

ID#: 20721508

Professor Chris Fletcher
GEOG 408
University of Waterloo
April 13, 2022

Contents

Contents	i
List of Tables	ii
List of Figures	iii
1 Introduction and History	1
1.1 Review of Recent Findings	2
1.2 Objectives, Aims, and Summary of Motivation	4
2 Research Questions	4
3 Methods	5
3.1 Datasets and Pre-Processing	5
3.2 Empirical Orthogonal Function (EOF) Analysis	6
3.3 IPO TPI Analysis	7
4 Results	8
5 Discussion and Conclusions	22
References	25
6 Appendix A: Code and Details	30
7 Appendix B: Extended Methods	30
7.1 Empirical Orthogonal Function (EOF) Analysis	30

List of Tables

3.1	CMIP5 Models Summary	5
3.2	CMIP3 Models Summary	5
3.3	Observational and Reanalysis Dataset Summary	5

List of Figures

1.1	Spatial Pattern of ENSO	1
1.2	Spatial Pattern of IPO	2
3.1	Eigenvalue Spectrum for CCSM3	6
3.2	TPI Visualization	8
4.1	EOF 1 from CMIP3 Models	9
4.2	EOF 1 from CMIP5 Models	10
4.3	EOF 2 from CMIP3 Models	11
4.4	EOF 2 from CMIP5 Models	12
4.5	EOF 3 from CMIP3 Models	13
4.6	EOF 3 from CMIP5 Models	14
4.7	EOFs From HadSST3 Reanalysis Dataset	15
4.8	EOFs From NOAA Observational Dataset	16
4.9	FFT of PCs from CMIP3 Models	17
4.10	FFT of PCs from CMIP5 Models	18
4.11	FFT of PCs from Observational and Reanalysis Datasets	19
4.12	TPI Timeseries for CMIP3 Models	20
4.13	TPI Timeseries for CMIP5 Models	21
4.14	TPI Timeseries for Reanalysis and Observational Datasets	22

1 Introduction and History

Coupled climate systems, as well as full Earth System Models (ESMs), have had great success recreating observed climate phenomena, and have been improving over time. However, it is not so clear whether these models are capable of simulating the underlying mechanisms of climate and sources of variability, specifically, for important mechanisms such as the El-Niño Southern Oscillation (ENSO) and Interdecadal Pacific Oscillation (IPO) [1]. ENSO is generally recognized as the major mode of variability of the climate on annual scales, and is defined by Sea Surface Temperature (SST) anomalies in the Pacific Ocean, west of South America as seen in Figure 1.1. [2]. ENSO does not have a single, regular period, it ranges between 2-7 years, and the phases have asymmetric magnitudes [2]. Phase change of ENSO involves a shift in the thermal structure of the ocean, as well as a change in atmospheric conditions [2]. The importance of ENSO cannot be understated: through teleconnections and other changes to circulation, ENSO affects the meteorology of the whole globe; these effects are reviewed by McPhaden, Zebiak, and Glantz, the most clear being the effects on surface pressure, trade winds, and tropical precipitation, which have ramifications for forest fires, air quality, water availability, and agriculture [2], [3].

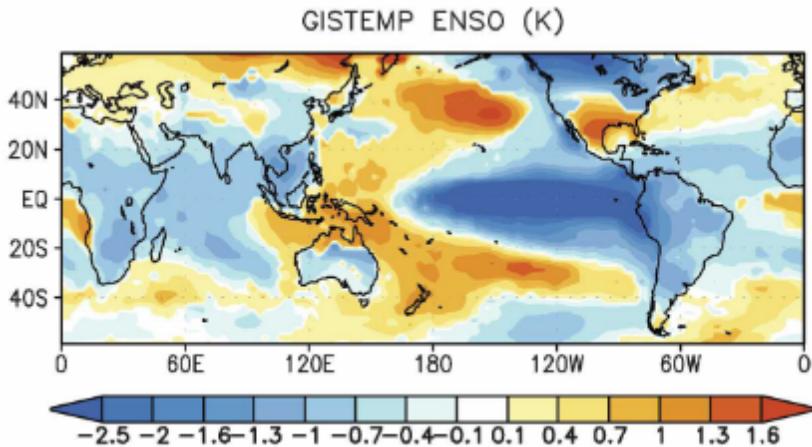


Figure 1.1: Spatial Pattern of ENSO based on EOF analysis of the GISTEMP reanalysis dataset, adapted from Chen *et al* [4].

Similarly, IPO, sometimes also referred to as Pacific Pan-decadal Variability (PDV) or Pacific Decadal Oscillation (PDO), is a strong source of decadal to interdecadal variability of SSTs in the Pacific basin, as can be seen in Figure 1.2 [4]–[6]. Herein, this phenomenon will be referred to only as IPO to avoid confusion, as there are occasionally minor disagreements in definitions between the three in literature [1], [4]. IPO is strongly linked to the changing over time of ENSO, as regime shifts in the North Pacific correspond with shifts in ENSO [4], [7], [8]. Furthermore, it was found that there were similar shifts occurring at 30-35 year intervals, an irregular period much like that of ENSO [4]. These shifts have also been linked to changes in the rate of anthropogenically forced global warming, including the “hiatus”

which ended in 2013 [1], [9].

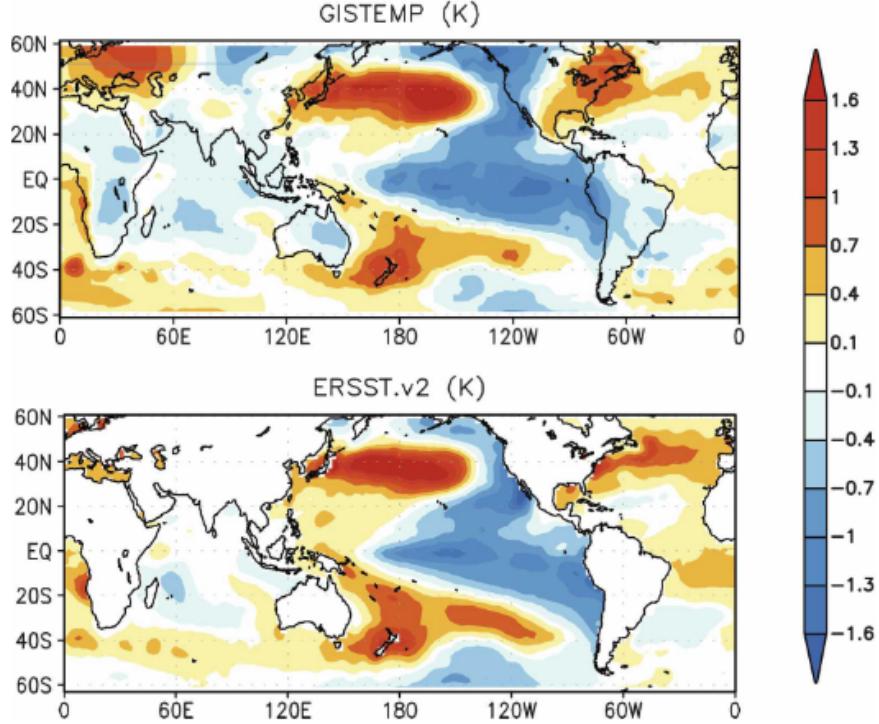


Figure 1.2: Spatial Pattern of IPO based on ENSO-removed EOF analysis of two reanalysis datasets, adapted from Chen *et al* [4].

Despite recent improvements in modelling and ESMs, it is not clear whether these are capable of simulating the underlying mechanisms of IPO/ENSO because, hitherto, there is little consensus about what the underlying mechanisms are; furthermore, all of these are very complex, involving interplay between ocean and atmosphere [2], [4]. This is further complicated by the complex inter-relation of IPO to ENSO, lacking historical climate data, and uneven spatial distribution and non-linearity of long-term global warming in SSTs [4]. As for ENSO itself, the atmosphere plays a major role, with convective processes being the most important [2], [10]. Specifically, the two main ENSO-driving atmospheric processes are typically modelled as the atmosphere causing SST anomalies to feedback and grow, as well as heat flux negative feedback causing damping [2], [11]. There are also several related nonlinear cloud processes that have been deemed necessary for correctly modelling ENSO [2], [12].

1.1 Review of Recent Findings

It is believed that ESMs are capable of simulating IPO accurately, and it is promising that some Coupled Model Inter-comparison (CMIP) 5 models surveyed in the 5th assessment

report (AR5) correctly predicted shifts in IPO, including models initialized in the early 1990s predicting the shift later that decade [1], [13], [14]. However, it has also been found that the CMIP3 models were generally unskillful in modelling IPO [15], [16]. As previously mentioned, the relation of IPO to ENSO is very important, as it is thought to directly influence ENSO and how it changes over time [1], [17]–[19]. Furthermore, there is evidence that the effects of IPO and ENSO can be statistically and significantly distinguished, particularly as they apply to the South Pacific Convergence Zone, for example, in Australian rainfall and flood risk [1], [20].

Moreso than IPO, ENSO has had much attention given to theoretical understanding, as well as correctly modelling and skill-benchmarking [1], [2], [21]. CMIP3 models were found to be able to simulate the more basic facets of ENSO, but were still struggling to simulate features like amplitude, irregular periodicity, seasonal phase-locked nature, and spatial variability [10], [22]–[26]. This is related to CMIP3 models poorly simulating long-term means and cycles in the tropical regions of the Pacific Ocean, which has been suggested as a source of bias for simulating ENSO, and inhibits the ability of the models to simulate individual related mechanisms [2], [22], [26]. An example of this would be that most CMIP3 models underestimate the response of wind to SST anomalies [2], [26].

There are also issues where ENSO statistics are correctly simulated, although for incorrect reasons; for example, the aforementioned underestimate impacting thermal damping, leading to the cancelling of errors [2], [26]. CMIP5 models are generally better, with over half falling within 25% of the observed ENSO amplitudes, where CMIP3 models had less falling in the same range [2]. There were also marked improvements in seasonal phase locking and location of maximum SST anomalies during peak ENSO, although CMIP5 performed about as poorly in simulating the spatial shape of SST anomalies [2]. CMIP5 models also showed some slight improvement in mean oceanic state, which is thought to be related to the improvements in ENSO representation [2]. As for the atmosphere, key processes identified did not show improvement and so errors cancelling is still a concern [2]. With these things in mind, it is also worth noting that some models do outperform others, and many of the previous conclusions are true only in the sense of a multi-model ensemble [2]. Furthermore, it was found that some models from CMIP5 incorrectly predict multi-decadal warming, which reduces confidence in other long-term predictions [1], [27].

It has been found that CMIP3 models were also generally unsatisfactory at modelling IPO, and also did not accurately depict SST modality or tropical-extratropical teleconnections [1], [15], [16], [28]. It has also been found that there was general improvement in modelling IPO and teleconnections, especially those to North American precipitation, from CMIP3 to CMIP5 [29]. Despite this, it was also concluded that CMIP5 models are still not properly representing mean sea level pressure variability in some regions inter-decadally, which is associated with improper simulation ENSO [1], [30]. There is also some conflict in projections of the IPO, CMIP3 models tend to show more negative phases of IPO, where it has been independently put forth that increasing greenhouse gasses (GHGs) and aerosols should instead cause more positive phases of IPO [1], [31], [32].

In summary, most CMIP3 models did not do a good job of simulating either ENSO or IPO, the underlying mechanisms or the effects, and most CMIP5 models showed only a modest improvement, if any improvement at all.

1.2 Objectives, Aims, and Summary of Motivation

It is clear that IPO and ENSO are important for global climate; hence, it is important to test how well models represent these. Therefore, the main goal of this report is to determine how well models from the CMIP3 and CMIP5 families represent variability due to ENSO and IPO, especially as compared to observational and reanalysis datasets. The objectives of this report are to answer the research questions outlined in Section 2 by performing an Empirical Orthogonal Function (EOF) analysis on SST fields from each dataset. This will provide information on the variability due to each of a set of so-called EOFs, with corresponding time series known as the Principal Components (PCs). This process will be outlined in Section 3. EOFs are a very powerful analytical tool which allows the representation of correlated fields via a relatively small set of functions and their eigenvalues, which explain variance in the original dataset due to each EOF [33], [34]. One advantage of this over other orthogonal function analysis techniques is the lack of a predefined form for the EOFs [33], [34]. This is especially nice for analyzing climate data, which do not have a known functional form and have intricate boundary conditions [33], [34]. This has lead to EOF analysis being among the most widely used tools in climate science, and thus it is fit for use in this report [34].

In addition, the so-called IPO Tripole Index proposed by Henley *et al* will be investigated herein, as it is both systematic and does not involve detrending SST data, nor does it involve any other assumptions about long-term and global warming processes [35]. The results will be presented in Section 4, and discussion will be presented in Section 5, as well as summary and conclusions.

2 Research Questions

The research questions to be addressed in this report are enumerated as follows:

- (RQ1) How well do models from CMIP3 and CMIP5 represent variability in SST due to phenomena such as ENSO and IPO?
- (RQ2) How does this compare to observational and reanalysis datasets?
- (RQ3) Do CMIP5 models improve over CMIP3 models?

3 Methods

This section will outline the methods of analysis used, as well as details of the datasets and any pre-processing performed on them.

3.1 Datasets and Pre-Processing

In this report, several datasets were analyzed, these are summarized in Tables 3.1 and 3.2 below, adapted from Polade *et al* [29]. For all models, the historical simulation case was used. The `tas` anomaly field, representing near-surface air temperature, was used for the ensuing analyses, which is analogous to SST. The observational and reanalysis datasets, are summarized in Table 3.3. All data was obtained via the KNMI Climate Explorer, details of which can be found in Section 6 [36].

Table 3.1: A list of CMIP5 models used in this report.

Institution	CMIP5 Name	Atmospheric Resolution (lon x lat)	Dates	Citation
CSIRO	CSIRO-Mk3-6-0	$1.875^\circ \times 1.875^\circ$	Jan. 1850 to Dec. 2005	[37]
MOHC	HadGEM2-ES	$1.875^\circ \times 1.25^\circ$	Dec. 1859 to Nov. 2005	[38]
NCAR	CCSM4	$1.25^\circ \times 0.9^\circ$	Jan. 1850 to Dec. 2005	[39]

Table 3.2: A list of CMIP3 models used in this report. Citations are the same as above.

Institution	CMIP3 Name	Atmospheric Resolution (lon x lat)	Dates
CSIRO	CSIRO-Mk3.5	$1.875^\circ \times 1.875^\circ$	Jan. 1871 to Dec. 2000
MOHC	HadGEM1	$1.875^\circ \times 1.25^\circ$	Jan. 1860 to Dec. 1999
NCAR	CCSM3	$1.4^\circ \times 1.4^\circ$	Jan. 1870 to Dec. 1999

Table 3.3: A list of observational and reanalysis datasets used in this report.

Institution	CMIP3 Name	Atmospheric Resolution (lon x lat)	Dates	Citation
UK Met Office	HadSST 3.1.1.0	$5^\circ \times 5^\circ$	Jan. 1850 to Aug. 2021	[40]
NOAA NCEP	OI SST V2	$1^\circ \times 1^\circ$	Sep. 1981 to Oct. 2021	[41]

Many of these datasets differ in spatial resolution, so, to simplify analysis, all data was interpolated to a common $1.5^\circ \times 1.5^\circ$ grid at monthly temporal resolution, within $50^\circ S - 50^\circ N$ and $120^\circ E - 100^\circ W$. The difference in time range was deemed unimportant. Whenever possible, the data used was an ensemble mean.

3.2 Empirical Orthogonal Function (EOF) Analysis

The mathematics behind the method are as follows: first, define the dataset as a matrix, X , where each column is a measurement at a specific time $t = 1 \dots N$, and each row is a different location [34]. Next, define the anomaly field, X' , by removing the time average. The goal is to find a vector with unit length, \mathbf{u} , which maximizes the variability of $X'\mathbf{u}$, which means solving the eigenvalue problem (7). These eigenvalues are normalized by their sum and explain variance due to each PC [34]. In theory, the next step is to simply define the covariance matrix $S = \frac{1}{N}X'^T X'$ and solve the eigenvalue problem:

$$S\mathbf{u} = \lambda^2\mathbf{u}, \quad (1)$$

but instead, a slightly easier computation is done: a Singular Value Decomposition (SVD) is used to decompose the anomaly matrix:

$$X' = A\Lambda U^T, \quad (2)$$

where the columns of the matrix A are the EOFs and the columns of the matrix U are corresponding time series, the PCs [34]. Additionally, the matrix Λ is diagonal by design, and each element on the diagonal is an eigenvalue [34]. This is more efficient because it is trivial to truncate the SVD wherever desired, which is valid as each succeeding EOF will account for less and less of the variance, as can be seen in Figure 3.1 [34].

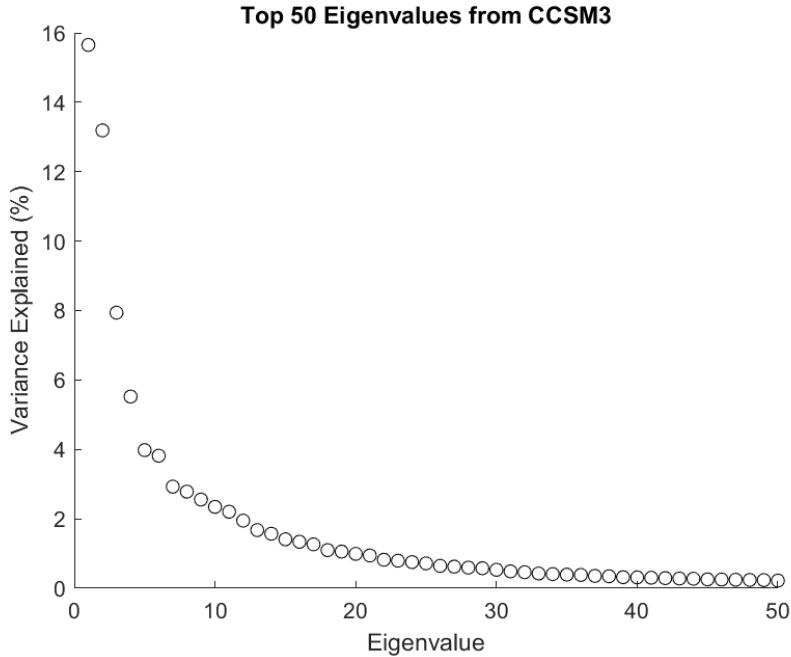


Figure 3.1: As an example, this figure shows the top 50 most significant eigenvalues of EOF analysis performed on CCSM3 with the global warming trend removed. As expected, each successive EOF accounts for less and less variability.

In addition, it is oftentimes necessary to remove global warming before computing EOFs, since it dominates the variance. For the example above, the EOF corresponding to global warming accounted for approximately 40% of the variance before removal. To do this removal, global warming was assumed to be a linear trend in time, and homogeneous in space, and then linearly de-trend the data. The time series according to each EOF can be analyzed as well by applying the Fast Fourier transform (FFT), then the periodicity can be compared to known values.

Generally it is also necessary to remove the effects of ENSO, as was done with global warming, before IPO can truly be investigated independently [4], [42]. Unfortunately, this is beyond the scope of this report; for example, Chen *et al* and Compo and Sardeshmukh provide methods which use complex regressions [4], [42]. Despite this, some attempt will be made to attribute EOFs to IPO.

3.3 IPO TPI Analysis

There is one major technique that can be used to confidently gauge IPO representation: the IPO Tripole Index (TPI) from Henley *et al* [1], [35]. Benefits include: having the original units, $^{\circ}C$, being tuned specifically for the spatial pattern of the IPO, not requiring detrending or making assumptions about global warming or ENSO [1], [35]. This is defined as the combination of mean SST anomalies over the following areas: T_1 corresponds to $25^{\circ}N - 45^{\circ}N$, $140^{\circ}E - 145^{\circ}W$; T_2 corresponds to $10^{\circ}S - 10^{\circ}N$, $170^{\circ}E - 90^{\circ}W$; and T_3 corresponds to $50^{\circ}S - 15^{\circ}S$, $150^{\circ}E - 160^{\circ}W$, shown in Figure 3.2 [1], [35]. These are combined:

$$TPI = T_2 - \frac{1}{2}(T_1 + T_3), \quad (3)$$

and then a low-pass Chebyshev filter using a 13 year cutoff period was applied, similar to Henley *et al* [1], [35]. They also discounted any IPO phases lasting less than 5 years, which will be done here as well [1], [35]. Then, the number of IPO phases per 100 years can be compared between each case.

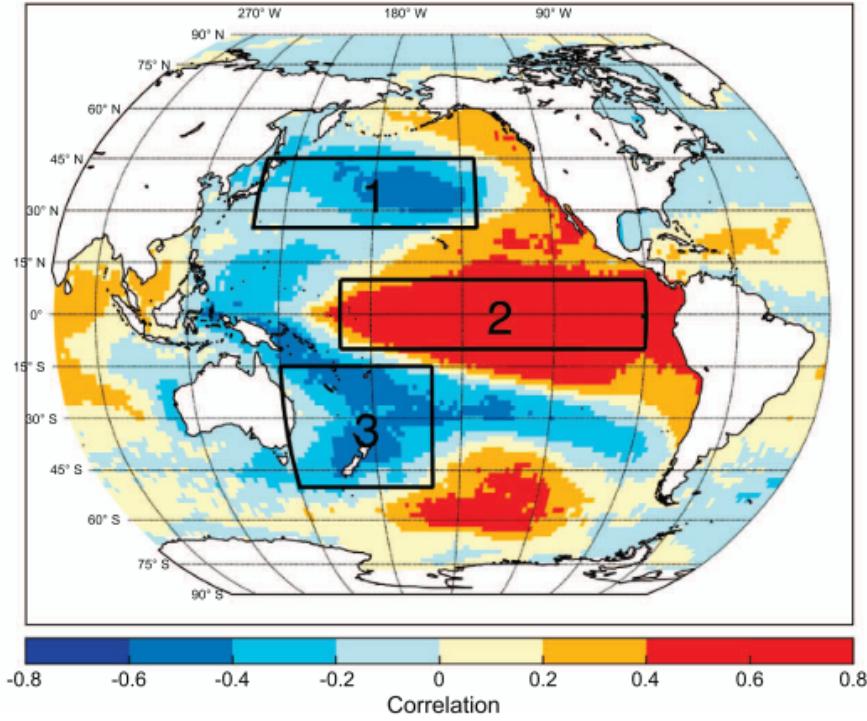


Figure 3.2: An illustration of TPI showing each area over which the mean SST anomalies are defined. The figure also displays spatial correlation, and the spatial pattern of the IPO. This figure is sourced from Henley *et al* [1].

4 Results

This section will focus on the results and of the computations described in the previous section, beginning with the EOFs. The first EOF of each of the CMIP3 models is displayed in Figure 4.1, where the generally expected spatial pattern is clearly visible, and each EOF accounts for approximately 10-20% of the variance in each dataset.

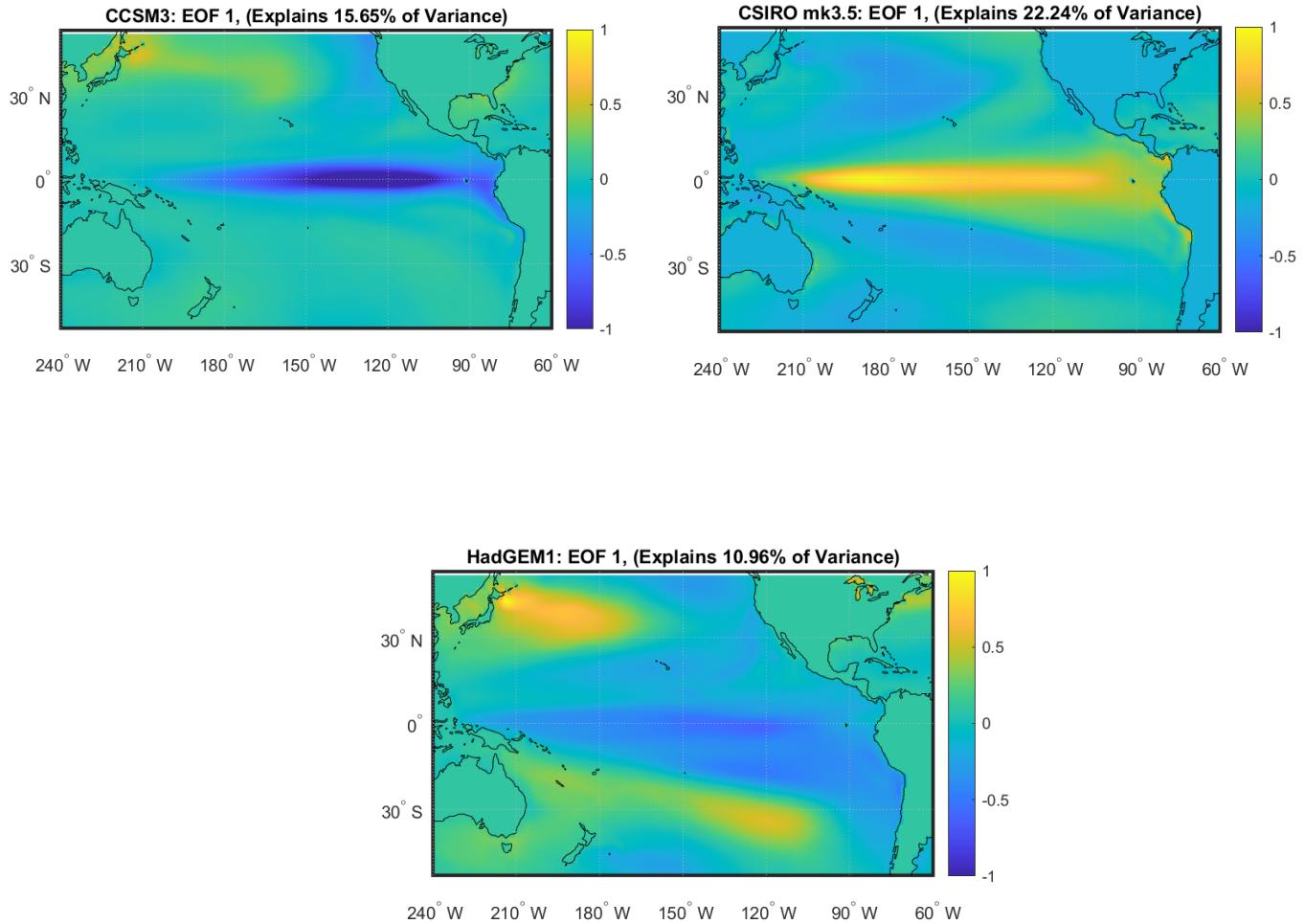


Figure 4.1: The first EOF of each of the CMIP3 models considered. All colorbars have been normalized to the same scale, corresponding to spatial correlation.

CCSM3 shows a narrower plume than would be expected, and HadGEM1 shows a lower magnitude but correctly shaped plume. As for the CMIP5 models, the first EOF of each is displayed in Figure 4.2, where once again each model displays a generally correct spatial pattern, and each accounting for approximately 15-20% of the variance from each dataset.

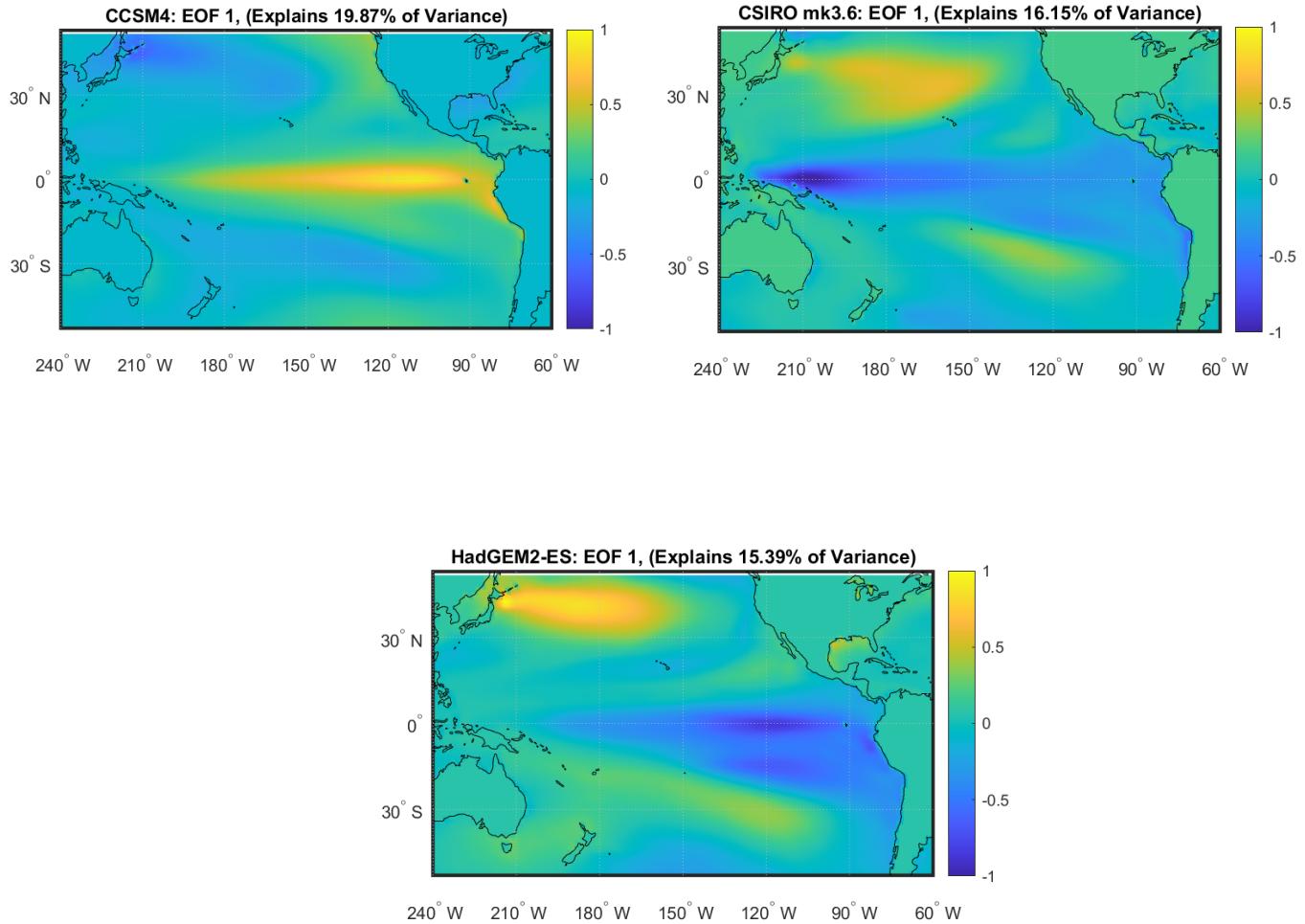


Figure 4.2: The first EOF of each of the CMIP5 models considered. All colorbars have been normalized to the same scale, corresponding to spatial correlation.

The CCSM4 and HadGEM2-ES plumes are once again well shaped, where the CSIRO mk3.6 plume is somewhat elongated.

The second EOF for each of the CMIP3 models is shown in Figure 4.3, where a rather diffuse pattern is observed, this EOF is likely unrelated to ENSO and IPO, and possibly the leftover nonlinear variance of global warming after having linearly de-trended the data.

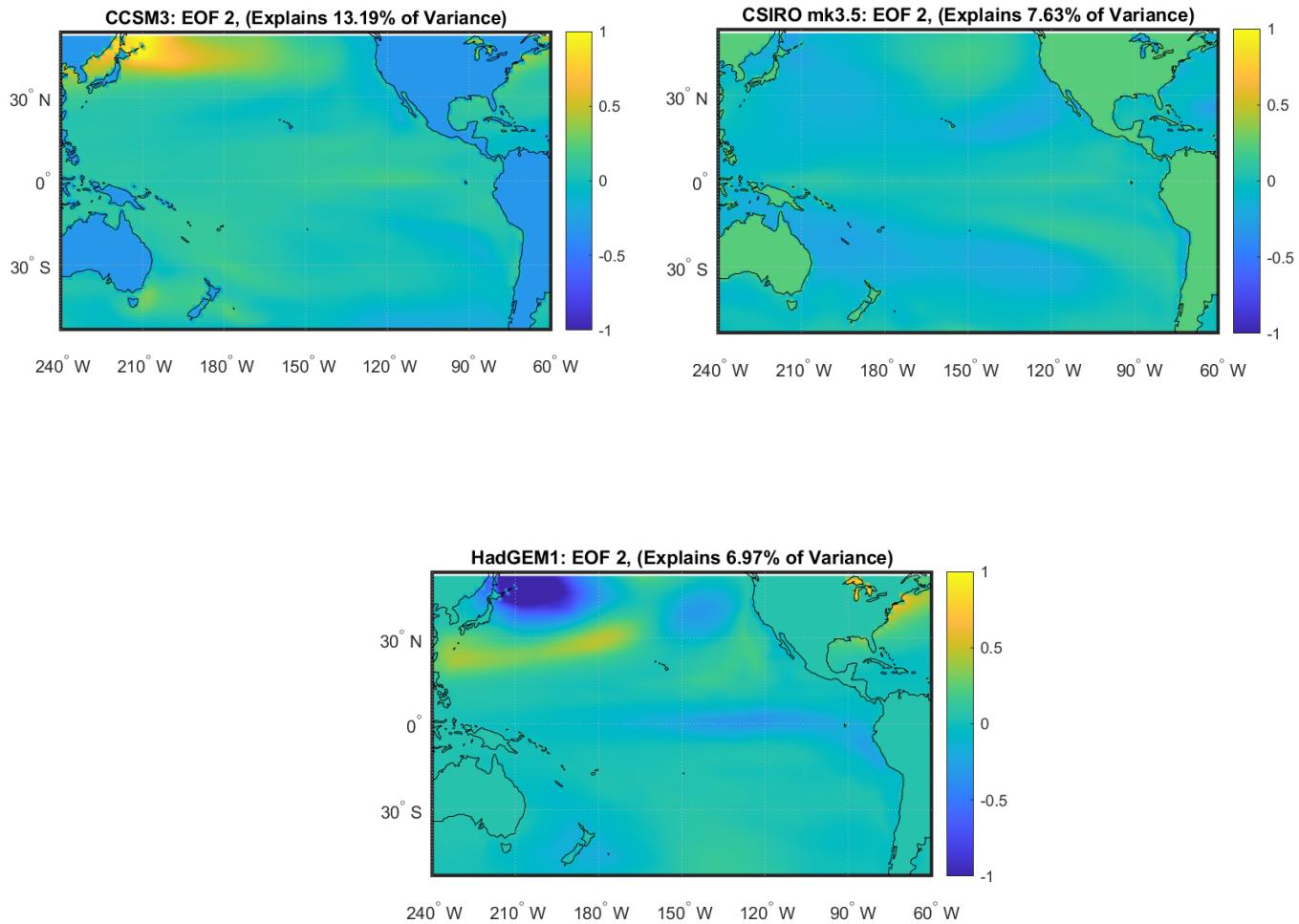


Figure 4.3: The second EOF of each of the CMIP3 models considered. All colorbars have been normalized to the same scale, corresponding to spatial correlation.

Much the same thing can be seen in the second EOF from CMIP5, as shown in Figure 4.4, where HadGEM2-ES also shows a slightly ENSO/IPO shaped plume at a low magnitude.

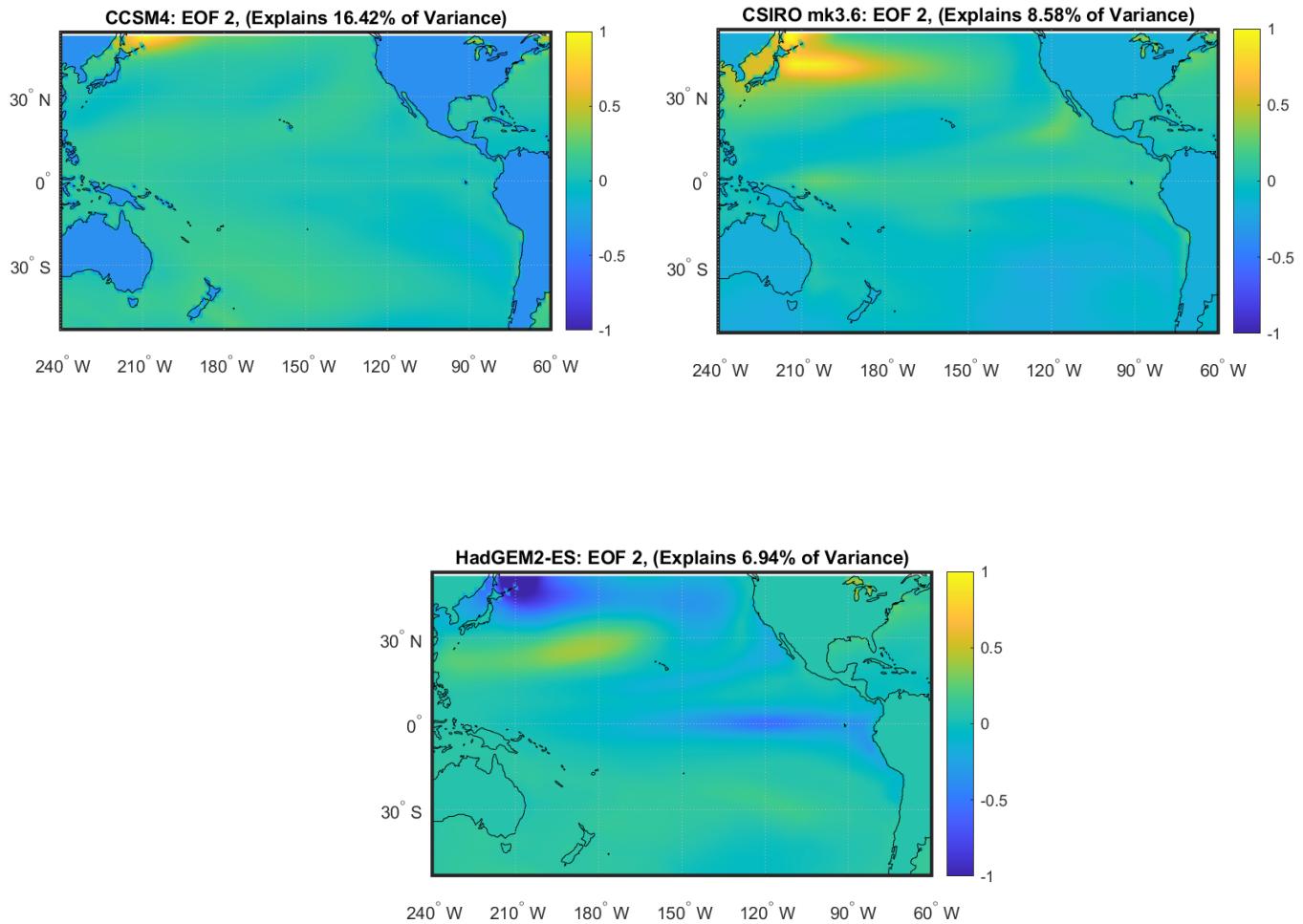


Figure 4.4: The second EOF of each of the CMIP5 models considered. All colorbars have been normalized to the same scale, corresponding to spatial correlation.

Moving on to the third EOFs of each model, those for CMIP3 are displayed in Figure 4.5, where some, like CCSM3 and HadGEM1, display ENSO-like spatial patterns, and CSIRO mk3.5 shows a pattern much like that of its second EOF.

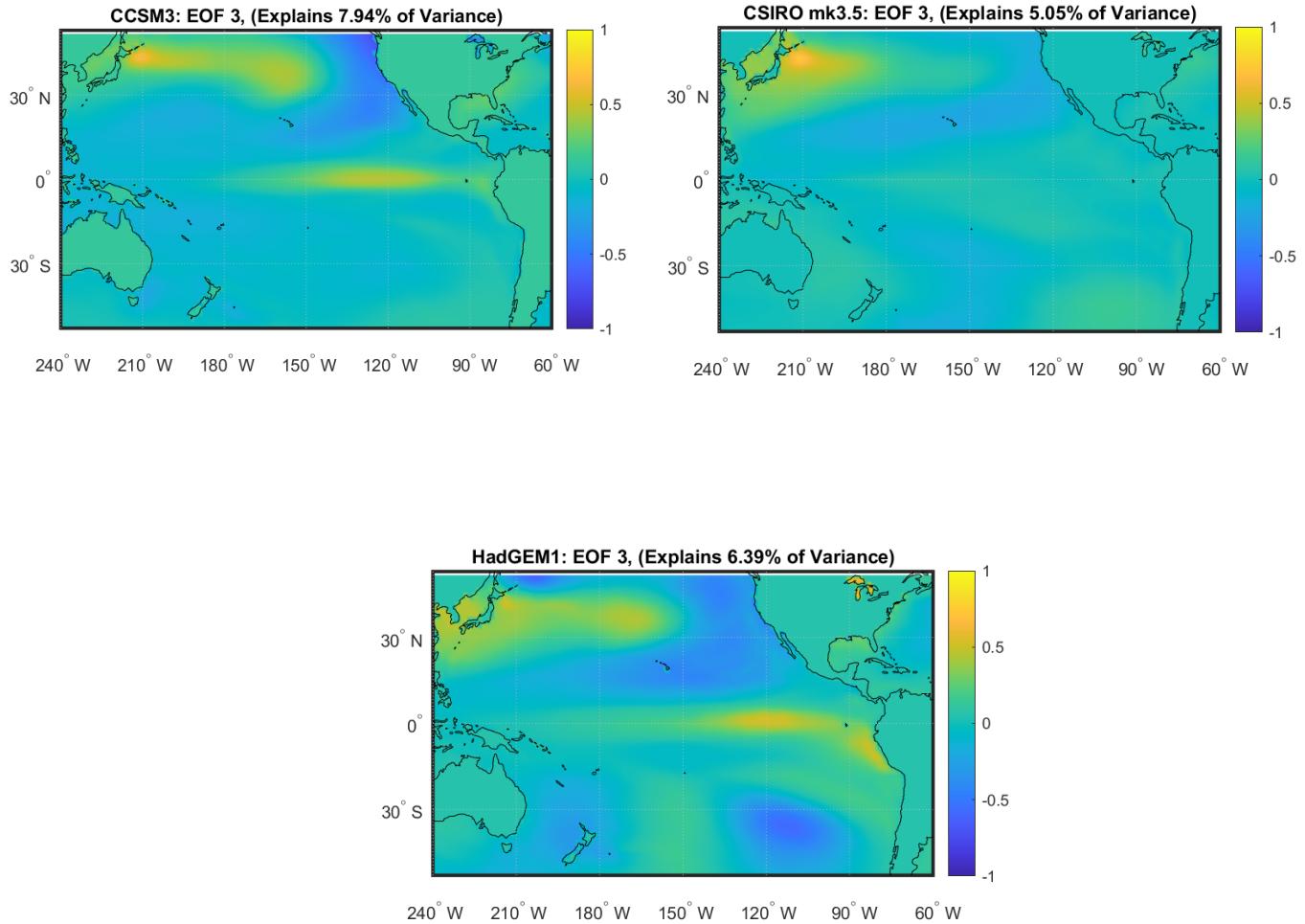


Figure 4.5: The third EOF of each of the CMIP3 models considered. All colorbars have been normalized to the same scale, corresponding to spatial correlation.

Next, the third EOF from the CMIP5 models is presented in Figure 4.6, where, once again, some weak ENSO/IPO-like spatial patterns are presented. The CSIRO mk3.6 model, much like in its first EOF, has a high correlation region in the west. Despite the variance explained decreasing the correlations off the coast of South America are more intense in the latter EOFs, which is observed in both generations of models.

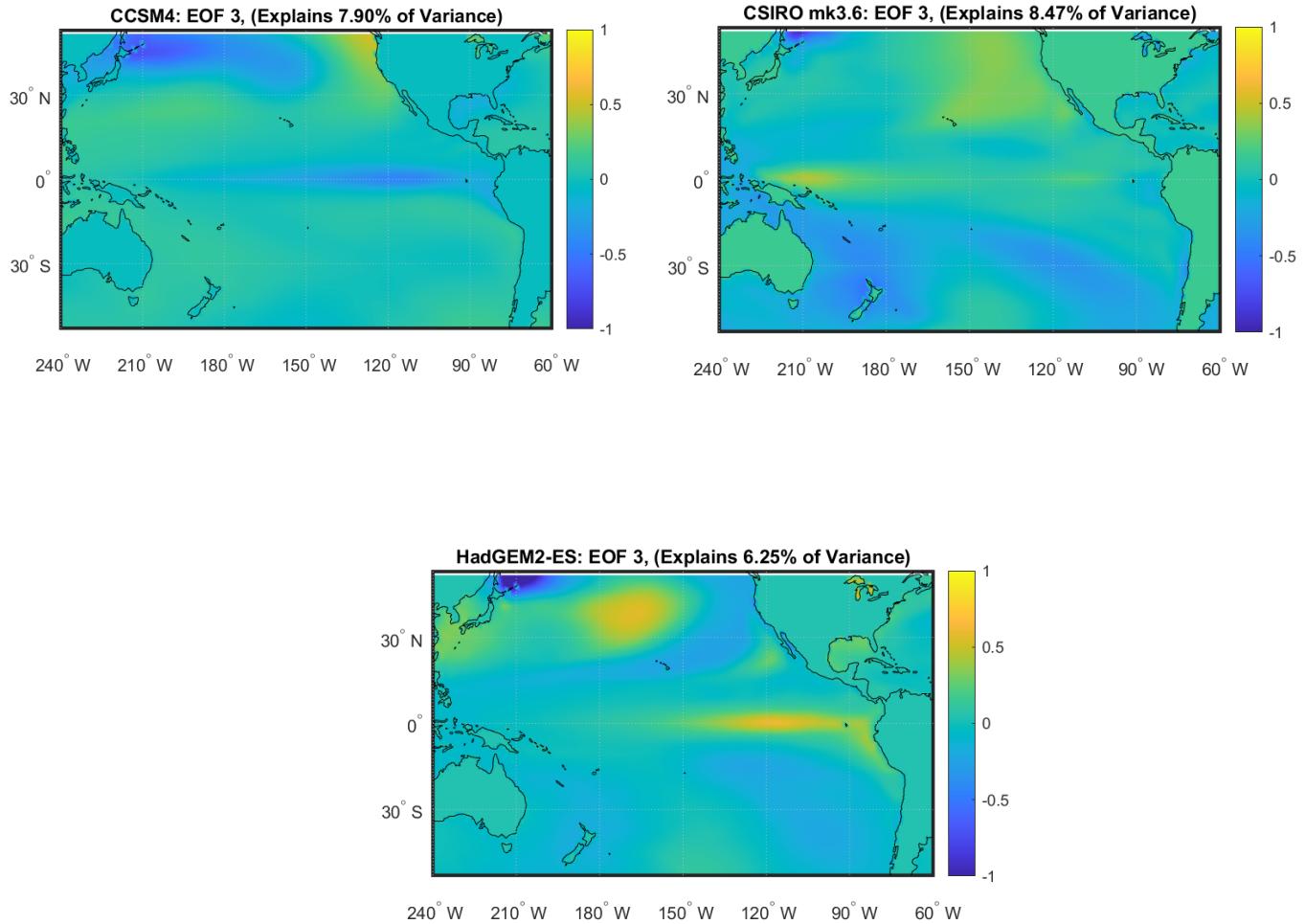


Figure 4.6: The third EOF of each of the CMIP5 models considered. All colorbars have been normalized to the same scale, corresponding to spatial correlation.

Overall, the first EOF in every case shows the desired spatial pattern of ENSO, the second EOF shows a pattern which could be related to the non-linear or non-homogeneous part of global warming, and the third EOF shows, in some cases, an ENSO/IPO-like pattern. The fourth EOF was in all cases similar in pattern to the third, and so is not shown here, although it will be present when analyzing the periodicity of each PC.

Next, these will be compared to the reanalysis (HadSST) and observational (NOAA) datasets, which are presented in Figure 4.7 and 4.8 respectively.

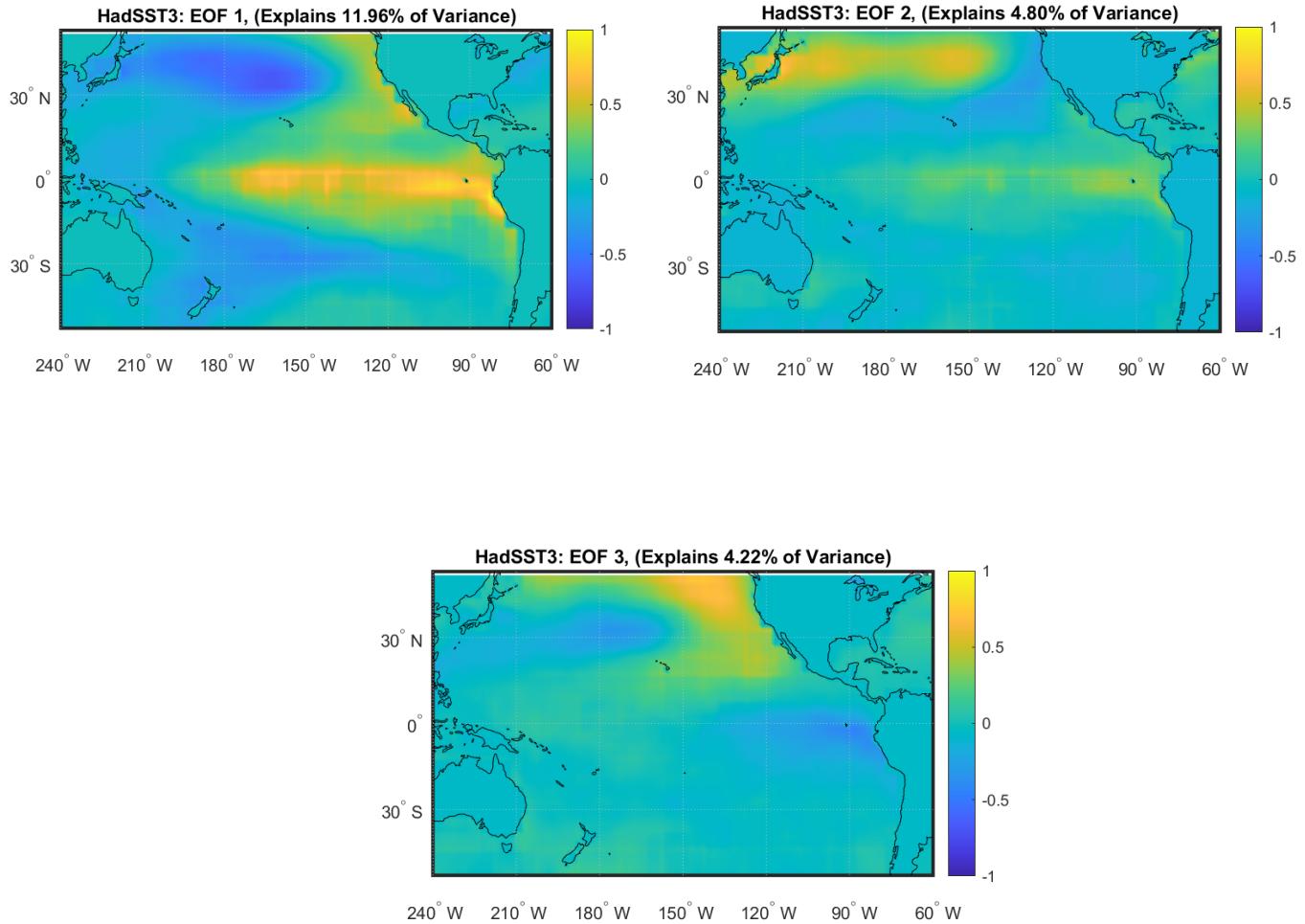


Figure 4.7: The first three EOFs of the HadSST3 reanalysis dataset. All colorbars have been normalized to the same scale, corresponding to spatial correlation. The larger pixels visible in some locations are due to the dataset being originally at a much lower resolution: $5^\circ \times 5^\circ$.

The reanalysis dataset shows the expected spatial shape for ENSO in the first and third EOFs, where the lack of expected pattern in the second EOF further supports that it is not simply a model artifact, but leftover global warming variations showing up. Interestingly, each EOF accounts for much less variance than for the other datasets here. As for the NOAA observational dataset, it can be seen that the first EOF has a very strong ENSO spatial pattern, and the third EOF has a weaker ENSO-like pattern. The second EOF is, once more, different from the expected pattern.

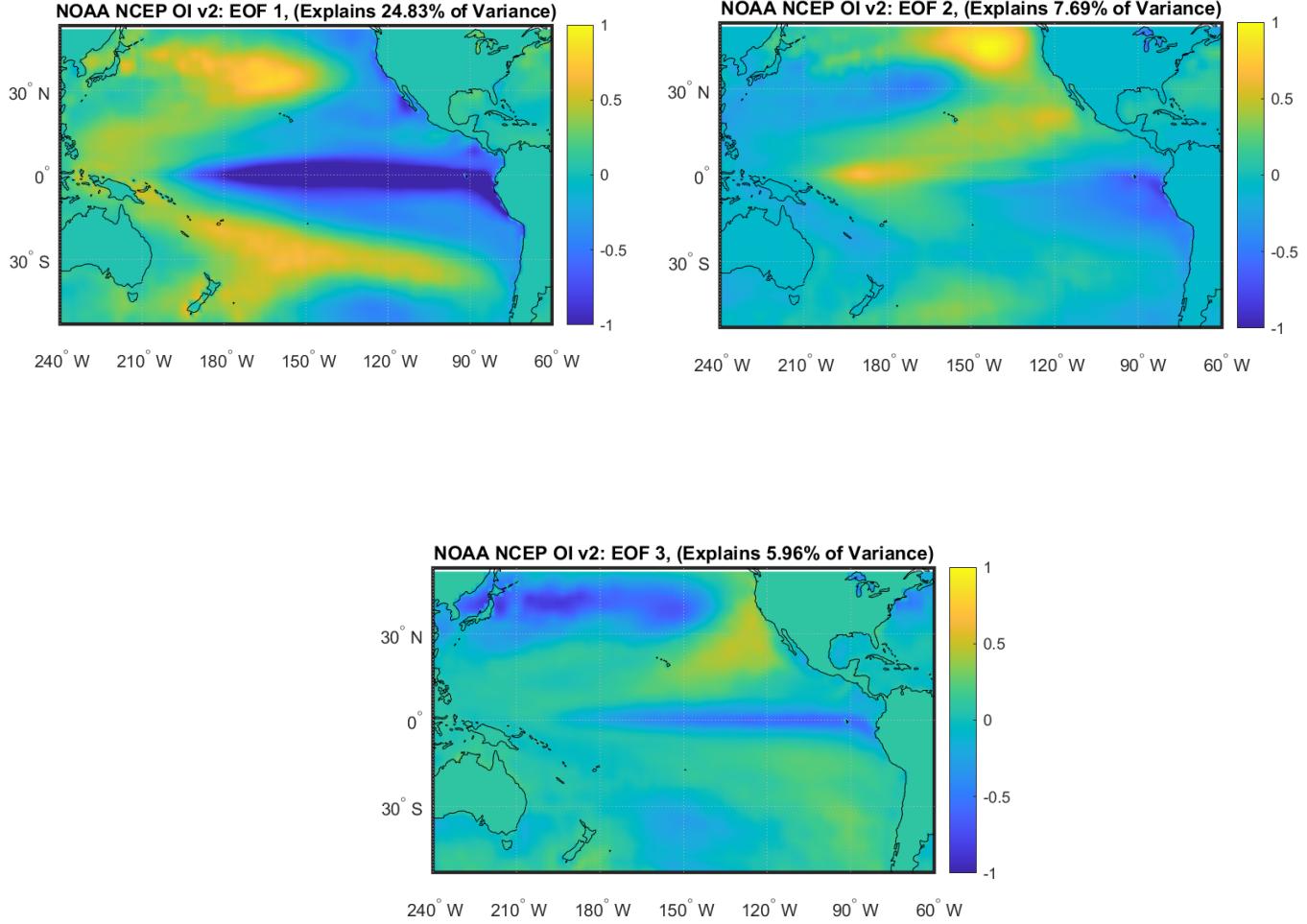


Figure 4.8: The first three EOFs of the NOAA observational dataset. All colorbars have been normalized to the same scale, corresponding to spatial correlation.

So, the CMIP3 and CMIP5 models generally match the pattern and approximately match the amount of variance represented in the reanalysis and observational datasets. Next, the PCs will be analyzed to understand how well these models/EOFs represent the periodicity of IPO and ENSO.

As detailed in Section 3, the Fourier transform of the PC time series can be analyzed to understand frequencies/periods associated with each EOF. Beginning with the CMIP3 models, the periodicity of each is presented in Figure 4.9, where, for CCSM3 and CSIRO mk3.5, there is some broad peaking around the 2-5 year mark in the first and third EOFs. In both cases, the first EOF also shows some peaks around the 6-9 year marks. As for HadGEM1, the first and third EOFs peak at years 3, 4, and 5, with the first EOF also

displaying a peak at exactly 10 years. This supports the hypothesis that the first and third EOFs are related to ENSO, as that matches the expected period of 2-7 years. However, all PCs also have considerable peaking in the interdecadal range, including where would be expected for IPO around 35 years.

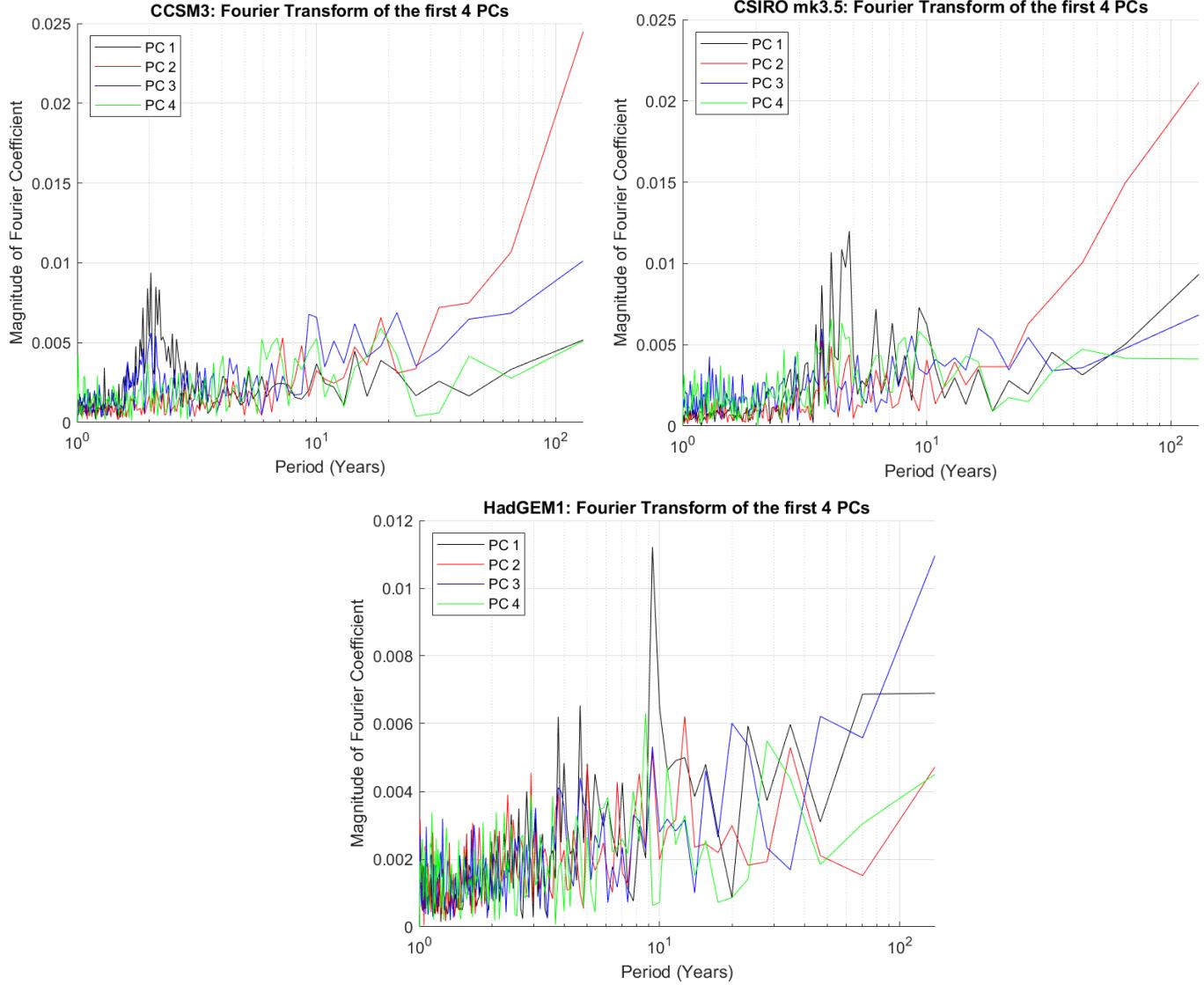


Figure 4.9: Each panel shows the FFT of the first four PCs corresponding to the first four EOFs for each CMIP3 model.

The peaking of the second PC right at the largest period supports the hypothesis that the second EOF is more related to global warming, at least for CCSM3 and CSIRO mk3.5, as the linear de-trending of the data would leave behind large low-frequency components, only removing up to first order effects. It would further seem that, for HadGEM1, PC2 is instead related to IPO and ENSO, and PC3 is related global warming, as each matches the respective patterns described above. Based on this, it would appear that, in all cases, EOF

1 is mostly representing ENSO, EOF2 is mostly representing global warming (EOF 3 for HadGEM1), and finally, EOF3 (EOF2 for HadGEM1) and EOF4 have mixed ENSO and IPO representation.

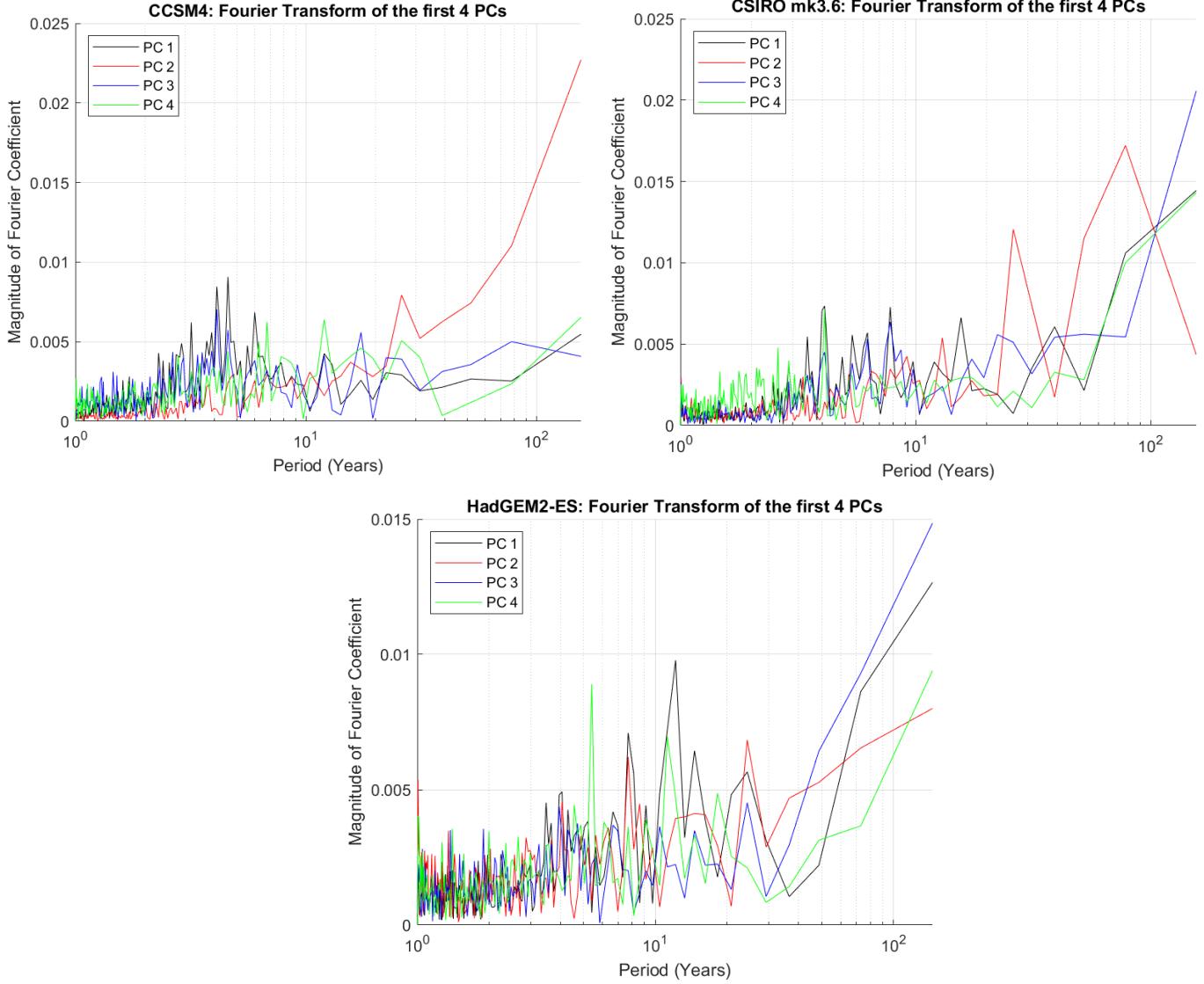


Figure 4.10: Each panel shows the FFT of the first four PCs corresponding to the first four EOFs for each CMIP5 model.

As for the CMIP5 generation of models, the FFT of the first four PCs are shown in Figure 4.10, where it can be seen that each CMIP5 model has a similar pattern to its CMIP3 counterpart. One notable exception is that CSIRO mk3.6 has less peaking from its first PC than its previous iteration, and its third PC peaks near the end similar to the global warming related PCs.

Finally, the FFTs of the first four PCs for the reanalysis and observational datasets

are available in Figure 4.11, where there is considerable peaking in the first PC in the 3-6 year range for both cases, where the second PC peaks near the end in both cases, and the third PC peaks around 30 years in both cases, although for the observational dataset it has a higher peak near the 10 year mark. The observational dataset is severely limited by data availability, so it would not necessarily be expected to see representation from IPO there. However, there are strong spikes around the expected period for ENSO, which does seem to be well represented here. It appears that the reanalysis dataset does a better job of representing both phenomena based on the FFT of the PCs and the EOFs.

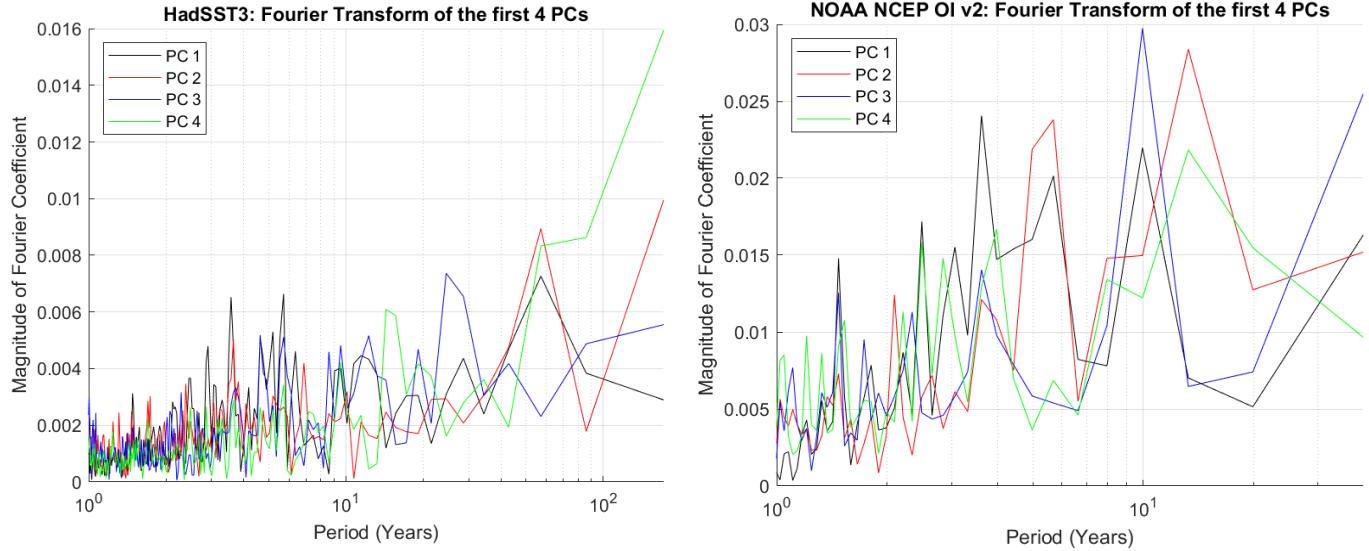


Figure 4.11: Each panel shows the FFT of the first four PCs corresponding to the first four EOFs for that dataset.

Finally, the TPI time series for each model from CMIP3 are presented in Figure 4.12; where, based on the filtered TPI, each model predicted 8.5, 4.6, and 7.9 IPO change events per 100 years respectively. Notably, CSIRO mk3.5 had much higher amplitude TPI than the other models.

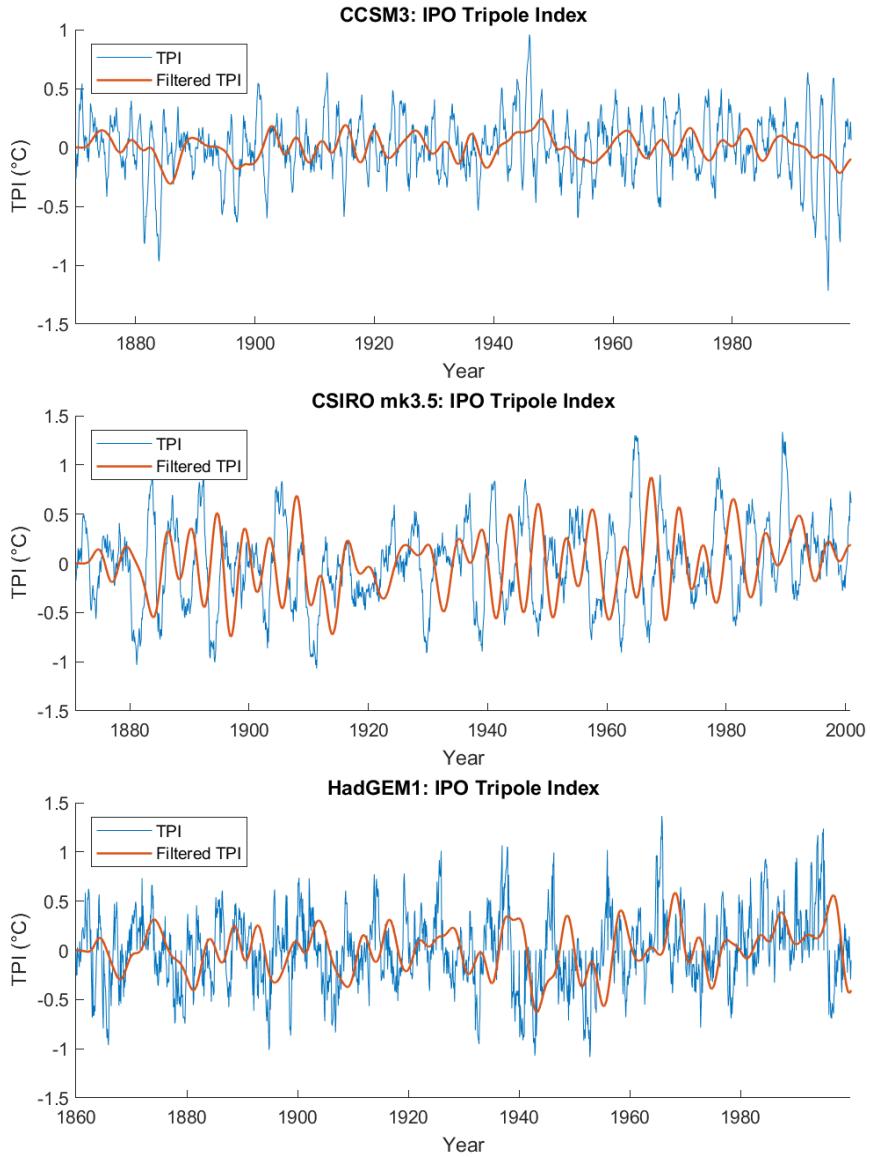


Figure 4.12: Each panel shows the filtered and unfiltered TPI for each of the CMIP3 models.

Next, the TPI time series for each model from CMIP5 are presented in Figure 4.13; where, based on the filtered TPI, each model predicted 3.2, 5.8, and 8.9 IPO change events per 100 years respectively.

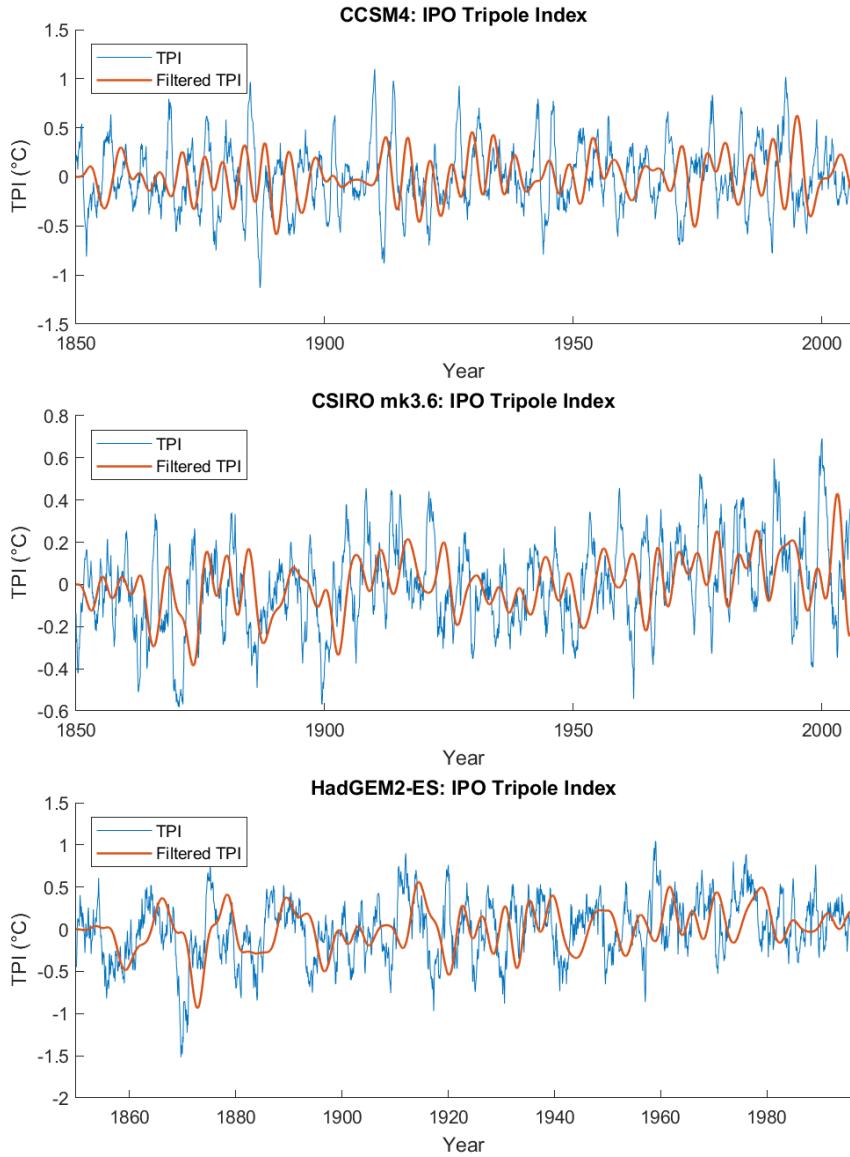


Figure 4.13: Each panel shows the filtered and unfiltered TPI for each of the CMIP5 models.

Lastly, the TPI time series for the reanalysis and observational datasets are presented in Figure 4.14; where, based on the filtered TPI, each dataset predicted 6.4 and 7.5 IPO change events per 100 years respectively.

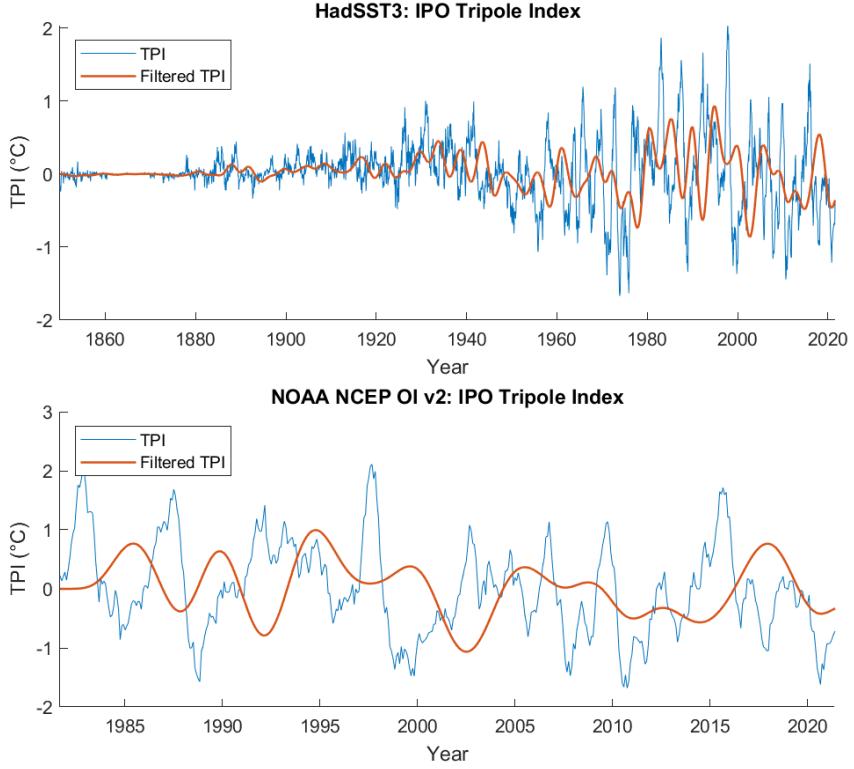


Figure 4.14: Each panel shows the filtered and unfiltered TPI for the reanalysis and observational datasets used.

On average, each generation of CMIP models included were quite close to the average of the reanalysis and observational datasets, with CMIP3 being slightly closer on average. The model predicting the closest number of IPO phase change events was HadGEM1. Interestingly, all CMIP models underestimate the amplitude of the TPI compared to the observational and reanalysis datasets, with CSIRO mk3.5 coming the closest to the true amplitude on average.

5 Discussion and Conclusions

At this point, the research questions outlined in Section 2 can be answered. In summary, for (RQ1), it was found that the first EOF in each case did a satisfactory job recreating those observed in the true data, specifically, the plumes off the west coast of South America; although most are missing the opposite phase spatial pattern surrounding that, which is most clearly seen in the NOAA dataset. Additionally, the amount of variance accounted for by the first EOF of each of the true datasets is higher than those from the CMIP models. In all cases, the third EOFs agreed well generally. As for the FFTs of the PCs, in all cases, these generally supported the first EOF being strongly linked to ENSO based on its associated periodicity matching that expected for ENSO; however, the peak periods were biased towards

the lower side of the expected range. Furthermore, these did show some amount of peaking around periods expected for IPO, suggesting the EOF analysis was unable to separate the two well. Finally, the TPI analysis showed that many of the models severely over- and under-estimated the number of IPO events per century, although the mean was close to the truth, this is attributed to sampling bias. Therefore, the answer to (RQ1) is: CMIP3 and CMIP5 models generally do a satisfactory job representing ENSO, with a tendency to underestimate the period and over-attribute variance to annual-scale variations; as for IPO, the models did a worse job representing this phenomenon, tending to attribute variations to much shorter time scales, and either over-estimating or under-estimating the number of IPO phase change events occurring. This conclusion agrees with those reached by Henley *et al* and Kociuba and Power, where both studies found that interannual variability was over-represented and interdecadal variability was under-represented, although the spatial patterns were well captured [1], [30]. Furthermore, the patterns thought to be related to IPO (EOF 3 for all models) generally do not agree well with those presented by Maher *et al* or Newman *et al*, both of whom show the IPO pattern as much more evenly distributed across the tropical Pacific, as opposed to the tongue-like shape found herein [9], [19].

As for (RQ2), the results obtained here for CMIP3 and CMIP5 models do strongly resemble those of observational and reanalysis datasets, but generally fail to capture the intricacies and details of those datasets, as was discussed briefly for (RQ1) above. The CMIP models generally matched the reanalysis dataset, HadSST3, better than the observational dataset, which could possibly be attributed to the significant difference in time scales considered. These results are supported in literature similarly to those for the previous research question.

Finally, for (RQ3), the answer is that CMIP5 models do not meaningfully improve over the CMIP3 models in their representations of ENSO and IPO, and, in some cases, the CMIP3 models appear to do a better job than their updated counterparts. This is because the spatial patterns are generally very similar between generations, and the periodicity associate with each PC are generally very similar as well, with the CMIP3 models having clearer peaks around periods expected for ENSO than the CMIP5 models had. As for IPO, the CMIP3 model mean of TPI events per century is closer to that of the observational/reanalysis datasets than that for CMIP5, but this is likely attributable to sampling bias. This is somewhat at odds with literature, as Polade *et al* found an improvement in IPO representation between these model generations [29]. However, Bellenger *et al* found modest improvement in some ENSO statistics between CMIP3 and CMIP5, although they attribute this to improved model tuning and not improved model physics [2].

There are many potential limitations to these analyses, first and foremost, the procedure of EOF analysis enforces both orthogonality between EOFs, and zero correlation between PCs, which is not reflective of physical processes, especially given that ENSO and IPO would not be expected to be uncorrelated [34]. Furthermore, the imposed orthogonality and zero correlation lead to difficulty interpreting the findings, as was seen when trying to attribute EOFs to IPO [34]. Finally, these impositions also introduce domain dependence into the analysis, which can hinder applicability and generalization of these findings [34].

And so, based on the relative simplicity of these analyses, their limitations, as well as the small number of models compared from each generation, these results are not able to be generalized to the full set of CMIP3 and CMIP5 models, despite the results generally agreeing with those found in literature. Potential modifications to rectify these might include analyzing several different spatial domains and comparing those or breaking the data into smaller time periods and analyzing those separately. Additionally, it could be desirable to do more investigation into TPI statistics, including mean, variance, and ratio of interdecadal to interannual variability, as was done in Henley *et al* [35].

References

- [1] B. J. Henley, G. Meehl, S. B. Power, *et al.*, “Spatial and temporal agreement in climate model simulations of the interdecadal pacific oscillation,” *Environmental Research Letters*, vol. 12, no. 4, p. 044 011, 2017. doi: [10.1088/1748-9326/aa5cc8](https://doi.org/10.1088/1748-9326/aa5cc8). [Online]. Available: <https://doi.org/10.1088/1748-9326/aa5cc8>.
- [2] H. Bellenger, E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, “ENSO representation in climate models: From CMIP3 to CMIP5,” *Climate Dynamics*, vol. 42, no. 7-8, pp. 1999–2018, Apr. 2013. doi: [10.1007/s00382-013-1783-z](https://doi.org/10.1007/s00382-013-1783-z). [Online]. Available: <https://doi.org/10.1007/s00382-013-1783-z>.
- [3] M. J. McPhaden, S. E. Zebiak, and M. H. Glantz, “ENSO as an integrating concept in earth science,” *Science*, vol. 314, no. 5806, pp. 1740–1745, Dec. 2006. doi: [10.1126/science.1132588](https://doi.org/10.1126/science.1132588). [Online]. Available: <https://doi.org/10.1126/science.1132588>.
- [4] J. Chen, A. D. D. Genio, B. E. Carlson, and M. G. Bosilovich, “The spatiotemporal structure of twentieth-century climate variations in observations and reanalyses. part II: Pacific pan-decadal variability,” *Journal of Climate*, vol. 21, no. 11, pp. 2634–2650, Jun. 2008. doi: [10.1175/2007jcli2012.1](https://doi.org/10.1175/2007jcli2012.1). [Online]. Available: <https://doi.org/10.1175/2007jcli2012.1>.
- [5] C. K. Folland, D. E. Parker, A. W. Colman, and R. Washington, “Large scale modes of ocean surface temperature since the late nineteenth century,” in *Beyond El Niño*, Springer Berlin Heidelberg, 1999, pp. 73–102. doi: [10.1007/978-3-642-58369-8_4](https://doi.org/10.1007/978-3-642-58369-8_4). [Online]. Available: https://doi.org/10.1007/978-3-642-58369-8%5C_4.
- [6] S. Power, T. Casey, C. Folland, A. Colman, and V. Mehta, “Inter-decadal modulation of the impact of ENSO on australia,” *Climate Dynamics*, vol. 15, no. 5, pp. 319–324, May 1999. doi: [10.1007/s003820050284](https://doi.org/10.1007/s003820050284). [Online]. Available: <https://doi.org/10.1007/s003820050284>.
- [7] N. E. Graham, “Decadal-scale climate variability in the tropical and north pacific during the 1970s and 1980s: Observations and model results,” *Climate Dynamics*, vol. 10, no. 3, pp. 135–162, Aug. 1994. doi: [10.1007/bf00210626](https://doi.org/10.1007/bf00210626). [Online]. Available: <https://doi.org/10.1007/bf00210626>.
- [8] K. E. Trenberth, “Recent observed interdecadal climate changes in the northern hemisphere,” *Bulletin of the American Meteorological Society*, vol. 71, no. 7, pp. 988–993, Jul. 1990. doi: [10.1175/1520-0477\(1990\)071<0988:roicci>2.0.co;2](https://doi.org/10.1175/1520-0477(1990)071<0988:roicci>2.0.co;2). [Online]. Available: [https://doi.org/10.1175/1520-0477\(1990\)071%3C0988:roicci%3E2.0.co;2](https://doi.org/10.1175/1520-0477(1990)071%3C0988:roicci%3E2.0.co;2).
- [9] N. Maher, A. S. Gupta, and M. H. England, “Drivers of decadal hiatus periods in the 20th and 21st centuries,” *Geophysical Research Letters*, vol. 41, no. 16, pp. 5978–5986, Aug. 2014. doi: [10.1002/2014gl060527](https://doi.org/10.1002/2014gl060527). [Online]. Available: <https://doi.org/10.1002/2014gl060527>.
- [10] E. Guilyardi, A. Wittenberg, A. Fedorov, *et al.*, “Understanding el niño in ocean–atmosphere general circulation models: Progress and challenges,” *Bulletin of the American Meteorological Society*, vol. 90, no. 3, pp. 325–340, Mar. 2009. doi: [10.1175/2008bams2387.1](https://doi.org/10.1175/2008bams2387.1). [Online]. Available: <https://doi.org/10.1175/2008bams2387.1>.

- [11] F.-F. Jin, S. T. Kim, and L. Bejarano, “A coupled-stability index for ENSO,” *Geophysical Research Letters*, vol. 33, no. 23, Dec. 2006. doi: [10.1029/2006gl027221](https://doi.org/10.1029/2006gl027221). [Online]. Available: <https://doi.org/10.1029/2006gl027221>.
- [12] J. Lloyd, E. Guilyardi, and H. Weller, “The role of atmosphere feedbacks during ENSO in the CMIP3 models. part III: The shortwave flux feedback,” *Journal of Climate*, vol. 25, no. 12, pp. 4275–4293, Jun. 2012. doi: [10.1175/jcli-d-11-00178.1](https://doi.org/10.1175/jcli-d-11-00178.1). [Online]. Available: <https://doi.org/10.1175/jcli-d-11-00178.1>.
- [13] G. A. Meehl, H. Teng, and J. M. Arblaster, “Climate model simulations of the observed early-2000s hiatus of global warming,” *Nature Climate Change*, vol. 4, no. 10, pp. 898–902, Sep. 2014. doi: [10.1038/nclimate2357](https://doi.org/10.1038/nclimate2357). [Online]. Available: <https://doi.org/10.1038/nclimate2357>.
- [14] G. A. Meehl and H. Teng, “CMIP5 multi-model hindcasts for the mid-1970s shift and early 2000s hiatus and predictions for 2016-2035,” *Geophysical Research Letters*, vol. 41, no. 5, pp. 1711–1716, Mar. 2014. doi: [10.1002/2014gl059256](https://doi.org/10.1002/2014gl059256). [Online]. Available: <https://doi.org/10.1002/2014gl059256>.
- [15] K. OSHIMA and Y. TANIMOTO, “An evaluation of reproducibility of the pacific decadal oscillation in the CMIP3 simulations,” *Journal of the Meteorological Society of Japan. Ser. II*, vol. 87, no. 4, pp. 755–770, 2009. doi: [10.2151/jmsj.87.755](https://doi.org/10.2151/jmsj.87.755). [Online]. Available: <https://doi.org/10.2151/jmsj.87.755>.
- [16] A. M. K. Stoner, K. Hayhoe, and D. J. Wuebbles, “Assessing general circulation model simulations of atmospheric teleconnection patterns,” *Journal of Climate*, vol. 22, no. 16, pp. 4348–4372, Aug. 2009. doi: [10.1175/2009jcli2577.1](https://doi.org/10.1175/2009jcli2577.1). [Online]. Available: <https://doi.org/10.1175/2009jcli2577.1>.
- [17] A. J., M. G., and M. A., “Interdecadal modulation of australian rainfall,” *Climate Dynamics*, vol. 18, no. 6, pp. 519–531, Feb. 2002. doi: [10.1007/s00382-001-0191-y](https://doi.org/10.1007/s00382-001-0191-y). [Online]. Available: <https://doi.org/10.1007/s00382-001-0191-y>.
- [18] G. A. Meehl and J. M. Arblaster, “Relating the strength of the tropospheric biennial oscillation (TBO) to the phase of the interdecadal pacific oscillation (IPO),” *Geophysical Research Letters*, vol. 39, no. 20, Oct. 2012. doi: [10.1029/2012gl053386](https://doi.org/10.1029/2012gl053386). [Online]. Available: <https://doi.org/10.1029/2012gl053386>.
- [19] M. Newman, M. A. Alexander, T. R. Ault, *et al.*, “The pacific decadal oscillation, revisited,” *Journal of Climate*, vol. 29, no. 12, pp. 4399–4427, Jun. 2016. doi: [10.1175/jcli-d-15-0508.1](https://doi.org/10.1175/jcli-d-15-0508.1). [Online]. Available: <https://doi.org/10.1175/jcli-d-15-0508.1>.
- [20] C. K. Folland, “Relative influences of the interdecadal pacific oscillation and ENSO on the south pacific convergence zone,” *Geophysical Research Letters*, vol. 29, no. 13, 2002. doi: [10.1029/2001gl014201](https://doi.org/10.1029/2001gl014201). [Online]. Available: <https://doi.org/10.1029/2001gl014201>.
- [21] E. Guilyardi, H. Bellenger, M. Collins, S. Ferrett, W. Cai, and A. Wittenberg, “A first look at enso in cmip5,” *CLIVAR Exchanges*, vol. 17, pp. 29–32, Feb. 2012.

- [22] E. Guilyardi, “El niño–mean state–seasonal cycle interactions in a multi-model ensemble,” *Climate Dynamics*, vol. 26, no. 4, pp. 329–348, Nov. 2005. doi: [10.1007/s00382-005-0084-6](https://doi.org/10.1007/s00382-005-0084-6). [Online]. Available: <https://doi.org/10.1007/s00382-005-0084-6>.
- [23] J.-Y. Yu and S. T. Kim, “Identification of central-pacific and eastern-pacific types of ENSO in CMIP3 models,” *Geophysical Research Letters*, vol. 37, no. 15, n/a–n/a, Aug. 2010. doi: [10.1029/2010gl044082](https://doi.org/10.1029/2010gl044082). [Online]. Available: <https://doi.org/10.1029/2010gl044082>.
- [24] K. AchutaRao and K. R. Sperber, “ENSO simulation in coupled ocean-atmosphere models: Are the current models better?” *Climate Dynamics*, vol. 27, no. 1, pp. 1–15, Feb. 2006. doi: [10.1007/s00382-006-0119-7](https://doi.org/10.1007/s00382-006-0119-7). [Online]. Available: <https://doi.org/10.1007/s00382-006-0119-7>.
- [25] J. Leloup, M. Lengaigne, and J.-P. Boulanger, “Twentieth century ENSO characteristics in the IPCC database,” *Climate Dynamics*, vol. 30, no. 2-3, pp. 277–291, Jul. 2007. doi: [10.1007/s00382-007-0284-3](https://doi.org/10.1007/s00382-007-0284-3). [Online]. Available: <https://doi.org/10.1007/s00382-007-0284-3>.
- [26] G. J. van Oldenborgh, S. Y. Philip, and M. Collins, “El niño in a changing climate: A multi-model study,” *Ocean Science*, vol. 1, no. 2, pp. 81–95, Oct. 2005. doi: [10.5194/os-1-81-2005](https://doi.org/10.5194/os-1-81-2005). [Online]. Available: <https://doi.org/10.5194/os-1-81-2005>.
- [27] S. Power, F. Delage, G. Wang, I. Smith, and G. Kociuba, “Apparent limitations in the ability of CMIP5 climate models to simulate recent multi-decadal change in surface temperature: Implications for global temperature projections,” *Climate Dynamics*, vol. 49, no. 1-2, pp. 53–69, Sep. 2016. doi: [10.1007/s00382-016-3326-x](https://doi.org/10.1007/s00382-016-3326-x). [Online]. Available: <https://doi.org/10.1007/s00382-016-3326-x>.
- [28] J. C. Furtado, E. D. Lorenzo, N. Schneider, and N. A. Bond, “North pacific decadal variability and climate change in the IPCC AR4 models,” *Journal of Climate*, vol. 24, no. 12, pp. 3049–3067, Jun. 2011. doi: [10.1175/2010jcli3584.1](https://doi.org/10.1175/2010jcli3584.1). [Online]. Available: <https://doi.org/10.1175/2010jcli3584.1>.
- [29] S. D. Polade, A. Gershunov, D. R. Cayan, M. D. Dettinger, and D. W. Pierce, “Natural climate variability and teleconnections to precipitation over the pacific-north american region in CMIP3 and CMIP5 models,” *Geophysical Research Letters*, vol. 40, no. 10, pp. 2296–2301, May 2013. doi: [10.1002/grl.50491](https://doi.org/10.1002/grl.50491). [Online]. Available: <https://doi.org/10.1002/grl.50491>.
- [30] G. Kociuba and S. B. Power, “Inability of CMIP5 models to simulate recent strengthening of the walker circulation: Implications for projections,” *Journal of Climate*, vol. 28, no. 1, pp. 20–35, Dec. 2014. doi: [10.1175/jcli-d-13-00752.1](https://doi.org/10.1175/jcli-d-13-00752.1). [Online]. Available: <https://doi.org/10.1175/jcli-d-13-00752.1>.
- [31] S. L. Lapp, J.-M. S. Jacques, E. M. Barrow, and D. J. Sauchyn, “GCM projections for the pacific decadal oscillation under greenhouse forcing for the early 21st century,” *International Journal of Climatology*, vol. 32, no. 9, pp. 1423–1442, May 2011. doi: [10.1002/joc.2364](https://doi.org/10.1002/joc.2364). [Online]. Available: <https://doi.org/10.1002/joc.2364>.

- [32] L. Dong, T. Zhou, and X. Chen, “Changes of pacific decadal variability in the twentieth century driven by internal variability, greenhouse gases, and aerosols,” *Geophysical Research Letters*, vol. 41, no. 23, pp. 8570–8577, Dec. 2014. doi: [10.1002/2014gl062269](https://doi.org/10.1002/2014gl062269). [Online]. Available: <https://doi.org/10.1002/2014gl062269>.
- [33] B. C. Weare, A. R. Navato, and R. E. Newell, “Empirical orthogonal analysis of pacific sea surface temperatures,” *Journal of Physical Oceanography*, vol. 6, no. 5, pp. 671–678, Sep. 1976. doi: [10.1175/1520-0485\(1976\)006<0671:eoao>2.0.co;2](https://doi.org/10.1175/1520-0485(1976)006<0671:eoao>2.0.co;2). [Online]. Available: [https://doi.org/10.1175/1520-0485\(1976\)006%3C0671:eoao%3E2.0.co;2](https://doi.org/10.1175/1520-0485(1976)006%3C0671:eoao%3E2.0.co;2).
- [34] A. Hannachi, I. T. Jolliffe, and D. B. Stephenson, “Empirical orthogonal functions and related techniques in atmospheric science: A review,” *International Journal of Climatology*, vol. 27, no. 9, pp. 1119–1152, 2007. doi: [10.1002/joc.1499](https://doi.org/10.1002/joc.1499). [Online]. Available: <https://doi.org/10.1002/joc.1499>.
- [35] B. J. Henley, J. Gergis, D. J. Karoly, S. Power, J. Kennedy, and C. K. Folland, “A tripole index for the interdecadal pacific oscillation,” *Climate Dynamics*, vol. 45, no. 11–12, pp. 3077–3090, Mar. 2015. doi: [10.1007/s00382-015-2525-1](https://doi.org/10.1007/s00382-015-2525-1). [Online]. Available: <https://doi.org/10.1007/s00382-015-2525-1>.
- [36] V. Trouet and G. J. V. Oldenborgh, “KNMI climate explorer: A web-based research tool for high-resolution paleoclimatology,” *Tree-Ring Research*, vol. 69, no. 1, pp. 3–13, Jan. 2013. doi: [10.3959/1536-1098-69.1.3](https://doi.org/10.3959/1536-1098-69.1.3). [Online]. Available: <https://doi.org/10.3959/1536-1098-69.1.3>.
- [37] L. D. Rotstayn, M. A. Collier, M. R. Dix, *et al.*, “Improved simulation of australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment,” *International Journal of Climatology*, n/a–n/a, 2009. doi: [10.1002/joc.1952](https://doi.org/10.1002/joc.1952). [Online]. Available: <https://doi.org/10.1002/joc.1952>.
- [38] W. J. Collins, N. Bellouin, M. Doutriaux-Boucher, *et al.*, “Development and evaluation of an earth-system model – HadGEM2,” *Geoscientific Model Development*, vol. 4, no. 4, pp. 1051–1075, Nov. 2011. doi: [10.5194/gmd-4-1051-2011](https://doi.org/10.5194/gmd-4-1051-2011). [Online]. Available: <https://doi.org/10.5194/gmd-4-1051-2011>.
- [39] S. J. Vavrus, M. M. Holland, A. Jahn, D. A. Bailey, and B. A. Blazey, “Twenty-first-century arctic climate change in CCSM4,” *Journal of Climate*, vol. 25, no. 8, pp. 2696–2710, Apr. 2012. doi: [10.1175/jcli-d-11-00220.1](https://doi.org/10.1175/jcli-d-11-00220.1). [Online]. Available: <https://doi.org/10.1175/jcli-d-11-00220.1>.
- [40] J. J. Kennedy, N. A. Rayner, C. P. Atkinson, and R. E. Killick, “An ensemble data set of sea surface temperature change from 1850: The met office hadley centre HadSST.4.0.0.0 data set,” *Journal of Geophysical Research: Atmospheres*, vol. 124, no. 14, pp. 7719–7763, Jul. 2019. doi: [10.1029/2018jd029867](https://doi.org/10.1029/2018jd029867). [Online]. Available: <https://doi.org/10.1029/2018jd029867>.
- [41] R. W. Reynolds, N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, “An improved in situ and satellite SST analysis for climate,” *Journal of Climate*, vol. 15, no. 13, pp. 1609–1625, Jul. 2002. doi: [10.1175/1520-0442\(2002\)015<1609:aiisas>2.0.co;2](https://doi.org/10.1175/1520-0442(2002)015<1609:aiisas>2.0.co;2). [Online]. Available: [https://doi.org/10.1175/1520-0442\(2002\)015%3C1609:aiisas%3E2.0.co;2](https://doi.org/10.1175/1520-0442(2002)015%3C1609:aiisas%3E2.0.co;2).

- [42] G. P. Compo and P. D. Sardeshmukh, “Removing ENSO-related variations from the climate record,” *Journal of Climate*, vol. 23, no. 8, pp. 1957–1978, Apr. 2010. doi: [10.1175/2009jcli2735.1](https://doi.org/10.1175/2009jcli2735.1). [Online]. Available: <https://doi.org/10.1175/2009jcli2735.1>.

6 Appendix A: Code and Details

The code used in this analysis and guide to obtaining the data is available on [Github](#).

7 Appendix B: Extended Methods

It pains the author that so much of the mathematical background had to be removed from the methods section to fit the word count, so a fuller version of that section is included here.

7.1 Empirical Orthogonal Function (EOF) Analysis

EOFs are a very powerful analytical tool which allows the representation of correlated fields via a relatively small set of functions (called EOFs), as well as their eigenvalues, which are taken to explain a certain amount of the variance in the original dataset due to each EOF [33], [34]. One advantage of this over other orthogonal function analysis techniques is the lack of a predefined form for the EOFs [33], [34]. Where, for example, Fourier analysis uses a linear combination of sine and cosine functions, EOFs have no such limitations; this is especially nice for analyzing SSTs (and climate in general), which do not have a known closed functional form and are subject to intricate boundary conditions [33], [34]. These facts have lead to EOF analysis being one of the most widely used tools in atmospheric and other climate-related sciences, and thus it is appropriate for use in this report [34]. Typically a weighting by latitude is applied, since the majority of these datasets are near the equator, weighting was not performed.

The mathematics behind the method are as follows: first, define the dataset as a collection of vectors forming a matrix:

$$X = (\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_N), \quad (4)$$

where each column is a vector representing a measurement at a specific time $t = 1\dots N$, and each entry in rows across these vectors is a different location [34]. Next, define the anomaly field by removing the time average:

$$X' = X - (1 \ 1 \ \dots \ 1) \bar{\mathbf{x}}, \quad (5)$$

Where the overbar defines a time average. The goal is to determine uncorrelated variables that explain the most variation and take linear combinations of those to reconstruct the data [34]. This corresponds to finding an expansion of the form:

$$\mathbf{x}'_t = \sum_{k=1}^r \lambda_k a_{tk} \mathbf{u}_k, \quad (6)$$

where \mathbf{x}'_t is the t 'th column of the anomaly matrix, a_{tk} is the tk 'th component of the EOF matrix, and \mathbf{u}_k is the k 'th PC. Mathematically, this means finding a vector with unit length, \mathbf{u} , which maximizes the variability of $X'\mathbf{u}$, that is, to solve the eigenvalue problem (7). These eigenvalues are normalized by their sum, and then interpreted as the amount of variance explained by each PC [34]. In theory, the next step is to simply define the covariance matrix $S = \frac{1}{N}X'X'$ and solve the eigenvalue problem:

$$S\mathbf{u} = \lambda^2\mathbf{u}, \quad (7)$$

but these computations are wasteful and unwieldy [34]. In practise, a Singular Value Decomposition (SVD) is used instead to decompose the anomaly matrix:

$$X' = A\Lambda U^T, \quad (8)$$

where the columns of the matrix A are the EOFs and the columns of the matrix U are corresponding time series, the PCs [34]. Additionally, the matrix Λ is diagonal by design, and each element on the diagonal is an eigenvalue corresponding to an EOF/PC [34]. This is more efficient because it is trivial to truncate the SVD wherever desired, which is valid as each succeeding EOF will account for less and less of the variance, as can be seen in Figure 3.1 [34].