

1 **Title Goes Here**

2 **Elio Campitelli · Leandro B. Díaz ·**  
3 **Carolina Vera ·**

4  
5 Received: date / Accepted: date

6 **Abstract** abstract  
7

8 **Keywords** ·

---

The research was supported by UBACyT20020170100428BA and the CLIMAX Project funded by Belmont Forum/ANR-15-JCL/-0002-01. Elio Campitelli was supported by a PhD grant from CONICET, Argentina.

---

Elio Campitelli

Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Ciencias de la Atmósfera y los Océanos. Buenos Aires, Argentina. CONICET – Universidad de Buenos Aires. Centro de Investigaciones del Mar y la Atmósfera (CIMA). Buenos Aires, Argentina. CNRS – IRD – CONICET – UBA. Instituto Franco-Argentino para el Estudio del Clima y sus Impactos (IRL 3351 IFAECI). Buenos Aires, Argentina.  
E-mail: [elio.campitelli@cima.fcen.uba.ar](mailto:elio.campitelli@cima.fcen.uba.ar)

Leandro B. Díaz

Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Ciencias de la Atmósfera y los Océanos. Buenos Aires, Argentina. CONICET – Universidad de Buenos Aires. Centro de Investigaciones del Mar y la Atmósfera (CIMA). Buenos Aires, Argentina. CNRS – IRD – CONICET – UBA. Instituto Franco-Argentino para el Estudio del Clima y sus Impactos (IRL 3351 IFAECI). Buenos Aires, Argentina.

Carolina Vera

Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Ciencias de la Atmósfera y los Océanos. Buenos Aires, Argentina. CONICET – Universidad de Buenos Aires. Centro de Investigaciones del Mar y la Atmósfera (CIMA). Buenos Aires, Argentina. CNRS – IRD – CONICET – UBA. Instituto Franco-Argentino para el Estudio del Clima y sus Impactos (IRL 3351 IFAECI). Buenos Aires, Argentina.

## 1 Introduction

introduction

One typical way of describing the zonally asymmetric circulation is by the amplitude and phase of zonal waves obtained by Fourier decomposition of geopotential height at each latitude (e.g. van Loon and Jenne, 1972; Trenberth, 1980). These methodologies rely on the assumption that the circulation can be meaningfully understood in terms of zonal waves of constant amplitude along a latitude circle. This is not valid for meridionally propagating waves or zonal waves with localised amplitudes. In the case of the wave 3, for example, Trenberth and Mo (1985) observed that wave 3 played a role in blocking events, but that in many cases, this was due to increased amplitude of a longitudinally localised wave train and not a hemispheric-scale zonal wave 3.

Another approach to characterising the Southern Hemisphere circulation is by using Empirical Orthogonal Functions (EOF, also known as Principal Component Analysis). Within the EOF framework, the Southern Annular Mode appears as the leading mode of variability of the Southern Hemisphere (SAM) (Fogt and Marshall, 2020) followed by the two Pacific–South American Patterns (PSA) (Mo and Paegle, 2001). The SAM represents a relatively zonally symmetric pattern of alternating low pressures in polar latitude and a ring of high pressures in high latitudes with an embedded wave 3 pattern that is more prominent in the Pacific sector. The PSA1 and PSA2 describe the two out-of-phase sides of a meridionally propagating wave train that originates in the Eastern equatorial Pacific and travels towards the South Atlantic following a great-circle arch along the Antarctic Peninsula. These patterns are derived by applying EOF to temporal anomalies, but Raphael (2003) applied EOF methods specifically to zonal anomalies.

EOFs are more flexible in the sense that they can characterise a richer set of spatial patterns, but they are restricted to standing oscillation modes and have difficulty dealing with propagating or phase-varying modes such as zonal waves. A single EOF can also represent a mixture of two or more physical modes.

The inability of a single EOF to describe travelling waves is what forces the PSA pattern to be described by two EOFs. Irving and Simmonds (2016) characterises the Pacific–South American Pattern in a Fourier framework. Being a meridionally propagating mode, the PSA cannot be correctly characterised by Fourier decomposition at each latitude circle, so they reprojected meridional wind fields so that the path of the PSA laid on the equator. On this new projection, they could identify the PSA using Fourier decomposition.

Finally, another, hybrid, methodology consist on identifying particular features of interest and creating indices using simple methods such as averages and differences. The Marshall SAM Index (Marshall, 2003), for example, tries to characterise the SAM by the average difference of sea level pressure between 40°S and 65°S, following Gong and Wang (1999). Raphael (2004) used a similar methodology to describe the wave 3 in the Southern Hemisphere. Instead of using Fourier to compute the amplitude and phase of the zonal wave, they averaged standardised geopotential height anomalies in three points rep-

resenting the location of the ridges of the zonal wave 3 of the climatological mean. This is equivalent to an index of the amplitude of the zonal wave 3 that projects into the phase of the zonal wave 3 of the mean field. Similarly, Hobbs and Raphael (2010) posit that the zonally asymmetric circulation in the Southern Hemisphere can be described by the strength and location of two anticyclones located in the sub-Antarctic western hemisphere.

These “feature-based” methods are grounded on other methods to identify the interesting features and can be useful to characterise features that are not readily apparent with Fourier or EOF methods but that nonetheless interesting or important. The use of data of single points, which makes them suitable for extending into the data-poor past also makes them not very robust.

In this paper we propose using complex Empirical Orthogonal Functions (Horel, 1984) as a robust extension of EOF analysis that is capable of describing rich structures with meridional and zonal propagation, and amplitude modulation such as the PSA. We apply this method to characterise the Spring-time zonally asymmetric circulation of the extratropical Southern Hemisphere.

## 2 Data and Methods

data-and-methods

### 2.1 Data

data

We used monthly geopotential height, air temperature, ozone mixing ratio, and Total Ozone Column (TOC) at  $2.5^\circ$  longitude by  $2.5^\circ$  latitude of horizontal resolution and 37 vertical isobaric levels from ERA5 (Hersbach et al, 2019) for the period 1979 to 2019. We restricted most of our analysis to the post-satellite era to avoid any confounding factors arising from the incorporation of satellite observations, but we also used the preliminary back extension of ERA5 from 1950 to 1978 (Bell et al, 2020) to look at long-term trends. We derived streamfunction at 200 hPa from ERA5 vorticity using the FORTRAN subroutine FISHPACK (Adams et al, 1999) and computed horizontal wave activity fluxes following Plumb (1985). For precipitation data we used monthly data from the CPC Merged Analysis of Precipitation (Xie and Arkin, 1997), with a  $2.5^\circ$  resolution in latitude and longitude. This rainfall gridded dataset is based on information from different sources such as rain gauge observations, satellite inferred estimations and the NCEP-NCAR reanalysis, and it is available since 1979 to present.

### 2.2 Methods

methods

We restrict our study to the September-October-November (SON) trimester. For that, we use seasonal means, averaging monthly values weighted by the number of days in each month. We also use 200 hPa to represent the high troposphere and 50 hPa to represent the lower stratosphere.

We computed the amplitude and phase of the zonal wave 1 by averaging (area-weighted) variables (temperature, geopotential height, ozone mixing

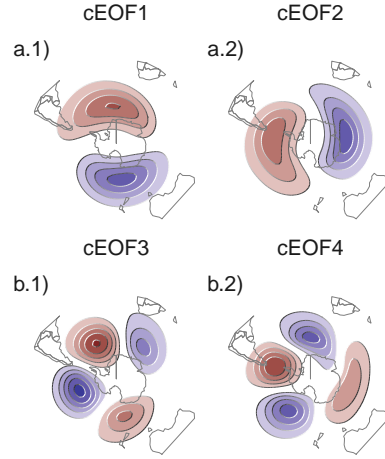


Fig. 1: Spatial patterns of the leading 4 EOFs of zonal anomalies of geopotential height at 50 hPa and 200 hPa for the SON trimester and the period 1979 – 2019 (arbitrary units).

ratio and Total Ozone Column) between 75°S and 45°S for each SON and extracting the wave-1 component of the Fourier spectrum. We chose this latitude band because it is wide enough to capture most of the relevant anomalies.

We computed the level-dependent Southern Annular Mode (SAM) index as the leading EOF of year-round monthly geopotential height anomalies south of 20°S at each level for the whole period (Baldwin and Thompson, 2009). We further split the SAM into its zonally symmetric and zonally asymmetric components (S-SAM and A-SAM indices respectively). These are obtained by projecting the zonally asymmetric and zonally symmetric part of the SAM spatial pattern onto monthly geopotential height fields (Campitelli et al, 2021). Since the analysis here is only for the SON trimester, monthly values were averaged across trimesters.

We also calculated the Pacific South American patterns (PSA1 and PSA2) as the third and fourth leading EOF of seasonal mean 500 hPa geopotential height with all seasons together following Mo and Paegle (2001).

Creo que habría que expandir un poco esta parte e ir comentando todas las metodologías que vas usando: regresiones, tendencias, cómo evaluar significancia en cada caso, etc.

## 2.3 Complex Empirical Orthogonal Functions (cEOF)

In traditional EOF analysis zonal waves appear as pairs of EOFs, usually degenerate, that represent similar patterns but shifted in phase (Horel, 1984). For instance, Figure 1 shows the leading 4 EOFs of zonally anomalous SON geopotential height at 50 hPa and 200 hPa. It is clear that at 50 hPa the first

Table 1: <sup>tab:corr-ceof-split</sup>Coefficient of determination ( $R^2$ ) of the absolute magnitude of complex EOFs between 200 hPa and 50 hPa computing EOF separately for each level.

200 hPa	50 hPa		
	cEOF1	cEOF2	cEOF3
cEOF1	0.28	0.02	0.02
cEOF2	0.00	0.60	0.02
cEOF3	0.00	0.00	0.01

two EOFs represent the same zonal wave 1 pattern and the last 2 represent the same zonal wave pattern with shorter wavenumber shifted by  $1/4$  wavelength. A similar EOF structure can be seen in 200 hPa (not shown). Since each pair of EOFs seem to represent the same phase-varying structure, it would be desirable to combine them into a single index with amplitude and phase.

Complex Empirical Orthogonal Functions (cEOF) is a useful method to characterise these waves (Horel, 1984). This method involves augmenting the original fields with the Hilbert transform of the data and then computing the Singular Value Decomposition of it. In this work, instead of applying the Hilbert transform to the time series at each point, we apply it to each latitude circle at each moment in time. Since the latitude circle is a periodic domain, this procedure does not suffer from edge effects. The result of the cEOF is a set of complex spatial patterns and complex time series. The real and imaginary part of each spatial pattern represent two phases wave-like spatial pattern that are in quadrature. The magnitude and argument of each complex time series represent the amplitude and phase of each zonal wave.

Table 1 shows the coefficient of determination between time series of the amplitude of each complex EOF (cEOF) across levels. There's a high degree of correlation between the respective cEOF1 and cEOF2 at each level. The spatial patterns of the 50 hPa and 200 hPa cEOFs are also similar (not shown)

Both the spatial pattern similarity and the high temporal correlation of cEOFs computed at 50 hPa and 200 hPa suggest that these are, to a large extent, modes of joint variability. This motivates the decision of performing complex EOF jointly between levels. The computation of the cEOFs was carried out using data from both levels at the same time, therefore, each cEOF has a spatial component that depends on longitude, latitude and level, and a temporal component that depends only on time.

The phase of principal components is defined up to an additive constant. For real principal components, this constant can be either 0 or  $\pi$ , corresponding to a change in sign. For complex principal components, it can be any number between 0 and  $2\pi$  (Horel, 1984). Since any choice is arbitrary and equally valid, we chose the phase of each cEOF so that the real and imaginary parts are aligned with meaningful phases in our analysis. This procedure does not create a spurious correlation, it only takes whatever relationship that already exist and aligns it with a specific phase.

For the first complex principal component, the phase was chosen so that the time series corresponding to the real part has the maximum correlation with the zonal wave 1 of Total Ozone Column between 75°S and 45°S. This also nearly minimises the correlation with the imaginary part.

For the second complex principal component, the phase was chosen so that the coefficient of determination between the Oceanic Niño Index (Bamston et al, 1997) and the real part was minimized, which also nearly maximizes the correlation with the imaginary part.

## 2.4 Computation procedures

### computation-procedures

We performed all analysis in this paper using the R programming language (R Core Team, 2020), using the data.table package (Dowle and Srinivasan, 2020) and the metR package (Campitelli, 2020). All graphics are made using ggplot2 (Wickham, 2009). We downloaded data from reanalysis using the ecmwfr package (Hufkens, 2020) and indices of the ENSO with the rsoi package (Albers and Campitelli, 2020). The paper was rendered using knitr and rmarkdown (Xie, 2015; Allaire et al, 2019).

## 3 Results

### results

### 3.1 cEOF patterns

#### ceof-patterns

To describe the spatial and temporal variability of the circulation zonal anomalies, the spatial and temporal parts of the first two leading cEOFs of zonal anomalies of geopotential height at 50 hPa and 200 hPa are shown in Figures 2 and 3. The first mode (CEOF1) explains 82% of the variance, while the second mode (CEOF2) explains a smaller fraction (7%). In the spatial patterns (Fig. 2), the real (in shading) and the imaginary components (in contour) are in quadrature by construction, so that each cEOF describe a single wave-like pattern whose amplitude and position (i.e. phase) is controlled by the magnitude and phase of the temporal cEOF. The wave patterns described by these cEOFs match the patterns seen in the traditional EOFs of Figure 1.

The first cEOF (Fig. 1 column a) is a hemispheric wave 1 pattern with maximum amplitude at high latitudes. At 50 hPa the Real cEOF1 has the maximum of the wave 1 at 150°E and at 200 hPa, the maximum is located at around 175°E indicating a westerly shift in phase.

The second cEOF (Fig. 1 column b) show also a zonal wave-like structure with maximum amplitude at high latitudes, but with shorter spatial scales. In particular, its dominant structure at both levels is a wave 3 but with larger amplitude in the pacific sector. This modulated amplitude is especially apparent at 200 hPa. There is no apparent phase shift, suggesting a rather barotropic structure.

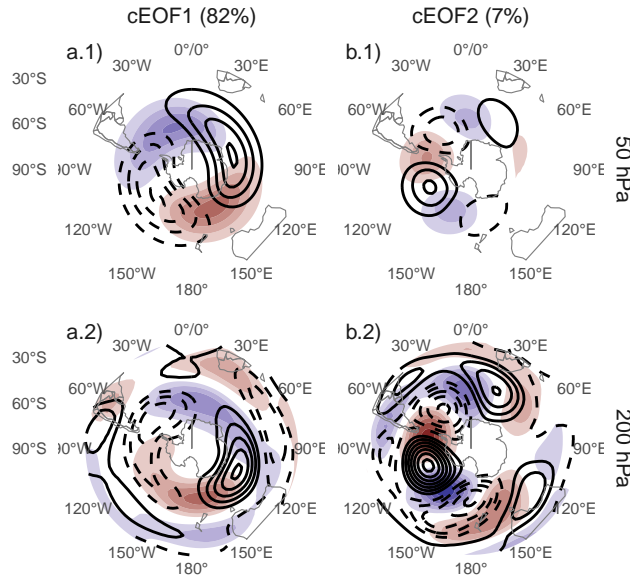


fig:ceofs-1  
 Fig. 2: Spatial patterns of the leading 2 cEOFs of zonal anomalies of geopotential height at 50 hPa and 200 hPa for the SON trimester and the period 1979 – 2019. Real part in shading, imaginary part in contours.

The temporal variability for both components of the cEOF1 have non zero mean (Fig. 3). This is due to the fact that the geopotential fields that enter into the cEOFs algorithm are anomalies with respect to the zonal mean, not the time mean. As a consequence, the variability associated with the first cEOF includes variability that projects onto the mean zonally anomalous field. There are no significant simultaneous correlation between the time series.

Both cEOF shot year-to-year variability but show no evidence of decadal variability. The relatively short timespan (40 years) hampers our ability to detect longer trends. Therefore, we extend the cEOF time series using the preliminary ERA5 back extension going back to 1950 by projecting the older geopotential height zonal anomaly fields onto the spatial patterns shown in Figure 2.

The extended cEOF timeseries are shown in Figure 4. There is a significant upward trend in the real component of cEOF1 (Fig. 4a.1,  $p$ -value  $< 0.001$ ) and no significant trend in any of the complex components of cEOF2. The positive trend in the real cEOF1 translates into a positive trend in cEOF1 magnitude, but not in the phase (not shown).

This long-term change indicates an increase in the magnitude of the high latitude zonal wave 1. A similar observation was made by Raphael (2003), who detected a step after around 1975 in the leading EOF of August-September-October 500 hPa zonal geopotential height anomalies.

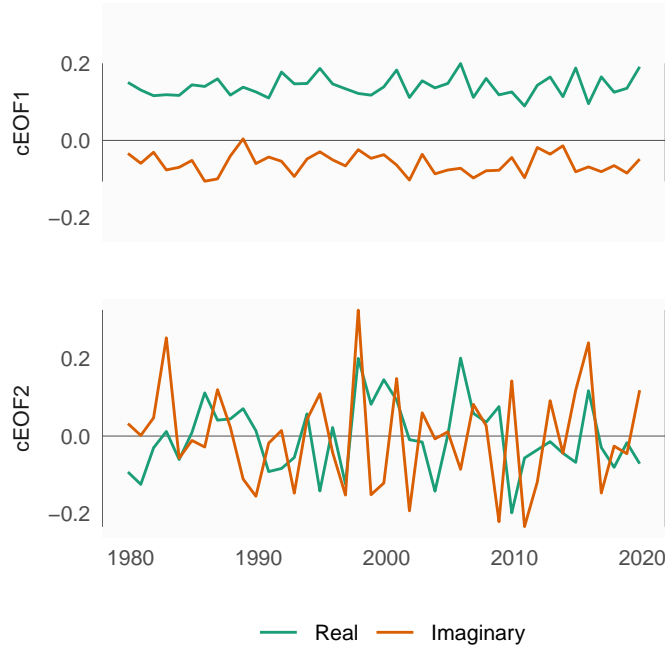


Fig. 3: Time series of the leading 2 cEOFs of zonal anomalies of geopotential height at 50 hPa and 200 hPa for the SON trimester and the period 1979 – 2019.

### 3.2 Regressions

#### regressions

The spatial patterns shown in Figure 2 are derived by removing the zonally symmetric circulation, so they might not include all the variability that is actually associated with the cEOF time series. To understand the full geopotential height field changes associated with each cEOF, we computed regression patterns of each cEOF with the complete geopotential fields.

Figure 5 shows regression patterns between geopotential height anomalies and cEOF1. At 50 hPa (Figure 5 row a), both the Real and Imaginary cEOF1 are associated with planetary wave 1 patterns, that are  $90^\circ$  off phase. Their phases coincide with the ones shown in Figure 2a.1, with the positive centre of the real cEOF1 located towards the dateline, and the one of the Imaginary cEOF1 located over Eastern Antarctica. However, the Real cEOF1 pattern is substantially altered by the zonally symmetric circulation. Instead of a clear wave 1 pattern, the regression pattern can be better describes as a monopole with its centre displaced from the South Pole.

The regression patterns at 200 hPa (Figure 5 row b) are similarly influenced by the zonally symmetric circulation. It is only possible to distinguish partially the wave 1 pattern in relation with the real cEOF1 (Figure 5b.1). The



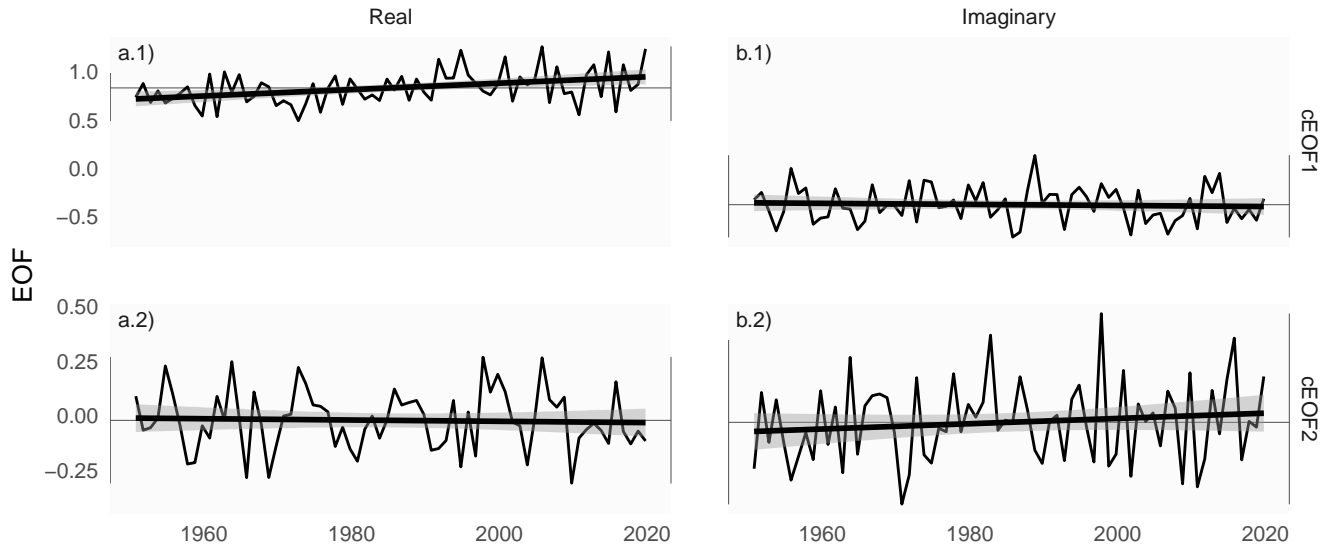


Fig. 4: Time series extended using ERA5 back extended preliminary edition (period 1950 – 1978) and ERA5 (period 1979 – 2019). Each series is computed by projecting monthly geopotential height zonal anomalies standardised by level south of 20°S onto the corresponding spatial pattern and scaled to zero mean and unit standard deviation.

fig:extended-series

imaginary cEOF1 shows a much more zonally symmetrical pattern resembling a negative Southern Annular Mode.

With the exception of the imaginary cEOF1, it is clear that these patterns are very different than the fully zonally asymmetric versions (Fig. 2), particularly at 200 hPa. Moreover, only in the stratosphere these patterns actually show a distinguishable wave 1 pattern shifted in phase by 90°. This should not be utterly surprising, since the spatial patterns of the *naïve* EOFs only show this feature at 50 hPa (Fig. 1), suggesting that using the cEOF method is artificially generating a wave structure at 200hPa. Therefore, the magnitude and phase of the cEOF1 are associated with the magnitude and phase of a zonal wave only in the stratosphere. While in the troposphere, they are associated with slightly off-centre monopoles.

Figure 6 shows the regression pattern of geopotential height and the cEOF2. Unlike for cEOF1, in this case the regression patterns are similar to the fully zonally asymmetric patterns from Figure 2. Although there are some differences (particularly in 50 hPa), the wave trains identified before are well characterised and patterns associated with the real cEOF2 are 90° out of phase with those associated with the imaginary cEOF2. Zonal wave 3 dominates all fields, but only in the western hemisphere, over the Pacific and Atlantic Oceans. cEOF2 then represents a equivalent barotropic wave train that is very similar to the the Pacific South American Patterns (Mo and Paegle, 2001).

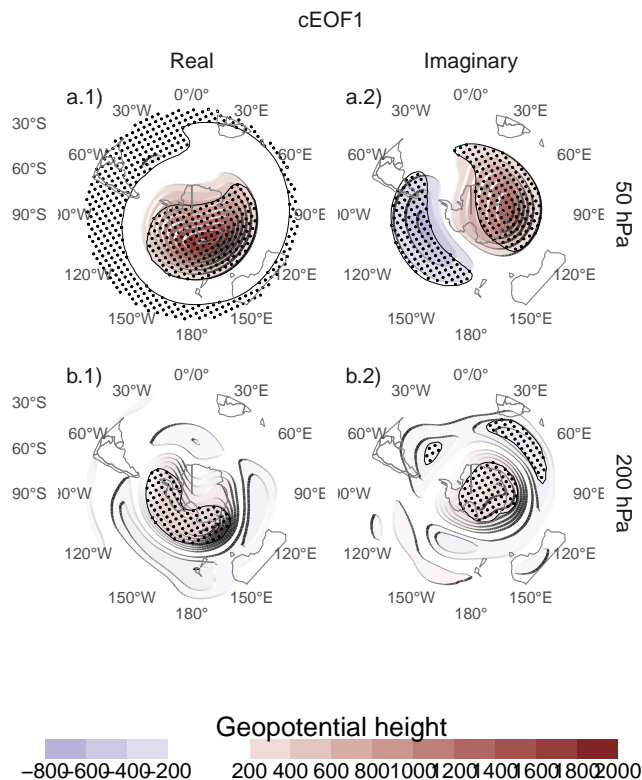


Fig. 5: Regression coefficients of the real and imaginary part of the first cEOF on SON geopotential height for the 1979 – 2019 period. These coefficients come from multiple linear regression involving the real and imaginary part of both cEOF2.

Comparing the location of the positive anomaly near 90°W in column b of Figure 6 with Figures 1.a and b from Mo and Paegle (2001), the real cEOF2 can roughly be identified with PSA2, while the imaginary cEOF2 resembles PSA1.

### 3.3 cEOFs relationship with other variables

ceofs-relationship-with-other-variables

#### 3.3.1 Temperature and Ozone

temperature-and-ozone

Figure 7 shows regression patterns of air temperature at 50hPa and 200hPa onto cEOF1. In both levels, the Real cEOF1 is associated with positive monopole over the South Pole with its centre moved slightly towards 150°E (Fig. 7 column 1). Regression maps on the Imaginary cEOF1, on the other hand, show a more clear wae 1 pattern with its maximum around 60°E. The distribution

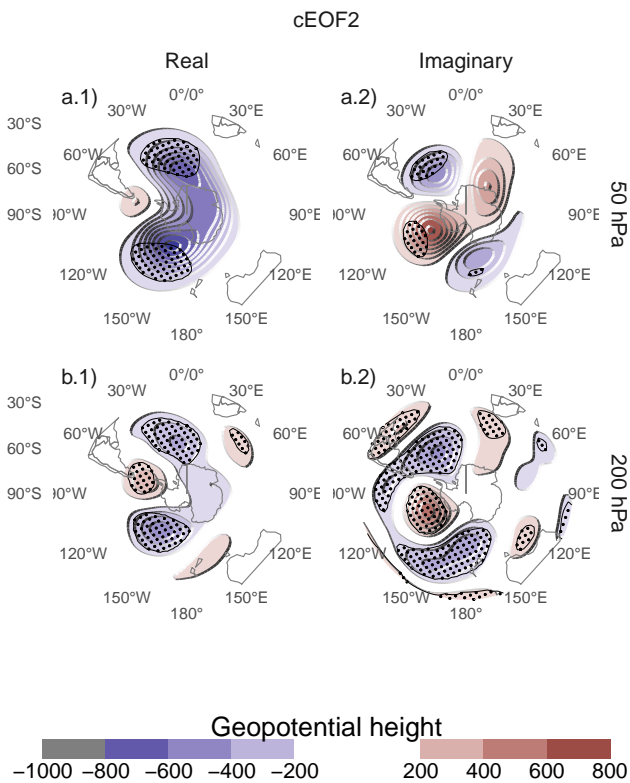


Fig. 6: Same as Figure 5 but for the second cEOF.

of temperature regression coefficients at 50 hPa and at 200 hPa mirror the geopotential height regression maps at 50 hPa (Fig. 5).

The vertical distribution of these patterns is shown in Figure 8) that shows the regression coefficients between cEOF1 and zonal anomalies of air temperature and zonal anomalies of ozone mixing ratio averaged between 75°S and 45°S. Temperature zonal anomalies associated with cEOF1 show a clear wave 1 pattern for both real and imaginary components throughout the atmosphere above 250 hPa with a change in sign above 10 hPa. Following hydrostatic balance, this is the level in which the geopotential anomaly have maximum amplitude (not shown).

The maximum ozone anomalies are co-located with the minimum temperature anomalies above 10 hPa and with the maximum temperature anomalies below 10 hPa (Fig. 8)); that is, the ozone zonal wave 1 is anticorrelated with the temperature zonal wave 1 in the upper stratosphere, and directly correlated in the upper stratosphere. This change in phase is observed in ozone anomalies forced by planetary waves that reach the stratosphere. In the photochemically-dominated upper stratosphere, cold temperatures inhibit the destruction of ozone, and the advectively-dominated lower stratosphere, ozone

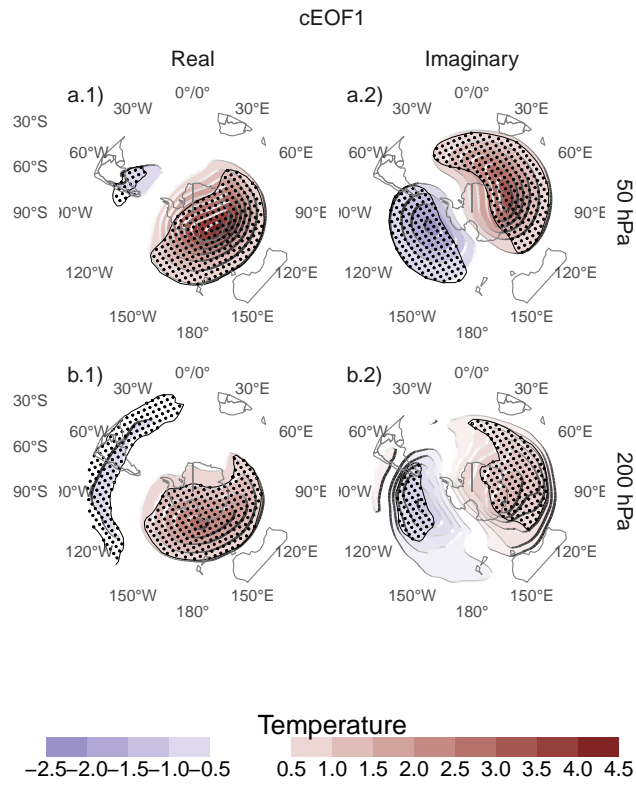


Fig. 7: Same as Figure 5 but for air temperature.

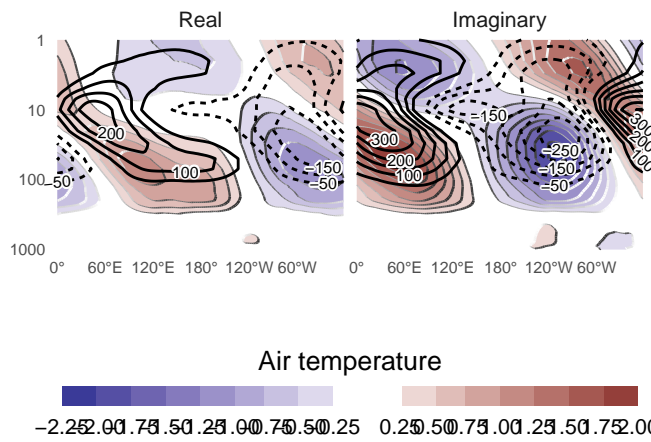


Fig. 8: Regression coefficients between cEOF1 and zonal anomalies of mean air temperature (shaded) and ozone mixing ratio (contours, negative contours with dashed lines, labels in parts per billion by mass) averaged between 75°S and 45°S for the period 1979 – 2019.

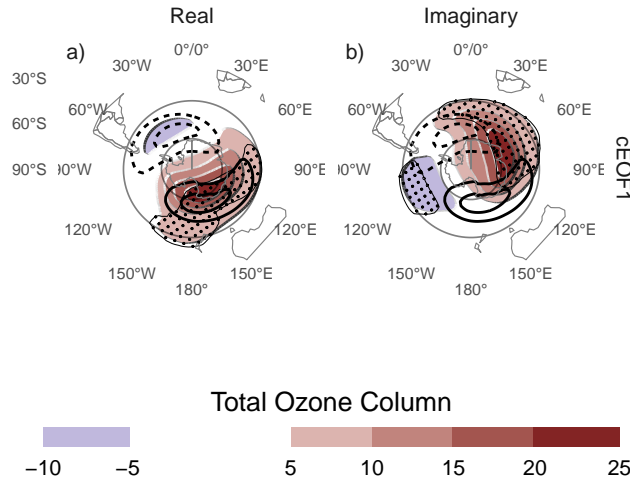


Fig. 9: Regression coefficients of SON mean Total Ozone Column anomalies onto the real (a) and imaginary (b) components of the cEOF1. On contours, the mean zonal anomaly of Total Ozone Column (negative contours in dashed lines, Dobson Units). For the 1979 – 2019 period.

anomalies are  $90^\circ$  out of phase with horizontal and vertical transport, which is  $90^\circ$  out of phase with temperature anomalies (Hartmann and Garcia, 1979; Wirth, 1993; Smith, 1995).

Figure 9 shows regression maps of cEOF1 with fields of Total Ozone Column (TOC). It shows zonal wave 1 patterns in TOC associated with both phases of cEOF1. Climatologically, the springtime Ozone minimum is located off the South Pole and towards the Weddell Sea (Wirth, 1993). Thus, the Real cEOF1 regression pattern (Figure 9a) coincides with the climatological position of the ozone hole while the one for the Imaginary cEOF1 is shifted by  $90^\circ$ .

Indeed, Figure 10 shows the relationship between amplitude and phase of the cEOF1 and the amplitude and phase of the planetary wave 1 in Total Ozone Column between  $75^\circ\text{S}$  and  $45^\circ\text{S}$ ; as shown in Figure 9, this is a wide latitude band that captures most of the area with significant regression coefficients. The correlation between the zonal wave 1 of ozone and cEOF1 (computed as the mean cosine of the difference in phase) is 0.86 (CI: 0.76 – 0.93). As expected from the location of anomalies in Figure 10, the Real cEOF1 drives the relationship with amplitude and the Imaginary cEOF1 drives the relationship with phase (not shown).

### 3.3.2 PSA

The relationship between PSAs and the modes described from the cEOFs is also studied. Figure 11 shows the relationship between the two PSA indices and the Real and Imaginary phase of cEOF2. As anticipated by Figure 6, there

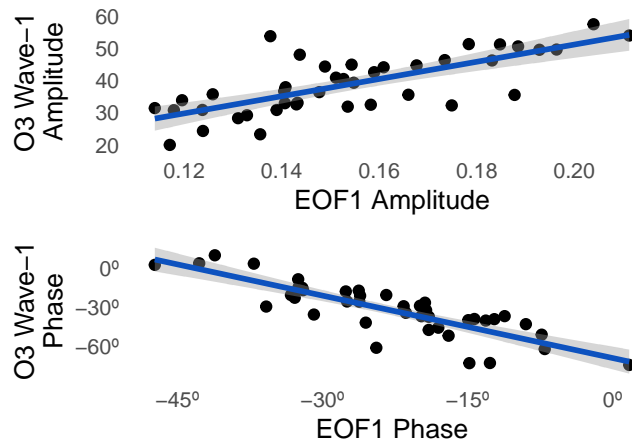


fig:wave1-o3  
 Fig. 10: Relationship between cEOF1 amplitude and phase with amplitude and phase of the zonal wave 1 in Total Ozone Column averaged between 75°S and 45°S for the period 1979 – 2019.

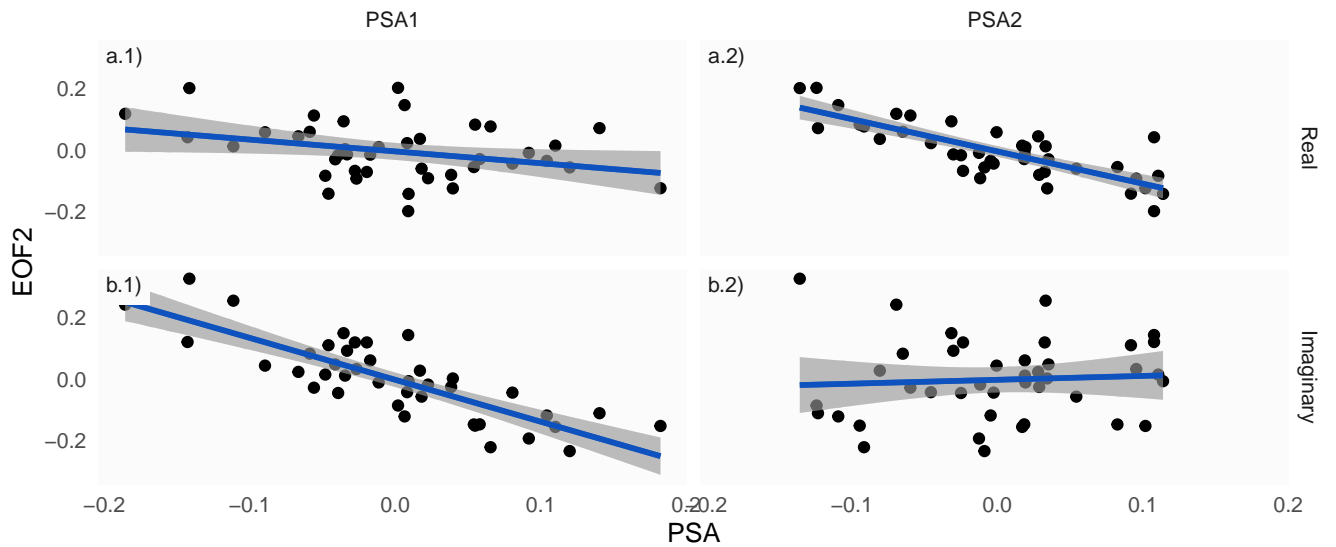


Fig. 11: Relationship between the Real and Imaginary parts of cEOF2 and the PSA1 and PSA2 modes computed as the second and third EOFs of seasonal geopotential height anomalies (Mo and Paegle, 2001). For the period 1979 – 2019.

fig:psa-eof2

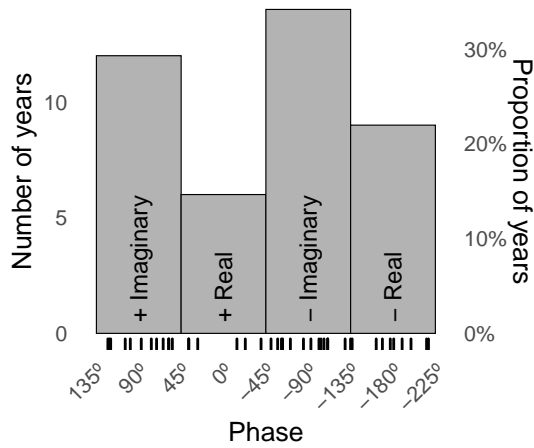


fig:phase-histogram  
 Fig. 12: Histogram of phase distribution of cEOF2 for the period 1979 – 2019. Bins are centred at  $90^\circ$ ,  $0^\circ$ ,  $-90^\circ$ ,  $-180^\circ$  with a binwidth of  $90^\circ$ .

is a strong correlation between PSA1 and Imaginary cEOF2 (Figure 11b.1), and between PSA2 and Real cEOF2 (Figure 11a.2). Conversely, there is no relationship between PSA1 and Real cEOF2, and between PSA2 and Imaginary cEOF2 (Figure 11 panels a.1 and b.2). So not only this cEOF2 represents well both the spatial structure and temporal evolution of the PSA modes, but it's also possible to make a rather clean association between its two phases and the two PSA modes.

This particular rotation of cEOF2 is the one which maximises the relationship between ENSO and Imaginary cEOF2. It is also the same that maximises this clean association between cEOF2 parts and PSA modes.

The reason the conventional EOF analysis arrives at the same separation than our particular cEOF rotation is probably the fact that not all phases are equal. Figure 12 is an histogram that counts the number of SON years in which the cEOF2 was near a particular phase with the observations marked as rugs on the horizontal axis. For instance, years with cEOF2 phase within  $45^\circ$  of  $0^\circ$  are nearest the “positive Real” phase. About two thirds of time cEOF2 has a phase similar to either the negative or positive Imaginary phase.

This phase-locking behaviour indicates that the Imaginary phase both is the direction that has the maximum relationship with ENSO and is also the most common phase. Therefore, the Imaginary cEOF2 explains more variance than the Real cEOF2 and conventional cEOF analysis will tend to separate the two.

This is similar to the method used by Irving and Simmonds (2016). From their Fourier decomposition of reprojected meridional wind fields, they show that the phase distribution is bimodal and isolate the PSA pattern from the rest of the PSA-like variability by selecting events that are near the peaks of the distribution (compare our Figure 12 with their Figure 6).

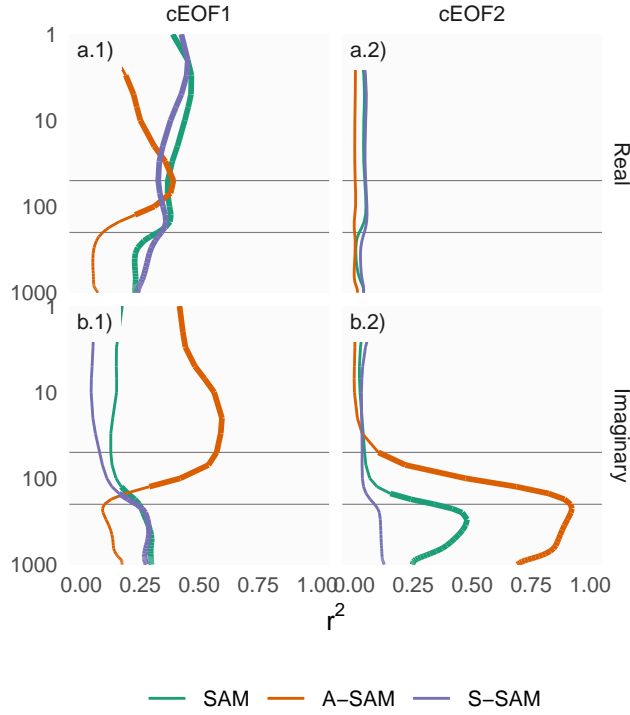


fig:sam-eof-vertical  
 Fig. 13: Coefficient of determination between the real and imaginary part of each cEOF and the SAM, Asymmetric SAM (A-SAM) and Symmetric SAM (S-SAM) indices computed at each level according to Campitelli et al (2021). Thick lines represent estimates with p-value < 0.01 corrected for False Detection Rate (Benjamini and Hochberg, 1995).

The advantage of our method is that it is much simpler to implement, it provides magnitude and phase naturally, and it facilitates the description of this mode as a propagating wave instead of as standing oscillation. As a consequence, the cEOF2 offers an alternative way of representing the PSA which has several advantages over using the second and third EOFs.

### 3.3.3 SAM

To explore the relationship between SAM and the modes described from the cEOF, we compute the coefficient of determination between the cEOFs time-series and the three SAM indices (SAM, A-SAM and S-SAM) at each level (Fig. 13). The SAM index is moderately correlated with the Real cEOF1 in all levels, and with the Imaginary cEOF1 and Imaginary cEOF2 in the troposphere. It is not correlated with the Real cEOF2 at any level.

The relationship between the tropospheric SAM and cEOF1 is explained entirely by the zonally symmetric component of the SAM as shown by the low



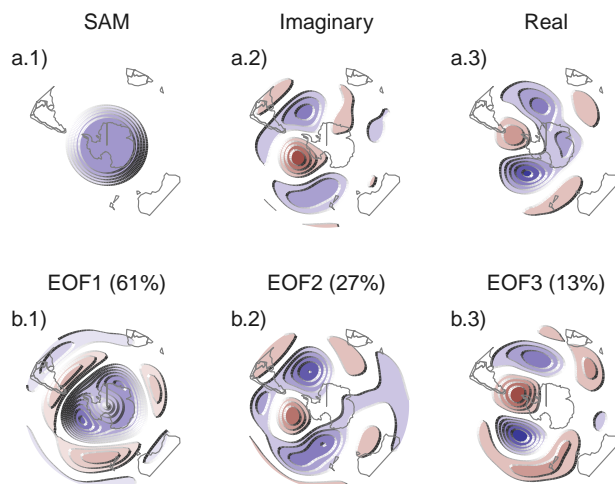


Fig. 14: Row a: The three components used to create a synthetic dataset of 41 years of geopotential height at 200 hPa. Row b: The three EOF computed from the synthetic dataset with the variance explained by each EOF in parenthesis. Units are arbitrary. Negative values in blue and positive values in red.

and statistically non-significant correlations between the A-SAM and either the Real or Imaginary cEOF1 and the high correlation with the S-SAM below 100 hPa. In the stratosphere, the Real cEOF1 is correlated with both A-SAM and S-SAM, while the Imaginary cEOF1 is highly correlated only with the A-SAM. These correlations are consistent with the regression patterns of geopotential height in Figure 5.

(La discusión de los próximos párrafos me resulta confusa. Deberías aclarar mejor por qué mostras cada cosa que mostras y cómo se relaciona con lo que mostraste antes. También me resulta como un resultado que se va un poco de tema, por eso creo que debería estar mejor contada esta parte para poder incluirla)

In the case of cEOF2, the moderate correlation between SAM and Imaginary cEOF2 becomes extremely high when only the zonally asymmetric variability of the SAM is considered. The Imaginary cEOF2 and the A-SAM share up to 92% of their variance, reached at 225 hPa (Figure 13.b2). Such extremely high correlation between the asymmetric SAM and the Imaginary cEOF2 suggests that these are not different features that share a causal link, but that they might be different ways of characterising the same phenomenon. And, in fact, suggests that the asymmetric part of the SAM might be statistical contamination from one phase of the PSA.

To test the hypothesis that a PSA-like feature can appear as statistic contamination of an otherwise zonally symmetrical and uncorrelated SAM, we create a synthetic dataset of 41 fields by reconstructing the 200 hPa geopotential field explained by the cEOF2 and adding a perfectly zonally symmetric

SAM-like structure with random sign and value. To this synthetic dataset we then apply traditional EOF. By construction this dataset contains two independent components: a zonally symmetric standing oscillation and a PSA of varying amplitude and phase (Figure 14 row a). The former one is a standing oscillation which could be represented by a single EOF, while the latter, being a wave which can change phase, needs two EOFs to be represented correctly. Thus, the full variance of this synthetic dataset can be perfectly explained by three EOFs.

The three EOFs of the synthetic dataset are shown in Figure 14 row b. The synthetic SAM can be seen in the first EOF while the PSA is explained by the second and third. However, even though the synthetic SAM is perfectly zonally symmetric by construction, the leading EOF is not. The EOF decomposition mixes this mode with parts of the zonally asymmetric cEOF2 pattern, resulting in a pattern eerily similar to the real SAM pattern. Of note, the three EOFs are well separated according to North's rule of thumb (North et al, 1982).

This suggests that the zonal asymmetries of the SAM could very well be mostly a statistical artifact. Since many surface impacts are mediated by the asymmetric component, as well as the relationship between SAM and ENSO (Campitelli et al, 2021), this potential issue could affect the interpretability of many results involving the SAM. Studies on the relationship between the SAM and the PSA pattern would be particularly difficult to interpret.

### 3.4 Tropical sources

#### tropical-sources

Figure 15 shows regression maps between cEOF2 and Sea Surface Temperatures (SST) Streamfunction at 200 hPa. The Imaginary cEOF2 is associated with strong positive SST anomalies on the Central Pacific and negative anomalies on an area across the North of Australia and New Zealand, the South Pacific Convergence Zone (SPCZ) (Figure 15.a2). This pattern is almost canonically positive ENSO and indeed, the correlation between the Imaginary cEOF2 and the Oceanic Niño Index (Barnston et al, 1997) is very high 0.76 (CI: 0.59 – 0.86). Streamfunction anomalies show a coherent picture. The Imaginary cEOF2 is associated with strong wave-like streamfunction anomalies emanating from the tropics (Figure 15.b2). This is consistent with what we know of the effect of ENSO on the extratropics: SST anomalies initiate anomalous convection that excites Rossby waves that propagate meridionally towards higher latitudes (REFS??).

Since the Real cEOF2 represents just a different phase of the same wave train, one would expect that it would show a similar forcing pattern to the Imaginary cEOF with a slight translation of its location. However, Figure 15.a1 and b1 show that the Real cEOF2 is not associated either with any significant SST nor streamfunction anomalies in the tropics. The correlation between the Real cEOF2 and ENSO is also not significant (-0.0017 (CI: -0.31 – 0.31)) (No

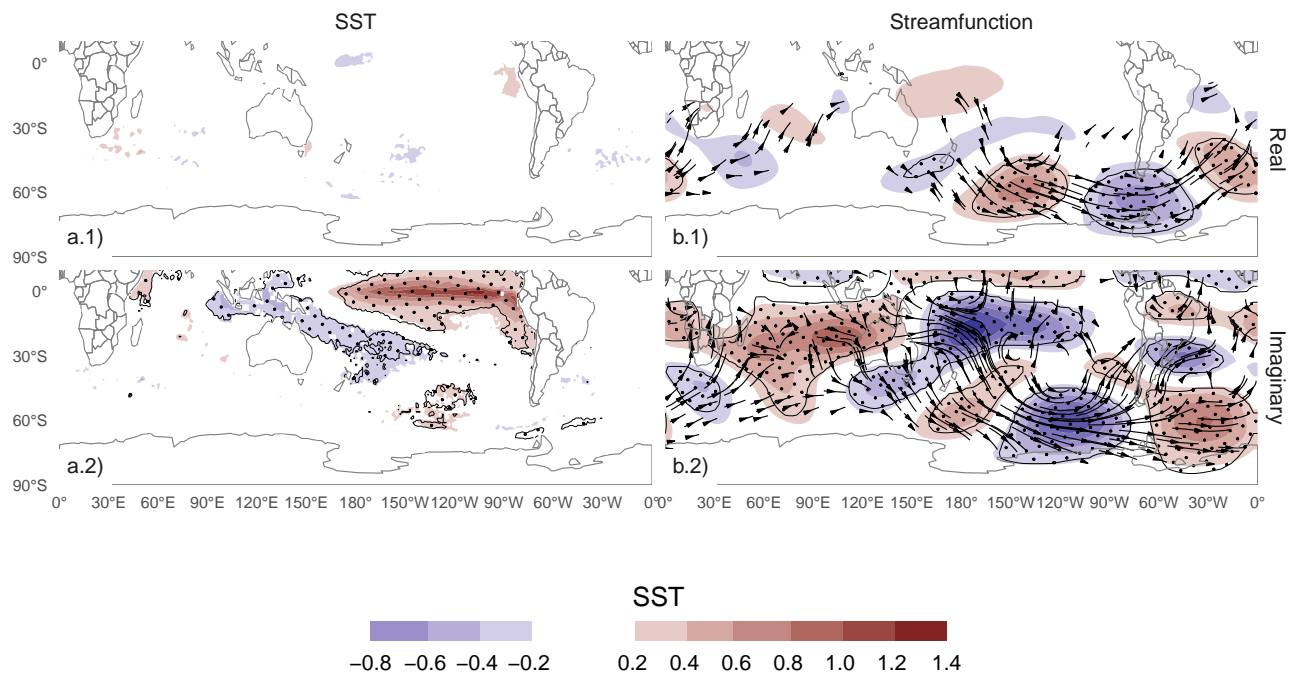


Fig. 15: Regression maps of cEOF2 with SST (column a) and streamfunction zonal anomalies with their corresponding activity wave flux (column b). Areas marked with dots have p-values smaller than 0.05 adjusted for FDR. fig:sst-psi-2

aparece el valor en el PDF). This lack of tropical signal suggests different natures of the different phases of the cEOF2 wave train.

To better explore the relationship between tropical forcing and phase of the cEOF2, Figure 16 plots the ONI index and the phase of the cEOF2 for each SON trimester between 1979 and 2019, highlighting years in which the magnitude of cEOF2 is above the median. In years with positive ONI, the phase of the cEOF2 is mostly around  $+90^\circ$  (corresponding with positive imaginary part) and vice versa. In years with near neutral ENSO, the phase of the cEOF2 is much more variable. The black line in Figure 16 is an harmonical fit of the relationship between ONI and cEOF2 phase. The  $r^2$  corresponding to the fit is 0.56, with p-value  $< 0.001$ , indicating a quasi-sinusoidal relation between these two variables.

Furthermore, Figure 16 suggest that strong cEOF2 years tend to coincide with strong ENSO years. The correlation between the absolute magnitude of the ONI and the magnitude of the cEOF2 is 0.44 (CI: 0.16 – 0.66). This relationship, however, appears to be driven only by the three years with strongest ENSO events in the period (2015, 1997, and 1982) which also coincide with the three years with strongest cEOF2 magnitude. If those years are removed, the correlation becomes non-significant (0.044 (CI: -0.28 – 0.36)). Further-

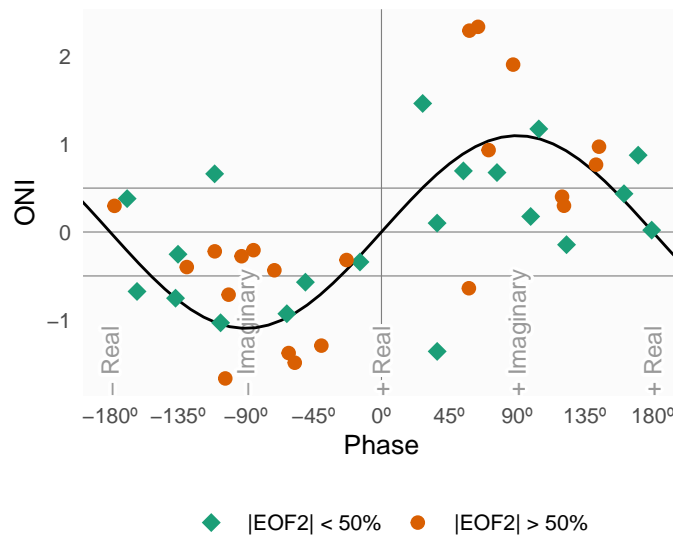


fig:enso-phase  
 Fig. 16: Relationship between ENSO and phase of cEOF2 for the period 1979 – 2019. Colours denote years with magnitude of cEOF2 greater or smaller than the 50th percentile. Black line is the fit  $ONI \sim \cos(\text{phase}) + \sin(\text{phase})$  computed by OLS weighted by the magnitude of cEOF2.

more, even when using all the datapoints, the Spearman correlation –which is robust to outliers– is also non-significant (0.2, p-value= 0.2). Therefore, the relationship between the magnitude of the cEOF2 train wave and ONI remains uncertain.

It could be concluded that the wave train represented by cEOF2 can be both part of the internal variability of the extratropical atmosphere or forced by tropical SSTs. In the former case, the wave train has little phase preference. However, when cEOF2 is excited by tropical SST variability, it tends to remain locked to the imaginary phase. This explains the relative overabundance of years with cEOF2 near positive and negative imaginary phase in Figure 12.

Figure 17 shows SST and streamfunction regression maps for cEOF1. There is no significant pattern of SST anomalies associated with either the Real or Imaginary cEOF1. Consistently, streamfunction anomalies do not show any tropical influence. Instead of that, the real and imaginary CEOF1 are associated with zonally wave activity fluxes in the extra-tropics, except for an equatorward flow from the coast of Antarctica around 150°E in the real component.

### 3.5 Precipitation

#### precipitation

The influence of cEOFs variability in continental rainfall of extra-tropical Southern Hemisphere is also explored. We focus on South America and Ocea-

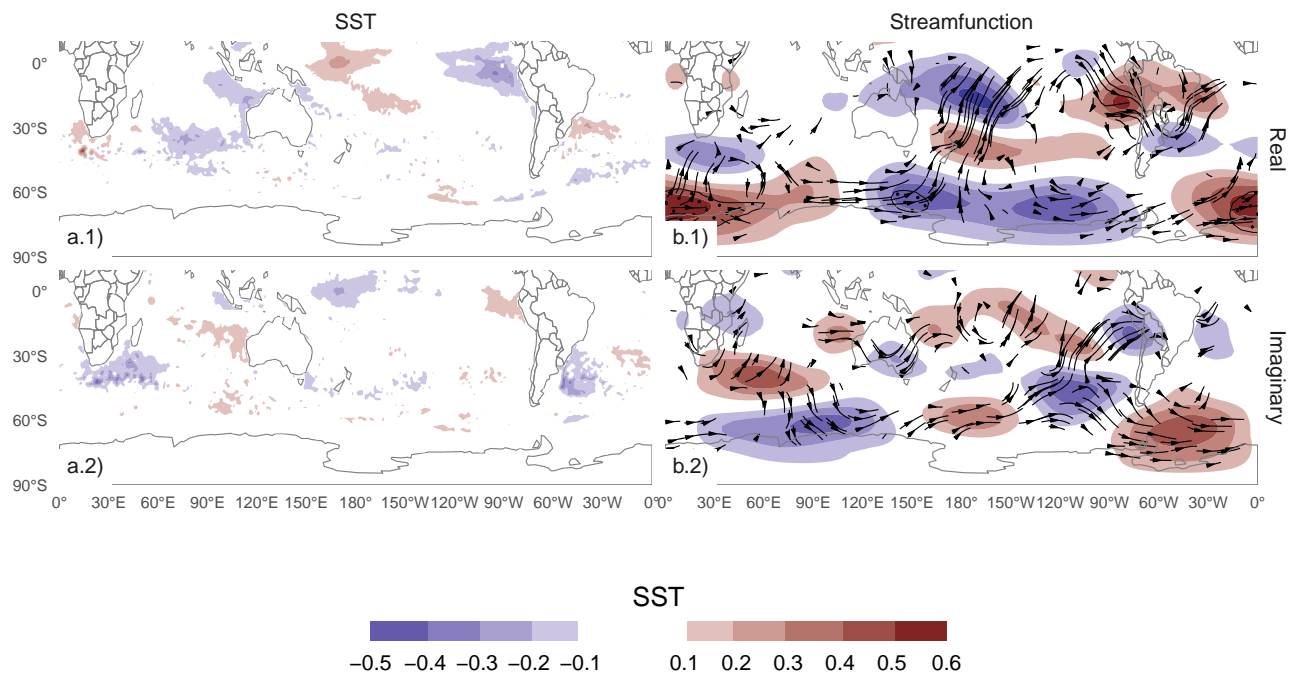


Fig. 17: Same as Figure 15 but for cEOF1.

fig:sst-psi-1

nia sectors, because there are not relevant signals in Southern Africa (not shown).

Figure 18 shows regression maps of seasonal precipitation with each cEOF in South America. cEOF1 is associated with slight decreases in precipitation in Southern Brazil and Paraguay, although these are not statistically significant (Figure 18 column a). The strongest precipitation anomalies are the ones associated with the Imaginary cEOF2. The positive anomalies on Southeastern South America (SESA) and Central Chile, and negative anomalies over South Atlantic Convergence Zone is a well known springtime precipitation signature of ENSO (Cai et al, 2020) and it is also similar to the precipitation anomalies associated with the A-SAM (Campitelli et al, 2021). This is not surprising considering the close relationship between the ONI, the Asymmetric SAM index and the Imaginary cEOF2 shown previously, but further consolidates the identification of this mode with the PSA pattern. The Real cEOF2, on the other hand, is associated with negative precipitation anomalies in a smaller area of SESA. There is a phase dependence of the precipitation anomalies in SESA (not shown), which resembles the relationship between ONI and the phase of cEOF2. (Podrías agregar acá el coeficiente de determinación)

The precipitation anomalies in Oceania associated with cEOF1 are also weak and not statistically significant (Figure 19 column a) whereas EFO2 is associated with large and expansive anomalies. The Real cEOF2 is associated

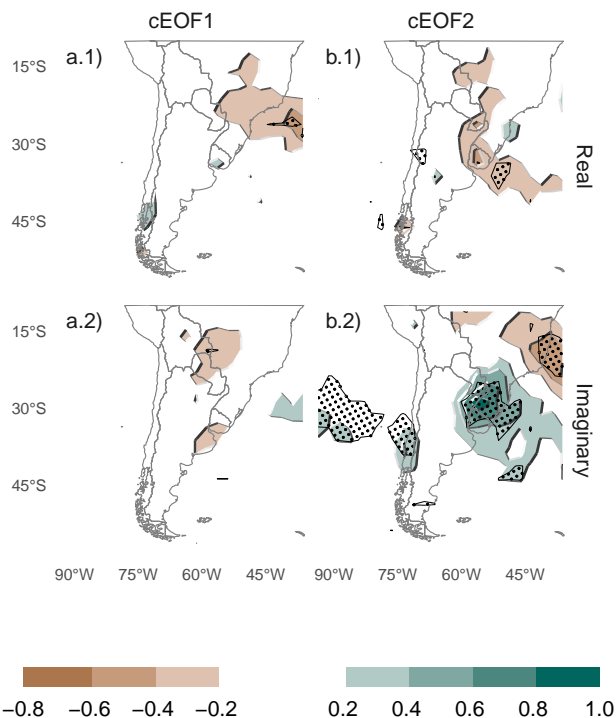


fig:pp-america  
 Fig. 18: Regression of SON mean precipitation anomalies in South America (mm per day, shaded) and (column a) cEOF1 the (row 1) Imaginary and (column 1) Real phase. For the 1979 – 2018 period. Black contours with dots indicate areas with p-value smaller than 0.05 controlling for False Detection Rate.

with positive seasonal anomalies in Northern and Eastern Australia, and negative ones over New Zealand, while precipitation anomalies associated with Imaginary cEOF2 are negative over all Eastern Australia. Again, these negative anomalies are similar to the springtime precipitation anomalies observed in relation with the Asymmetric SAM (Campitelli et al, 2021). (Habría que agregar alguna referencia del ENSO que es más conocido)

To better understand the relationship between precipitation anomalies over Western Australia and the cEOF2, Figure 20 plots precipitation anomalies averaged over Eastern Australia (box in Figure 19) as a function of cEOF2 phase. The black line is the equation  $\text{Precipitation} = -0.26 \sin(\text{phase}) + 0.22 \cos(\text{phase})$ , whose coefficient we fitted by weighted least squares using the magnitude of the EFO2 as weight. (Como te puse antes la ecuación como mucho la pondría sobre la figura) The  $r^2$  corresponding to the fit is 0.37, with p-value < 0.001. There is a strong relationship between the phase of the cEOF2 and precipitation anomalies in Eastern Australia, but Real and Imaginary phases we chose are not exactly aligned with the direction that

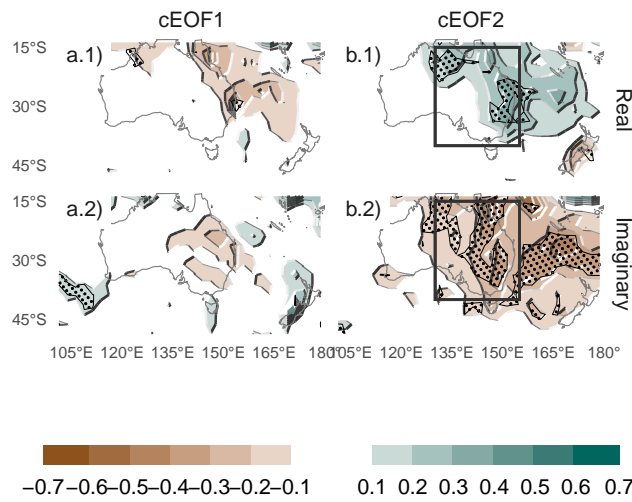


fig:pp-oceania  
Fig. 19: As Figure 18 but for Oceania.

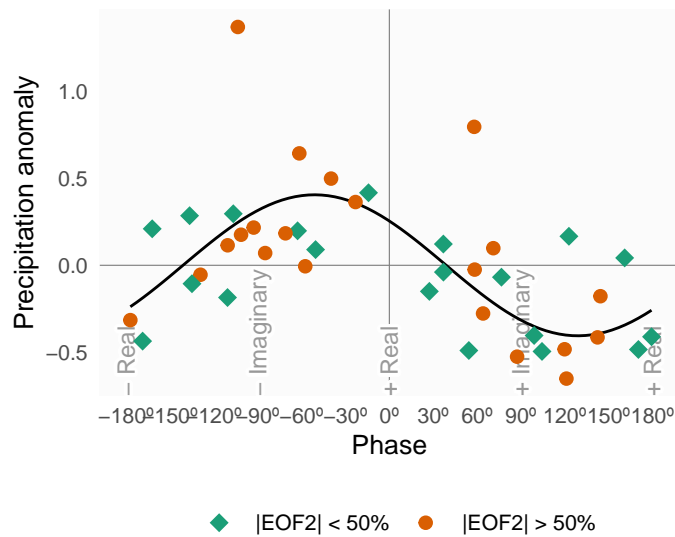


fig:australia-pp-phase  
Fig. 20: Same as Figure 16 but for precipitation anomalies averaged between 130°E and 155°E and 40°S and 15°S (box shown in Figure 19 column b).

maximises this relationship. If one were to perform a more in-depth analysis of the relationship between this pattern and precipitation in this region, one would conceivably align one of the axis (either the Real or the Imaginary) to the one that maximises it.

## 4 Conclusions

conclusions

We studied the two leading modes of co-variability of September-October-November tropospheric and stratospheric zonal anomalies by applying complex Empirical Orthogonal Functions.

The first complex EOF represents the variability of the zonal wave 1 in the stratosphere, but represents a more zonally symmetric monopole in the troposphere. There is a statistically positive trend in the magnitude of this cEOF, which is consistent with previous studies which showed secular changes in springtime wave-1-like patterns (e.g. Raphael, 2003). This mode is closely related to anomalies in Total Ozone Column.

The second complex EOF represents a wave-3 pattern with maximum magnitude in the Pacific sector. Essentially, it is an alternate representation of the PSA1 and PSA2 patterns (Mo and Paegle, 2001). We show that the Imaginary cEOF2 can be identified with the PSA1 and the Real cEOF2 with the PSA2.

We show that this cEOF prefers a phase aligned with the Imaginary cEOF2. This is because when ENSO is active, it forces the cEOF2 to be in the Imaginary phase; positive for El Niño and negative for La Niña. When ENSO is neutral, the cEOF2 is still active, but with no particular phase. This mirrors the results of Cai and Watterson (2002), who showed that the CSIRO Model can create PSA-like variability even in the absence of ENSO forcing (with a climatological run), but the variability of one of the PSA modes was enhanced when adding the ENSO signal.

We further show that the Imaginary cEOF2 is closely related to the Southern Annular Mode in the troposphere. In fact, it is a virtually perfect indicator of the zonally asymmetric portion of the SAM. To the extent that the leading EOF of geopotential height with the cEOF2 removes results in an annular mode that has none of the typical zonal asymmetries of the observed SAM.

This raises the possibility that the asymmetric component of the SAM is actually a statistical contamination of the PSA mode.

Precipitation anomalies in South America associated with the Imaginary cEOF2 show a clear ENSO-like impact, with positive anomalies in South-Eastern South America, negative anomalies in Southern Brazil and positive anomalies in central Chile. Precipitation anomalies associated with the Real cEOF2 are low and not statistically significant, showing that the Imaginary phase is optimally aligned with the direction of maximum precipitation impacts.

On the other hand, over Australia, both the Real and Imaginary phase are associated with significant precipitation anomalies, and we further show that the direction of maximum impact is not aligned with our chosen rotation of cEOF2. However, this underscores the benefit of using complex EOF, since it would be trivial to rotate it.



Code availability

code-availability

A version-controlled repository of the code used to create this analysis, including the code used to download the data can be found at <https://github.com/eliocamp/shceof>.

## 5 References

references

### References

- Adams JC, Swartztrauber PN, Sweet R (1999) FISHPACK, a package of Fortran subprograms for the solution of separable elliptic partial differential equations. <https://www2.cisl.ucar.edu/resources/legacy/fishpack>
- Albers S, Campitelli E (2020) Rsoi: Import Various Northern and Southern Hemisphere Climate Indices
- Allaire J, Horner J, Xie Y, Marti V, Porte N (2019) Markdown: Render Markdown with the c Library 'Sundown'
- Baldwin MP, Thompson DWJ (2009) A critical comparison of stratosphere-troposphere coupling indices. *Quarterly Journal of the Royal Meteorological Society* 135(644):1661–1672, DOI 10.1002/qj.479
- Bamston AG, Chelliah M, Goldenberg SB (1997) Documentation of a highly ENSO-related sst region in the equatorial pacific: Research note. *Atmosphere-Ocean* 35(3):367–383, DOI 10.1080/07055900.1997.9649597
- Bell B, Hersbach H, Berrisford P, Dahlgren P, Horányi A, Muñoz Sabater J, Nicolas J, Radu R, Schepers D, Simmons A, Soci C, Thépaut JN (2020) ERA5 monthly averaged data on pressure levels from 1950 to 1978 (preliminary version). Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (Accessed on <19-02-2021>), <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels-monthly-means-preliminary-back-extension?tab=overview>
- Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57(1):289–300, DOI 10.1111/j.2517-6161.1995.tb02031.x
- Cai W, Watterson IG (2002) Modes of Interannual Variability of the Southern Hemisphere Circulation Simulated by the CSIRO Climate Model. *Journal of Climate* 15(10):1159–1174, DOI 10.1175/1520-0442(2002)015<1159:MOIVOT>2.0.CO;2
- Cai W, McPhaden MJ, Grimm AM, Rodrigues RR, Taschetto AS, Garreaud RD, Dewitte B, Poveda G, Ham YG, Santoso A, Ng B, Anderson W, Wang G, Geng T, Jo HS, Marengo JA, Alves LM, Osman M, Li S, Wu L, Karamperidou C, Takahashi K, Vera C (2020) Climate impacts of the El Niño–Southern Oscillation on South America. *Nature Reviews Earth & Environment* 1(4):215–231, DOI 10.1038/s43017-020-0040-3
- Campitelli E (2020) metR: Tools for Easier Analysis of Meteorological Fields

- Campitelli E, Diaz L, Vera C (2021) Assessment of zonally symmetric and asymmetric components of the Southern Annular Mode using a novel approach. *Climate Dynamics* (under review)
- Dowle M, Srinivasan A (2020) `Data.table`: Extension of 'data.frame'
- Fogt RL, Marshall GJ (2020) The Southern Annular Mode: Variability, trends, and climate impacts across the Southern Hemisphere. *WIREs Climate Change* 11(4):e652, DOI 10.1002/wcc.652
- Gong D, Wang S (1999) Definition of Antarctic Oscillation index. *Geophysical Research Letters* 26(4):459–462, DOI 10.1029/1999GL900003
- Hartmann DL, Garcia RR (1979) A Mechanistic Model of Ozone Transport by Planetary Waves in the Stratosphere. *Journal of the Atmospheric Sciences* 36(2):350–364, DOI 10.1175/1520-0469(1979)036<0350:AMMOOT>2.0.CO;2
- Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut JN (2019) ERA5 monthly averaged data on pressure levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (Accessed on <08-03-2021>), DOI 10.24381/cds.6860a573
- Hobbs WR, Raphael MN (2010) Characterizing the zonally asymmetric component of the SH circulation. *Clim Dyn* 35(5):859–873, DOI 10.1007/s00382-009-0663-z
- Horel JD (1984) Complex Principal Component Analysis: Theory and Examples. *Journal of Applied Meteorology and Climatology* 23(12):1660–1673, DOI 10.1175/1520-0450(1984)023<1660:CPCATA>2.0.CO;2
- Hufkens K (2020) `Ecmwfr`: Programmatic interface to the two European Centre for Medium-Range Weather Forecasts API services
- Irving D, Simmonds I (2016) A New Method for Identifying the Pacific–South American Pattern and Its Influence on Regional Climate Variability. *Journal of Climate* 29(17):6109–6125, DOI 10.1175/JCLI-D-15-0843.1
- Marshall GJ (2003) Trends in the Southern Annular Mode from Observations and Reanalyses. *J Climate* 16(24):4134–4143, DOI 10.1175/1520-0442(2003)016<4134:TITSAM>2.0.CO;2
- Mo KC, Paegle JN (2001) The Pacific–South American modes and their downstream effects. *International Journal of Climatology* 21(10):1211–1229, DOI 10.1002/joc.685
- North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling Errors in the Estimation of Empirical Orthogonal Functions. *Monthly Weather Review* 110(7):699–706, DOI 10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2
- Plumb RA (1985) On the Three-Dimensional Propagation of Stationary Waves. *J Atmos Sci* 42(3):217–229, DOI 10.1175/1520-0469(1985)042<0217:OTTDPO>2.0.CO;2
- R Core Team (2020) `R`: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
- Raphael M (2003) Recent, Large-Scale Changes in the Extratropical Southern Hemisphere Atmospheric Circulation. *J Climate* 16(17):2915–2924, DOI

- 10.1175/1520-0442(2003)016<2915:RLCITE>2.0.CO;2
- Raphael MN (2004) A zonal wave 3 index for the Southern Hemisphere. *Geophysical Research Letters* 31(23), DOI 10.1029/2004GL020365
- Smith AK (1995) Numerical simulation of global variations of temperature, ozone, and trace species in the stratosphere. *Journal of Geophysical Research: Atmospheres* 100(D1):1253–1269, DOI 10.1029/94JD02395
- Trenberth KE (1980) Planetary Waves at 500 mb in the Southern Hemisphere. *Mon Wea Rev* 108(9):1378–1389, DOI 10.1175/1520-0493(1980)108<1378:PWAMIT>2.0.CO;2
- Trenberth KF, Mo KC (1985) Blocking in the Southern Hemisphere. *Mon Wea Rev* 113(1):3–21, DOI 10.1175/1520-0493(1985)113<0003:BITSH>2.0.CO;2
- van Loon H, Jenne RL (1972) The zonal harmonic standing waves in the southern hemisphere. *Journal of Geophysical Research* 77(6):992–1003, DOI 10.1029/JC077i006p00992
- Wickham H (2009) *Ggplot2: Elegant Graphics for Data Analysis*. Use R!, Springer-Verlag, New York, DOI 10.1007/978-0-387-98141-3
- Wirth V (1993) Quasi-stationary planetary waves in total ozone and their correlation with lower stratospheric temperature. *Journal of Geophysical Research: Atmospheres* 98(D5):8873–8882, DOI 10.1029/92JD02820
- Xie P, Arkin PA (1997) Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. *Bull Amer Meteor Soc* 78(11):2539–2558, DOI 10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2
- Xie Y (2015) *Dynamic Documents with R and Knitr*, 2nd edn. Chapman and Hall/CRC, Boca Raton, Florida