

Patrones de variabilidad espacial y temporal de la temperatura de la superficie del mar en el Pacífico Ecuatorial

 Edward Alburqueque^{*1,2} y Joel Rojas²

¹*Instituto del Mar del Perú, Lima, Perú*

²*Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Físicas, Lima, Perú*

Recibido 13 Feb 2024 – Aceptado 24 Set 2024 – Publicado 01 Dic 2024

Resumen

En el marco de esta investigación, se utiliza el análisis de Funciones Ortogonales Empíricas (FOE) para simplificar un conjunto de datos que contiene información sobre la temperatura de la superficie del mar (TSM). El objetivo es transformar este conjunto original en otro que posea una cantidad significativamente menor de variables, asegurando que estas nuevas variables conserven la mayor cantidad posible de la información del conjunto inicial. Este segundo conjunto se construye después de calcular los autovalores y autovectores de la matriz de covarianza. La meta es identificar los patrones más cruciales en términos espaciales y temporales (componentes principales) de la variabilidad de la TSM en el Océano Pacífico Ecuatorial (Latitud: (30°N - 30°S), Longitud: (140°E-70°O)) y luego asociar estos patrones (o modos) con fenómenos como El Niño y la Oscilación Sur (ENOS). Finalmente, para comprobar la validez de los patrones obtenidos, estos se contrastan con los patrones desarrollados por organizaciones tanto nacionales como internacionales.

Palabras clave: Covarianza, modos, TSM, autovalores, autovectores.

Spatial and temporal variability patterns of sea surface temperature in the Equatorial Pacific

Abstract

In this research, Empirical Orthogonal Function (EOF) analysis is applied to reduce the number of variables in a sea surface temperature (SST) dataset to a second dataset containing a much smaller number of variables. The condition is that these new variables retain the maximum possible fraction of information from the original dataset. This second dataset is derived by finding the eigenvalues and eigenvectors of the covariance matrix. The objective is to identify the most significant spatial and temporal patterns (principal components) of SST variability in the Equatorial Pacific Ocean (Latitude: (30°N - 30°S), Longitude: (140°E-70°O)) and subsequently associate these patterns (or modes) with phenomena such as the El Niño-Southern Oscillation (ENSO). Finally, to validate the estimated patterns, they are compared with those obtained by national and international institutions.

Keywords: Covariance, modes, SST, eigenvalues, eigenvectors.

1. Introducción

Los datos de TSM proporcionados por agencias internacionales como la NASA, NOAA, ESA, INPE, entre otras, tienden a estar distribuidos con mayor densidad espacial que temporal. Debido a esto, la interpretación

de dichos datos exige el uso de técnicas estadísticas capaces de evaluar tanto las variaciones en el tiempo como en el espacio. Entre las metodologías utilizadas para comprimir la variabilidad en los datos, una de las más comunes en el campo de la oceanografía es el análisis de Funciones Ortogonales Empíricas (FOE). Este análisis,

*ealburqueque@imarpe.gob.pe

© Los autores. Este es un artículo de acceso abierto, distribuido bajo los términos de la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea debidamente citada de su fuente original.



también conocido como Análisis de Componentes Principales (ACP), es una técnica de reducción de datos que guarda similitudes con el Análisis de Factores (AF), un método ampliamente utilizado en las ciencias sociales.

1.1. Antecedentes

Como Preisendorfer expone en su obra [1], uno de los elementos fundamentales de este método fue desarrollado en 1873 por el geómetra italiano Beltrami, quien formuló una versión moderna para resolver matrices cuadradas mediante la descomposición en valores singulares (DVS). Independientemente, el algebrista francés Jordan también realizó este descubrimiento en 1874. La primera aparición del análisis FOE en el ámbito occidental parece haber sido en 1889, cuando Sylvester abordó el problema de la descomposición de una matriz cuadrada en sus valores singulares dentro del campo del álgebra abstracta. Aproximadamente una década después, Pearson [2] reformuló el análisis de regresión lineal, introduciendo una representación geométrica clara del método en el espacio euclidiano. La primera utilización del método en meteorología parece haber sido realizada en 1948 por G.P. Wadsworth y sus colaboradores en el Massachusetts Institute of Technology (MIT). Casi simultáneamente, Fu-kuoka [3] aplicó el método de manera totalmente independiente. Con la llegada de la computadora Whirlwind al MIT en la década de 1950, E.N. Lorenz utilizó el trabajo de Wadsworth y sus colegas para llevar a cabo estudios de predicción de anomalías en una cuadrícula de 64 puntos que abarcaba el continente estadounidense, el sur de Canadá y partes de los océanos circundantes. El trabajo de Lorenz [4] se ha convertido en un referente clásico en el campo de la estadística aplicada a la predicción climática. La aplicación de las FOE a conjuntos de datos oceanográficos surgió casi una década después, con el estudio de Trenberth [5], que relacionó las oscilaciones atmosféricas en el hemisferio sur con las observaciones de TSM. Diversos autores [6], [7], [8], [9], [10], [11], [12], entre otros, han llevado a cabo estudios aplicados de las FOE en campos de TSM.

1.2. Relevancia del análisis FOE

La relevancia de este análisis se fundamenta en que los datos geofísicos que presentan una alta correlación pueden ser representados de manera efectiva mediante un número reducido de funciones ortogonales y sus correspondientes coeficientes ortogonales temporales, conocidos como componentes principales. A medida que aumenta la correlación entre el conjunto de datos, disminuye la cantidad de funciones y coeficientes necesarios para describir la varianza [13], [6]. Los autovectores resultantes, que representan patrones espaciales, son herramientas valiosas para la interpretación física de los

procesos complejos presentes en un conjunto de datos, y su contribución a la varianza total se expresa a través de los autovalores asociados [14]. Las funciones ortogonales empíricas se organizan de forma decreciente, siendo la primera aquella que abarca la mayor fracción de la varianza total. Las funciones de orden superior presentan escalas espaciales reducidas, con coeficientes caracterizados por menores amplitudes y frecuencias más elevadas, las cuales a menudo se relacionan con el ruido [15].

1.3. Limitaciones del análisis FOE

Este análisis se emplea exclusivamente para identificar oscilaciones estacionarias. Para examinar la propagación de fenómenos ondulatorios, se recurre a las Funciones Ortogonales Empíricas Extendidas (FOEE), que requieren el uso de la matriz de covarianzas rezagadas mencionada en la referencia 16, o a las Funciones Ortogonales Empíricas Complejas (FOEC) en el ámbito de las frecuencias (revisar [17] y [18]). Aunque el análisis FOE es no paramétrico, lo cual puede considerarse una ventaja en principio, esta característica también representa una de sus debilidades. El análisis FOE se restringe a expresar el conjunto original de datos como una combinación "lineal" de vectores ortonormales. Es precisamente la ortonormalidad de estos vectores lo que impone un límite al análisis, ya que en ciertas distribuciones de conjuntos de datos los resultados obtenidos no son satisfactorios. Shlens [19] presenta dos ejemplos ilustrativos. En el primero, se analiza la posición de una persona en una rueda de la fortuna. Al considerar las posiciones, se observa que la distribución de los datos es circular, y toda la dinámica puede ser representada por la variable θ (el ángulo de precesión de la rueda). Si el objetivo es identificar las direcciones principales de variabilidad, resulta inviable el uso de las componentes principales para representar la varianza de los datos (Figura 1(A)).

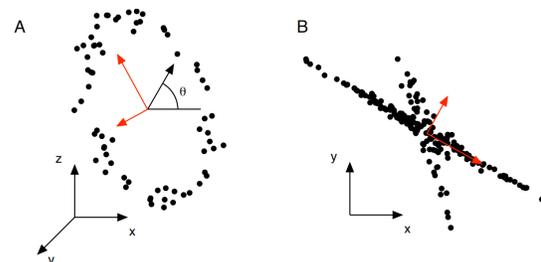


Figura 1: Algunos casos en los que el análisis FOE presenta fallos (ver las flechas rojas, [19])

El segundo ejemplo implica la travesía en automóvil por una ciudad: "Conduciendo a través de caminos no ortogonales". En esta analogía, el análisis FOE exige que cada nuevo camino explorado sea ortogonal al anterior; sin embargo, este requerimiento es demasiado restrictivo,

ya que, en general, los caminos urbanos están organizados a lo largo de ejes no ortogonales (ver Figura 1(B)). En este caso, un conjunto de datos que presenta una distribución no gaussiana y ejes no ortogonales provoca que los ejes con la mayor varianza no correspondan a una respuesta adecuada. Al aplicar el análisis FOE a datos de TSM, uno de los desafíos principales es la interpretación física de los patrones espaciales y temporales que se obtienen. Es importante recordar que los patrones son construcciones matemáticas seleccionadas para representar la varianza en el dominio de interés de la manera más eficiente posible [20].

1.4. Área de Estudio

En la región del Océano Pacífico Ecuatorial, la TSM es de suma importancia, ya que actúa como un indicador de la aparición del fenómeno El Niño, que se presenta de manera natural en esta área y tiene efectos globales significativos para la humanidad. Además, muchos investigadores consideran esta zona para calcular diversos índices climáticos que faciliten la comprensión de las variaciones entre los eventos, tanto en sus patrones espaciales como en sus desarrollos a lo largo del tiempo. El área elegida para llevar a cabo el estudio FOE es el Océano Pacífico Ecuatorial (30°N - 30°S, 140°E - 70°W).

1.5. Objetivos

Objetivo General:

- Reducir la dimensionalidad de los productos de Temperatura de la Superficie del Mar (ERSST, HadISST y OISST), asegurando que se conserve la mayor cantidad de información posible, e identificar los patrones principales de variabilidad en el Pacífico Ecuatorial.

Objetivos Específicos:

- Demostrar que el análisis FOE permite asociar los patrones temporales y espaciales de la TSM con fenómenos bien documentados como El Niño y El Niño Modoki.
- Estimar dos nuevos índices (a través de la rotación de los patrones iniciales ENOS y MODOKI), denominados índice E y C, que describen de manera óptima la estructura espacial no lineal de los eventos ENOS.
- Combinar el análisis FOE y la regresión lineal para estimar índices que representen los dos tipos de ENOS: uno correspondiente al Océano Pacífico Central (ENOS-CP) y otro al Pacífico Este (ENOS-EP).

2. Funciones Ortogonales Empíricas

2.1. Planteamiento del Problema

En la vida cotidiana, a menudo nos encontramos con problemas en los que las mediciones presentan un alto nivel de ruido o son redundantes. La razón para aplicar el análisis FOE se ilustra a través del siguiente ejemplo extraído del artículo de Shlens [19]. Imaginemos el movimiento de un resorte ideal. Este sistema está compuesto por una esfera de metal con una masa "m", que está conectada a un resorte que carece de masa y fricción (ver Figura 2).

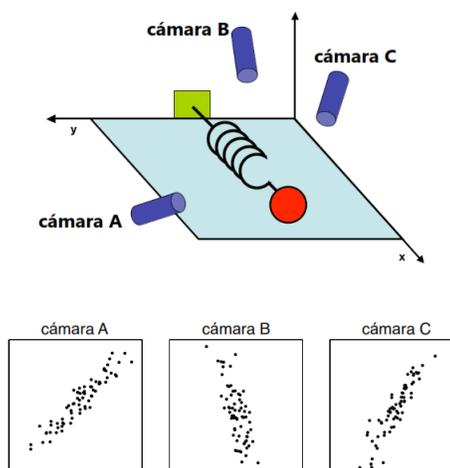


Figura 2: Modelo propuesto por Shlens [19] para explicar el funcionamiento del Análisis FOE

Una esfera se desplaza a cierta distancia del punto de equilibrio, lo que provoca que el resorte se estire. Este resorte ideal oscila indefinidamente a lo largo del eje x con una frecuencia fija. Este problema puede resolverse mediante una función que varía en el tiempo, expresada como $f(t) = A \cos(\omega t + \omega_0)$. Si nos colocamos en la posición de un investigador principiante, optaríamos por medir la ubicación de la esfera en un espacio tridimensional. Para ello, situaríamos tres cámaras de cine alrededor de nuestro sistema de interés. A una frecuencia específica, cada cámara capturaría una imagen bidimensional que mostraría la posición de la esfera (una proyección). Sin embargo, debido a nuestra falta de conocimiento, no tenemos claridad sobre cuáles son los ejes 'x', 'y' y 'z', por lo que elegimos tres posiciones para las cámaras con ángulos arbitrarios en relación con el sistema A, B, C. Los ángulos entre nuestras mediciones ni siquiera forman ángulos rectos. A continuación, se procederá a registrar el movimiento de la esfera durante varios minutos.

La pregunta crucial es: **¿Cómo podemos, a partir de estos datos, derivar una ecuación sencilla de x en función del tiempo?**

Sabemos de antemano que, si fuéramos experimentadores astutos, habríamos medido la posición a lo largo del eje 'x' utilizando solo una cámara. Sin embargo, esto no es lo que ocurre en la práctica. No estamos seguros de cuáles mediciones representan de manera más precisa la dinámica de nuestro sistema; además, registramos más dimensiones de las que realmente son necesarias y enfrentamos el inconveniente del ruido. En este contexto, el ruido se refiere a factores como la fricción del aire, las imperfecciones de la cámara o incluso la fricción en un resorte que no es ideal. Este ruido que afecta las mediciones solo sirve para complicar aún más la comprensión de la dinámica.

Este ejemplo representa el reto que los científicos abordan diariamente. Por esta razón, el análisis FOE se convierte en la herramienta adecuada para extraer los patrones más relevantes, tanto espaciales como temporales, que están relacionados con la dinámica de un fenómeno específico.

2.2. Nomenclatura

El tema de las Funciones Ortogonales Empíricas (FOE) se considera complicado y enrevesado, siendo la principal causa de esta confusión la abundancia de términos relacionados, especialmente en publicaciones como libros o artículos (Wilks [21]). La bibliografía resulta bastante ambigua en cuanto a las distinciones entre el enfoque FOE y el método ACP. Aunque algunos autores diferencian claramente ambos métodos, hay otros que emplean FOE y ACP para referirse al mismo procedimiento. Por ejemplo, el método FOE expuesto en el anexo del libro de Peixoto [22] es exactamente igual al método ACP que describe Preisendorfer [1]. Dado que la literatura es en gran medida confusa respecto a las diferencias entre estos enfoques, se ha decidido utilizar las siglas FOE y ACP de manera intercambiable. La mayoría de los autores nombran los patrones espaciales como FOE, aunque algunos los denominan "componentes principales de los patrones" o simplemente "componentes principales". Las series temporales son referidas como "coeficientes de expansión de las series temporales", "coeficientes de expansión", "componentes principales de las series temporales" o simplemente "componentes principales". En este estudio, los patrones espaciales se designarán como FOE, mientras que los patrones temporales se llamarán componentes principales.

2.3. Perspectiva Geométrica

Desde una perspectiva geométrica, el propósito del análisis FOE es representar un conjunto de puntos en un espacio de dimensiones reducidas de manera "óptima", lo que implica "minimizar la distorsión de las distancias entre los individuos". La Figura 3 ilustra la vis-

ta frontal de lo que podría ser un camello o un dromedario. La elección del ángulo para esta figura complica un poco su identificación. ¿Qué aspectos diferencian las perspectivas de la parte izquierda y la derecha? En la parte derecha, las distancias están menos alteradas y la representación abarca más espacio en la imagen. La silueta presentada es una proyección de un objeto tridimensional en un plano bidimensional.

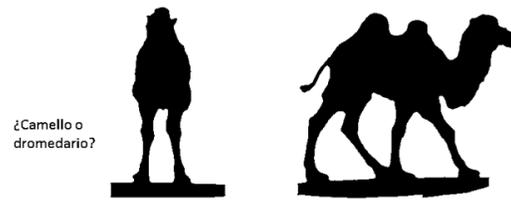


Figura 3: Modelo presentado por [23] para ilustrar cómo funciona el análisis FOE

Para que una representación sea efectiva, es necesario elegir un ángulo de visión adecuado. En términos generales, las FOE implican encontrar el espacio de representación óptimo (de dimensiones reducidas), lo que facilita la visualización más clara de la estructura de un conjunto de datos.

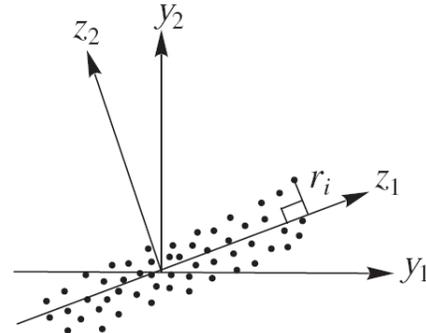


Figura 4: La formulación del problema del Análisis de Componentes Principales se plantea como una minimización de la suma de r_i^2 , donde r_i representa la distancia más corta desde el i -ésimo punto del conjunto de datos hasta el eje Z_1 [24]

2.4. Relaciones de Ortogonalidad

A través del análisis de componentes principales, se obtienen los autovalores y autovectores de la matriz de covarianza \mathbf{C} . Los autovectores ortonormales representan una base, lo que significa que la matriz $\mathbf{y} = [y_1, \dots, y_m]^T$ puede expresarse en función de estos autovectores. $\mathbf{e} = [e_1, \dots, e_j, \dots, e_m]$

$$\mathbf{y} - \bar{\mathbf{y}} = \sum_{j=1}^m a_j(t) \mathbf{e}_j$$

donde $a_j(t)$ representan las componentes principales. Para obtener $a_j(t)$ a la ecuación anterior se le multiplica

\mathbf{e}_i^T por la izquierda, y se usa la relación ortonormal de los autovectores.

$$\mathbf{e}_i^T \mathbf{e}_j = \delta_{ij}$$

(siendo δ_{ij} la función delta de Kronecker, que es 1 si $i=j$ y 0 si $i \neq j$). Realizado el cálculo se obtiene:

$$a_j(t) = \mathbf{e}_j^T (\mathbf{y} - \bar{\mathbf{y}})$$

Esto significa que $a_j(t)$ se obtiene proyectando el vector $\mathbf{y} - \bar{\mathbf{y}}$ sobre el autovector “e” (producto escalar entre ambos vectores). La terminología en la literatura varía considerablemente: las a_j son denominadas componentes principales (CP), coeficientes temporales y amplitudes; en cambio, los \mathbf{e}_j también se conocen como patrones espaciales, vectores principales o FOE, los \mathbf{j} son referidos como modos. Es importante tener en cuenta que para las series de tiempo, a_j representa una función del tiempo, mientras que \mathbf{e}_j representa una función del espacio. Por consiguiente, los términos coeficientes temporales y espaciales son correctamente utilizados ([25] y [26]).

2.5. Descomposición en Valores Singulares (DVS)

En lugar de abordar el problema relacionado con los autovalores de la matriz de covarianza \mathbf{C} , una opción más práctica desde la perspectiva computacional para llevar a cabo el Análisis de Componentes Principales es implementar la DVS sobre \mathbf{Y} , que tiene un tamaño de $m \times n$ (Kelly [27]). Considerando que $m \geq n$, podemos representar la DVS de la siguiente manera ([28], [29]):

$$\mathbf{Y} = \mathbf{E}\mathbf{S}\mathbf{F}^T$$

| | | |
|-------------------------------|-------------------------------|---------------------|
| \mathbf{E} | \mathbf{S} | \mathbf{F}^T |
| \mathbf{E}' $m \times n$ | \mathbf{S}' $n \times n$ | 0 $n \times n$ |
| 0 $m \times m$ | 0 $m \times n$ | |

$$\mathbf{Y}_{m \times n} = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{m1} & \cdots & y_{mn} \end{bmatrix} =$$

$$\begin{bmatrix} e_{11} & \cdots & e_{m1} \\ e_{12} & \ddots & e_{m2} \\ \vdots & \ddots & \vdots \\ e_{1m} & \cdots & e_{mm} \end{bmatrix} \begin{bmatrix} s_{11} & \cdots & 0 \\ 0 & & 0 \\ \vdots & s_{22} & \vdots \\ 0 & \cdots & s_{nn} \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} f_{11} & \cdots & f_{n1} \\ \vdots & \ddots & \vdots \\ f_{1n} & \cdots & f_{nn} \end{bmatrix}$$

La matriz \mathbf{E} posee dimensiones $m \times m$ y contiene una submatriz \mathbf{E}' de dimensiones $m \times n$, bajo la condición de que $m > n$ algunas columnas son nulas. En cuanto a \mathbf{S} , esta es una matriz de dimensiones $m \times n$ que incluye una submatriz diagonal \mathbf{S}' de dimensiones $n \times n$, y con algunas filas de vectores nulas. Asimismo, \mathbf{F}^T se refiere a otra matriz con dimensiones $n \times n$. \mathbf{E} y \mathbf{F} son matrices ortonormales, lo que implica que satisfacen:

$$\mathbf{E}^T \mathbf{E} = \mathbf{I}, \quad \mathbf{F}^T \mathbf{F} = \mathbf{I}$$

En este contexto, \mathbf{I} representa la matriz identidad. Las n columnas que se encuentran en la parte izquierda de \mathbf{E} albergan los n vectores singulares izquierdos, mientras que las columnas de \mathbf{F} contienen los n vectores singulares derechos. Por su parte, los elementos en la diagonal de \mathbf{S}' corresponden a los valores singulares.

2.6 FOE's significantes (reglas de selección)

En la práctica, se descartan aquellos modos que presentan una varianza asociada más baja, ya que tienden a contener ruido. ¿Cómo determinamos cuántos modos debemos conservar? Existen ciertos criterios de selección". Un enfoque sencillo, aunque bastante subjetivo, consiste en representar gráficamente los autovalores λ_j versus el número de modos j . Lo que se busca es identificar una transición brusca entre los autovalores altos y bajos en torno al modo k . Si se observa esto, se pueden conservar los k primeros modos. Luego, exploraremos algunos métodos más sofisticados para distinguir la señal del ruido.

2.6.1 Regla de Preisendorfer (regla N)

Asumimos que para cada l , el conjunto de autovalores es ordenado en orden descendiente

Esta regla, propuesta por [30], se fundamenta en el método de Monte Carlo, que consiste en generar matrices de datos aleatorios $\mathbf{R}_l(l = 1, \dots, L)$ con la misma dimensión que la matriz \mathbf{Y} . Los elementos aleatorios deben seguir una distribución normal y poseer la misma varianza que los datos reales. Se aplica el análisis FOE a cada una de las matrices aleatorias, obteniendo así los autovalores $\lambda_j^{(l)}$. Suponemos que para cada l , los autovalores se ordenan de manera descendente.

$$\lambda_j^{(l)} = [\lambda_1^{(l)}, \lambda_2^{(l)}, \lambda_3^{(l)} \dots \lambda_p^{(l)}].$$

A continuación, calculamos el % de varianza explicada por los autovalores mediante la siguiente fórmula:

$$U_j(l) = \lambda_j^{(l)} \left[\sum_{j=1}^p \lambda_j^{(l)} \right]^{-1} \times 100$$

Para $l = 1, \dots, L$ y $j = 1, \dots, p$.

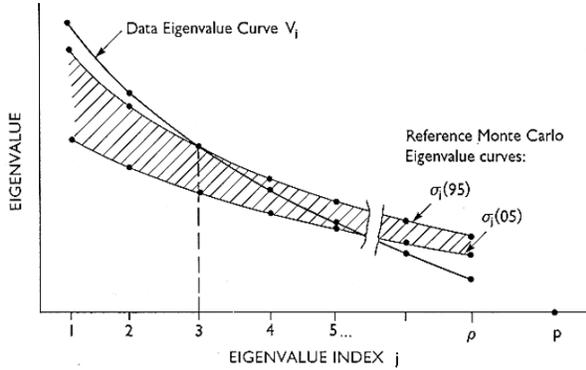


Figura 5: Geometría de la regla de Preisendorfer. Los modos relevantes son aquellos que se sitúan por encima de la curva $\sigma_j(95)$. La ilustración representa el escenario en el que el total de modos significativos es 2.

2.6.2 Regla de North

De acuerdo a North [26], el error al estimar el autovalor λ_α es definida por:

$$\delta\lambda_\alpha \approx \sqrt{\frac{2}{N}} \lambda_\alpha$$

donde la cantidad de realizaciones es N . Si otro autovalor λ_β está muy próximo a λ_α tal que:

$$\Delta_{\alpha\beta} = |\lambda_\alpha - \lambda_\beta| \leq \delta\lambda_\alpha$$

En otras palabras, cuando el intervalo entre los autovalores es similar al error en la estimación, se presenta una “degeneración efectiva”. Esta degeneración de un par de autovalores genera una ambigüedad inherente en la caracterización de los autovectores y sus componentes asociados, ya que cualquier combinación lineal de los autovectores también puede considerarse un autovector. El total de modos significativos será equivalente a los primeros autovalores que no satisfacen la ecuación mencionada anteriormente.

3. Datos y Metodología

Los productos TSM utilizados en este trabajo son los siguientes:

- **ERSST:** Los datos ERSST (The extended reconstructed sea surface temperature) son datos mensuales de TSM (desde 1854) y son derivados del conjunto de datos de ICOADS (International Comprehensive Ocean-Atmosphere Data Set). Los datos tienen una resolución espacial de 2° latitud x 2° longitud con un tamaño global de 89×180 píxeles (88° N - 88° S, 0° E - 358° E). Estos datos están disponibles en: <https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.ersst.v5.html>

- **HadISST:** El Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST) es una combinación de campos mensuales completos de TSM y concentración de hielo marino a nivel mundial (desde 1871). HadISST utiliza la interpolación óptima de espacio reducido aplicada a las TSM del Marine Data Bank (principalmente rastreos de barcos) e ICOADS hasta 1981 y una combinación de TSM derivadas de datos satélites e in situ y ajustadas para 1982 en adelante. Estos datos en formato netcdf están disponibles en: <https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html>

- **OISST:** Optimum Interpolation (OI) Sea Surface Temperature (SST). Estos datos de 1° de resolución espacial proporcionan campos globales que se basan en una combinación de observaciones de temperatura del océano desde plataformas satelitales e in situ. Este dato se puede descargar del siguiente link: <https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html>

La técnica empleada para implementar el análisis FOE en los datos previamente citados es la siguiente:

- 1) Analicemos un grupo de N mapas a lo largo de los períodos $t = 1 \dots N$, en los cuales cada mapa incluye datos sobre el campo ψ en las ubicaciones $p = 1 \dots P$. De este modo, disponemos de P series temporales $\psi_p(t)$, cada una con una longitud de N , cumpliendo que $P > N$.
- 2) Se eliminan las fluctuaciones estacionales al sustraer del campo analizado el promedio climatológico anual μ_p (lo que genera anomalías de TSM o ATSM). Es posible trabajar tanto con las anomalías como con las anomalías estandarizadas. La estandarización resulta crucial cuando se examinan dos o más campos en conjunto, garantizando así que ningún campo predomine sobre los demás.

$$X_p(t) = \frac{\psi_p(t) - \mu_p}{\sigma_p} \quad ; \quad \mu_p = \frac{1}{N} \sum_{t=1}^N \psi_p(t)$$

- 3) Se elabora la matriz $X_{N \times P}$

$$\mathbf{X} = \begin{bmatrix} y_1(1) & \cdots & y_P(1) \\ \vdots & \ddots & \vdots \\ y_1(N) & \cdots & y_P(N) \end{bmatrix}$$

- 4) La matriz C tiene una forma simétrica y cuadrada, incluso si X no presenta esta misma estructura. Si las series de datos contenidas en X han sido normalizadas, entonces C representará la matriz de correlación en lugar de la matriz de covarianza.

- 5) Una vez obtenida la matriz C se da solución al problema de autovalores-autovectores:

$$CE = E\Lambda$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}$$

Por lo general, los autovalores de la matriz Λ se disponen en orden descendente. Dado que la matriz de datos X es de tipo real, la matriz de covarianza C se considera definida positiva, lo que implica que todos sus autovalores son mayores o iguales a

cero. Aunque Λ tiene una dimensión de $N \times N$, normalmente solo los primeros K autovalores son diferentes de cero, en donde $K \leq \min(N, M)$. Así, la dimensión de Λ es en realidad $K \times K$, lo que significa que solo se pueden identificar K modos. Cada autovalor que es diferente de cero λ_k , se relaciona con un autovector columna E^k de la matriz E .

- 6) La evolución en el tiempo del k -ésimo FOE se logra al proyectar sobre los autovectores los datos originales, obteniéndose las Componentes Principales:

$$PC^k(t) = \sum_{m=1}^M E_m^k X_m(t)$$

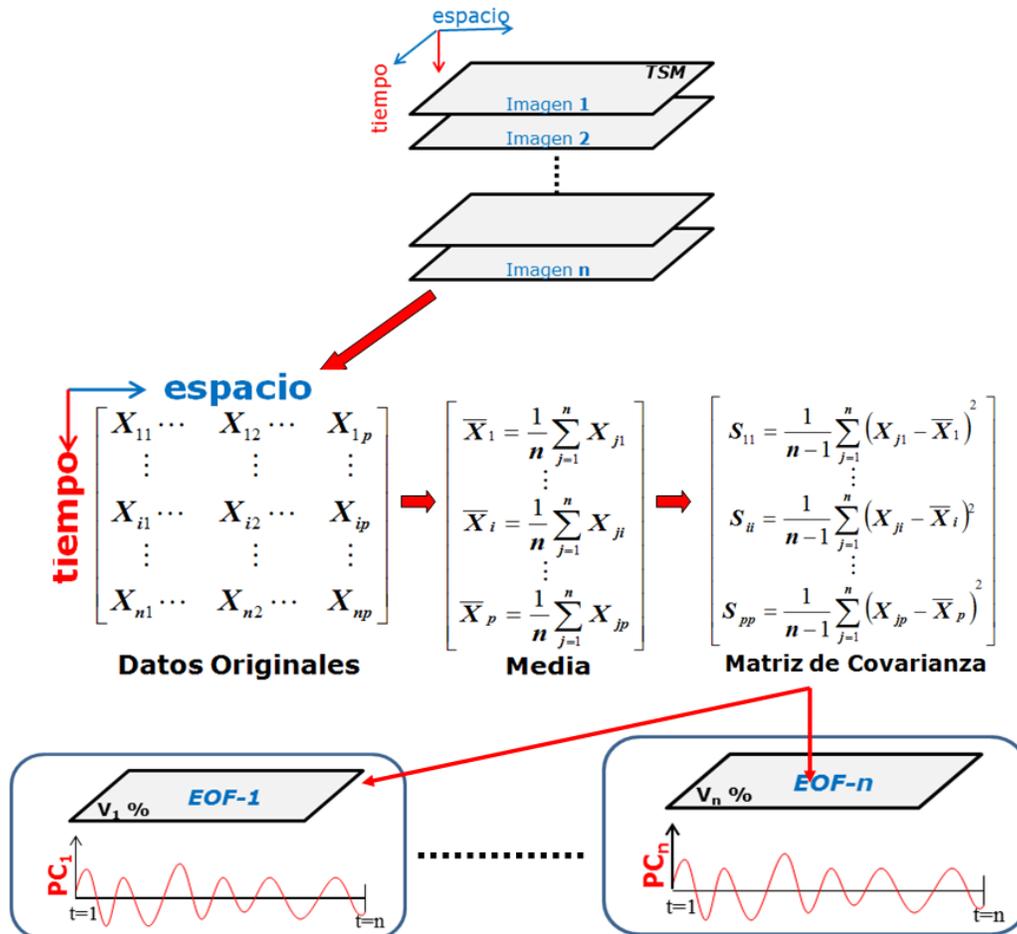


Figura 6: Metodología empleada para la obtención de las FOE

- 7) Cada autovalor λ_k está relacionado con la proporción de la varianza explicada del campo X que se representa a través del modo k .

$$\% \text{ Varianza} = \frac{\lambda_k}{\sum_{i=1}^k \lambda_i} * 100$$

- 8) El campo X original se puede reconstruir completamente al multiplicar cada autovector por su respectiva PC y luego sumar todos los productos resultantes.

$$X_m(t) = \sum_{k=1}^K E_m^k PC^k(t)$$

- 9) No obstante, la meta de la descomposición FOE es lograr una reconstrucción del campo X original que sea aproximada, menos ruidosa y compacta, utilizando únicamente los primeros H modos, donde $H < K$. El valor de H se selecciona de acuerdo con los criterios establecidos en la sección 2.6.

$$\hat{X}_m(t) = \sum_{k=1}^H E_m^k PC^k(t)$$

- 10) Los patrones o modos del análisis FOE constan de tres elementos: los autovectores (que representan los patrones espaciales), los autovalores y las PC 's (que se refieren a los patrones temporales).

4. Resultados y Discusión

4.1 Variabilidad Térmica en el Océano Pacífico Tropical

Esta región es conocida por “*El Niño*”. [31]. Se produce un calentamiento súbito de las aguas costeras frente a Perú (cada 2-10 años). Durante este fenómeno, el afloramiento cesa abruptamente, lo que provoca una crisis en la productividad pesquera. Un evento de El Niño fuerte puede resultar en un aumento de temperatura de 5°C o más en las aguas superficiales de Perú. En ocasiones, se presenta un fenómeno opuesto al de El Niño, caracterizado por la aparición anómala de aguas frías en el Pacífico ecuatorial. Este fenómeno se conoce como La Niña.

Este fenómeno ejerce una influencia global y genera significativas repercusiones socioeconómicas. Comprender las variaciones en los patrones de El Niño es fundamental para la interpretación y previsión de este fenómeno y sus efectos. Los eventos canónicos de El Niño se vinculan a una anomalía de temperatura cálida que alcanza su máximo en la parte oriental del Océano Pacífico ecuatorial, conocido como El Niño del Pacífico Este (El Niño-PE). En las últimas décadas, se ha observado la

aparición frecuente de un nuevo tipo de El Niño, caracterizado por una anomalía máxima en la parte central del Pacífico ecuatorial. Este fenómeno es conocido por diversos nombres tales como El Niño-PC, El Niño del Pacífico central ([32], [33]), El Niño de la piscina cálida ([34]) o El Niño Modoki (Ashok [10]). Ambos tipos de El Niño presentan distintas teleconexiones y efectos climáticos ([35], [36], [37]). Un estudio de [38] identificó los años en los que se manifestaron tanto El Niño-PE como El Niño-PC durante la temporada de verano. (ver Tabla 1).

| | El Niño-PE | El Niño-PC |
|-------|------------------|------------------|
| 1870s | 1876, 1877 | |
| 1880s | 1888 | |
| 1890s | 1896, 1899 | |
| 1990s | 1902, 1904, 1905 | |
| 1910s | 1911, 1913, 1914 | |
| 1920s | 1925 | |
| 1930s | 1930, 1939 | |
| 1940s | 1940, 1941 | |
| 1950s | 1951, 1957 | |
| 1960s | 1963, 1965, 1969 | 1968 |
| 1970s | 1972, 1976, 1979 | 1977 |
| 1980s | 1982, 1986, 1987 | |
| 1990s | 1991, 1997 | 1990, 1992, 1994 |
| 2000s | 2002, 2003, 2006 | 2001, 2004, 2009 |

Tabla 1: Eventos de El Niño [38]

A partir de la información presentada en la tabla 1, se elaboró un compuesto de los años en que se registraron los eventos EP y PC (Figuras 7 y 8). Se observó que los eventos El Niño-PE se caracterizan por máximas anomalías de TSM en el Pacífico Ecuatorial Este. En contraste, el punto de máximas anomalías para los eventos El Niño-PC se localiza en los 180°W. Para examinar estas posiciones con mayor precisión, se generó una gráfica de la estructura zonal en ambos compuestos, utilizando un promedio entre 2°N y 2°S (Figura 9). Esta figura muestra de manera clara que el centro de las anomalías máximas para El Niño-PC se encuentra desplazado hacia el oeste en comparación con El Niño-PE.

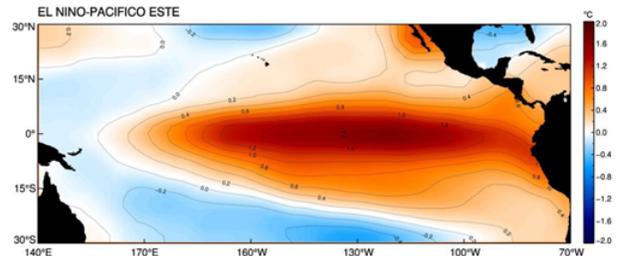


Figura 7: Promedio de ATSM para los eventos El Niño-PE. Datos ERSST

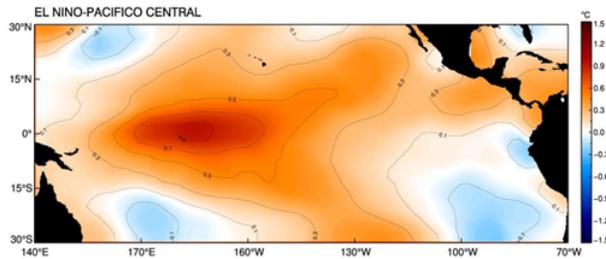


Figura 8: Promedio de ATSM para los eventos El Niño-PC. Datos ERSST

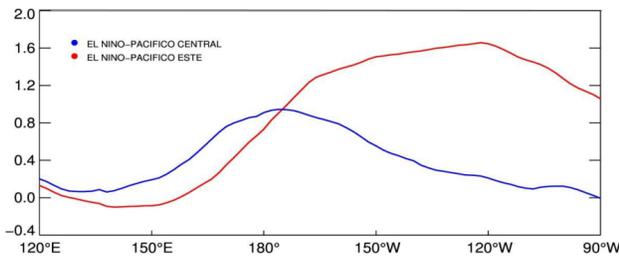


Figura 9: Promedio de ATSM (estructura zonal) para los eventos El Niño-PE y el Niño-PCM entre los 2°N y 2°S. Datos ERSST.

Aunque investigaciones previas han señalado un aumento (en las últimas décadas) en la frecuencia de eventos de El Niño, no se han registrado cambios en la magnitud de dichos fenómenos. El trabajo de [39] indica que los eventos del ENOS no se limitan exclusivamente al Pacífico Central o al Pacífico Este, sino que representan combinaciones lineales diferentes de estos dos tipos. En el caso de los eventos El Niño-PE, las anomalías suelen manifestarse en el extremo oriental del Pacífico en primavera y se desplazan hacia el oeste durante el verano y el otoño. Por otro lado, en los eventos El Niño-PC, las anomalías se extienden desde los subtropicos orientales hacia el Pacífico Central en primavera y verano. Ambos fenómenos alcanzan su máxima intensidad en el verano austral.

4.2 Índices de los eventos El Niño

Para profundizar en la diversidad del ENOS, se han creado varios índices destinados a distinguir distintos tipos de eventos, centrándose especialmente en la fase cálida del ENOS (El Niño). Estos índices se han desarrollado a partir de la temperatura superficial del mar. Las definiciones de los distintos tipos de El Niño suelen diferir según el enfoque empleado (ver [39]).

- a) **Índice Niño 3-4:** Un evento se clasifica como "Piscina caliente" si sus anomalías promedio en la región Niño4 (5°S - 5°N, 160°E - 150°W) superan una desviación estándar y también sobrepasan la anomalía promedio en la región Niño3 (5°S - 5°N, 90°

- 150°W). Por otro lado, los eventos denominados "Lengua Fría" se caracterizan porque las anomalías en la región Niño-3 superan una desviación estándar y son mayores que las anomalías en Niño-4. Revisar [34] y [38].

- b) **Índice Niño Modoki - EMI:** Este índice se define como la diferencia entre la ATSM promedio en el Pacífico Central (10°S - 10°N, 165°E - 140°W) y la ATSM promedio en el Pacífico Oeste (10°S - 20°N, 160°E - 150°W) y en el Pacífico Este (15°S - 5°N, 110°E - 70°W). Su propósito es resaltar la relación desfasada entre las anomalías del Pacífico Central y las del Pacífico Oeste y Este. Revisar [10].

- c) **Índice CP - EP:** Se emplea la regresión de la ATSM con el índice Niño 1+2 (en la región de 0° a 10°S, 90° a 80°W) con el objetivo de eliminar la componente de la ATSM vinculada al calentamiento del Pacífico Oriental o Este. Posteriormente, se realiza un análisis de componentes principales para identificar el modo espacial y el índice temporal relacionado con los eventos CP. De manera análoga, la regresión de la ATSM con el índice Niño-4 se utiliza para excluir la componente de ATSM asociada al calentamiento en el Pacífico Central. Luego, se aplica un análisis PCA para determinar el modo espacial y el índice temporal de los eventos "EP". Revisar [32] y [33].

- d) **Índice C y E :** Este enfoque utiliza dos ejes perpendiculares que se giran 45° en relación con las PCA's de la ATSM en la región del Pacífico Tropical. Las proyecciones de la temperatura superficial del mar en estos ejes rotados generan el índice E, el cual representa El Niño en el Pacífico Oriental, y el índice C, que refleja tanto El Niño como La Niña en el Pacífico Central. Revisar [11].

Se aplicará el análisis FOE para calcular los índices C y E, siguiendo los pasos empleados por [11], y se obtendrán los índices CP y EP mediante el método sugerido por [32]. Las PC's estandarizadas se presentan en grados centígrados. Asimismo, los patrones o modos de ATSM también se expresan en grados centígrados y fueron obtenidos a través de la regresión de la ATSM sobre las series temporales normalizadas (PCs). Además, para caracterizar de forma más precisa los modos de variabilidad, se han correlacionado los patrones temporales con la ATSM de las regiones Niño 1+2, Niño 3, Niño 4 y Niño 3+4 (Figura 10), las cuales se pueden encontrar en el siguiente enlace: <http://www.cpc.ncep.noaa.gov/data/indices/>

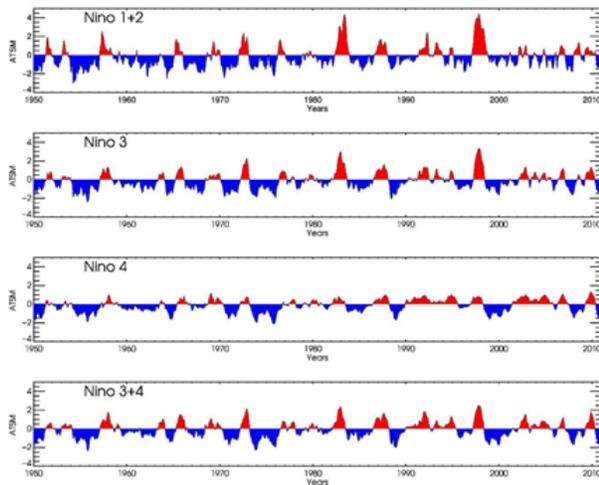


Figura 10: ATSM de las Regiones Oceánicas El Niño.

4.3 Patrones (No Estacionales) de Variabilidad

El intervalo considerado para el cálculo de las FOE será de 1950 a 2010. Para determinar las anomalías, se ha elegido como periodo base el intervalo de 1982 a 2010. Esta selección se justifica por tres motivos: (a) la alta calidad de las observaciones disponibles desde la era post-satélite y la red de boyas TAO, (b) consistencia necesaria en el cálculo de anomalías entre todas las bases de datos de TSM, ya que los datos OISST inician en 1982, y (c) lo recomienda la Organización Meteorológica Mundial (WMO), que sugiere que los periodos base tengan una duración aproximada de 30 años.

El primer modo o patrón de ATSM corresponde al modo canónico del ENOS, tal como lo describieron [40]. Este patrón se caracteriza por una amplia área de anomalías con el mismo signo, que se extiende a lo largo de la línea ecuatorial y frente a la costa sudamericana, y representa el 45.10% de la varianza total de la ATSM. La Figura 11 muestra el patrón predominante basado en los datos mensuales de ATSM obtenidos de HadISST.

Se obtienen resultados similares al emplear los datos de ERSST y OISST. El patrón espacial relacionado con la fase cálida del ENOS se caracteriza por anomalías positivas en la región oriental del Océano Pacífico Ecuatorial, rodeadas por débiles anomalías negativas en el extremo occidental del Pacífico Tropical. Este patrón (en forma de herradura) se extiende hacia el Pacífico Norte y Sur. Además, se encuentran anomalías positivas a lo largo de la costa oeste de América del Norte y del Sur.

El índice Niño 3-4 es prácticamente igual a la PC1 (serie temporal del primer modo de variabilidad dominante). Esta serie temporal, PC1, presenta una tendencia lineal ascendente y muestra una fuerte correlación con las anomalías en las regiones de El Niño (ver Figura 12).

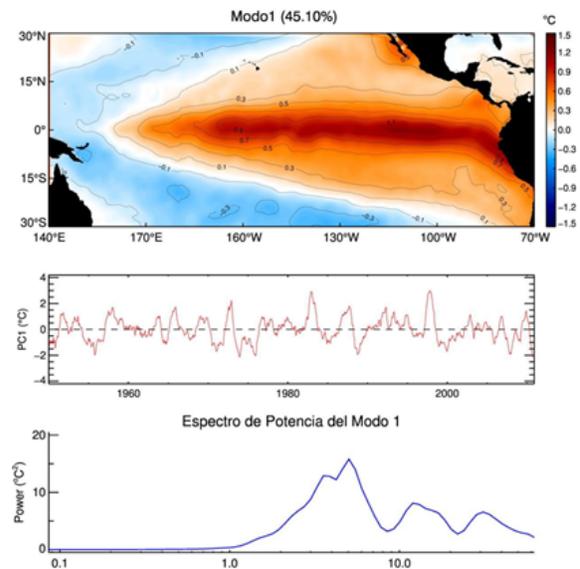


Figura 11: Primer patrón de variabilidad (espacial y temporal). Datos HadISST (1950-2010).

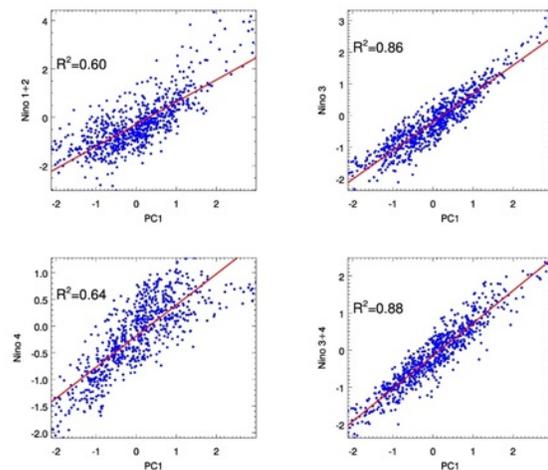


Figura 12: Correlación entre el primer modo (PC1) y la ATSM de las regiones oceánicas El Niño.

Las series de tiempo (PC1) de los datos OISST, HadISST y ERSST (Figura 13) muestran eventos fríos y cálidos, los cuales generalmente tienen una duración de aproximadamente 1 a 1.5 años y se repiten en intervalos de alrededor de 3 a 8 años. Los eventos cálidos más intensos (1982 y 1997) presentan amplitudes cercanas a 2°C - 3°C, mientras que los eventos cálidos de menor magnitud oscilan entre 1°C y 2°C. Por otro lado, los eventos fríos tienden a ser más prolongados y menos intensos que sus equivalentes cálidos. El espectro de potencia del primer patrón indica que el rango de periodos más dominante se sitúa entre 2.5 y 8 años.

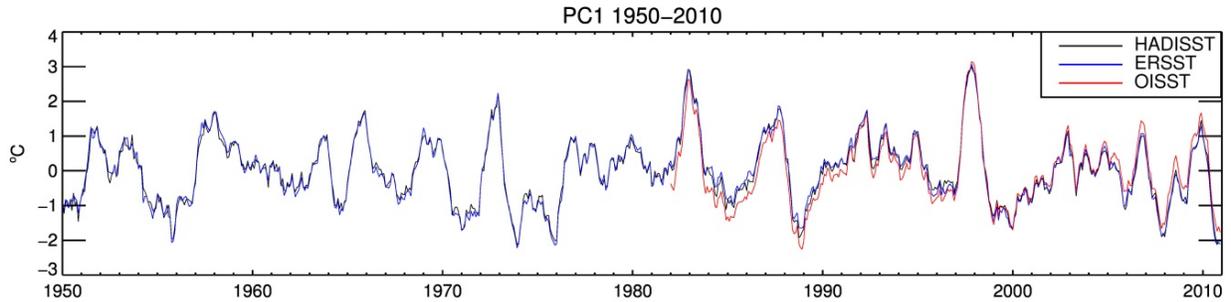


Figura 13: PC1 obtenido de los datos OISST, HadISST y ERSST en el periodo 1950-2010.

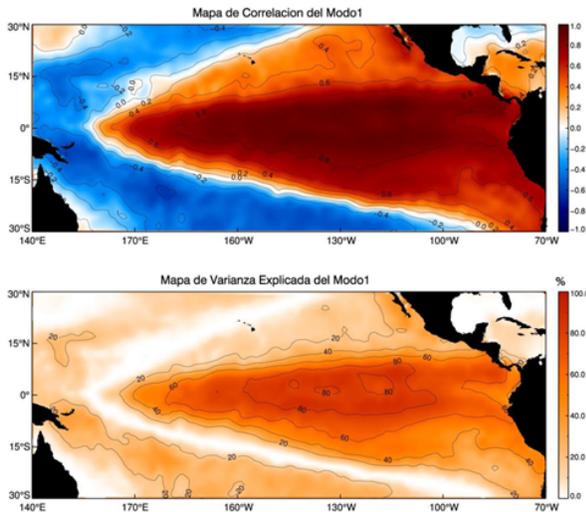


Figura 14: Distribuciones espaciales de la Correlación y Varianza explicada asociadas al primer modo.

La representación gráfica de correlación presentada en la Figura 14 ofrece una visión sobre las ubicaciones del dominio donde se observa una correlación más intensa o más débil entre la PC1 y las anomalías de los datos originales. Por otro lado, el gráfico de varianza explicada se obtiene al elevar al cuadrado los valores de correlación, revelando así las posiciones que tienen un mayor impacto en la varianza total del primer modo, que, en este caso, representa el 45.10 % de la varianza.

El segundo modo representa aproximadamente el 10 % de varianza explicada y refleja un dipolo zonal (ver Figura 15). Las anomalías en los extremos este y oeste presentan el mismo signo, mientras que las anomalías en el Centro del Pacífico Tropical muestran signos opuestos. En las latitudes altas, los valores negativos en el Pacífico Ecuatorial Central se propagan hacia el oeste tanto en el hemisferio Norte como en el Sur. En el estudio realizado por [10] se analizó datos de HadISST durante el período de 1979 a 2004 y se denominó a este modo como El Niño o La Niña MODOKI (un término japonés que se traduce como "similar pero diferente").

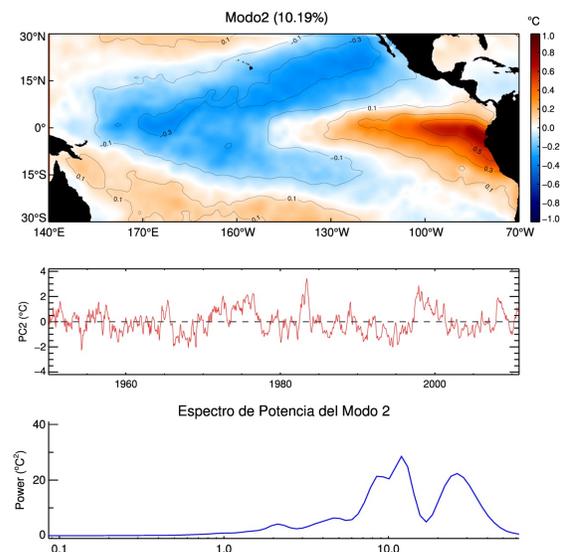


Figura 15: Segundo patrón de variabilidad (espacial y temporal). Datos HadISST (1950-2010).

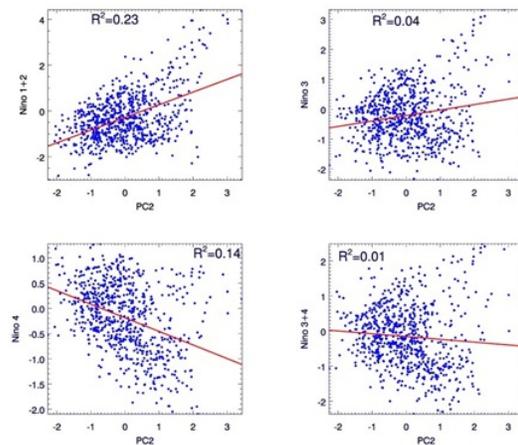


Figura 16: Correlación entre el segundo modo (PC2) y la ATSM de las regiones oceánicas El Niño.

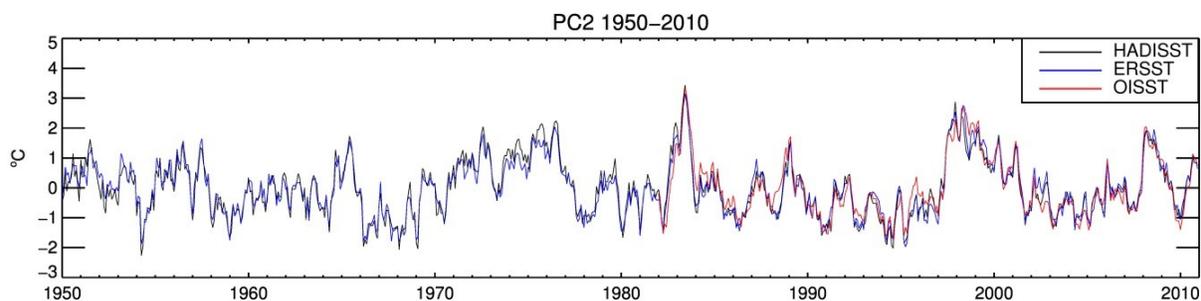


Figura 17: PC2 obtenido de los datos OISST, HadISST y ERSST en el periodo 1950-2010.

Esto se comprueba al analizar la Figura 16, donde los coeficientes de determinación son bajos entre la PC2 y las anomalías en las regiones El Niño. Existe evidencia que sugiere que este tipo de El Niño ha aumentado en frecuencia en las últimas décadas ([38], [41]). Este patrón también se asemeja al denominado "Modo del Pacífico Meridional", que puede anticipar eventos del ENOS ([42], [43]). En la Figura 17 se presenta la serie temporal de la PC2, obtenido a partir de los tres conjuntos de datos empleados, mostrando solo pocas variaciones en la serie del OISST, lo cual podría explicarse por el hecho de que esta base de datos solo contiene registros desde 1981. Es fundamental destacar que este modo difiere del primer modo, ya que se observa una señal decadal, claramente visible en el espectro de la Figura 15.

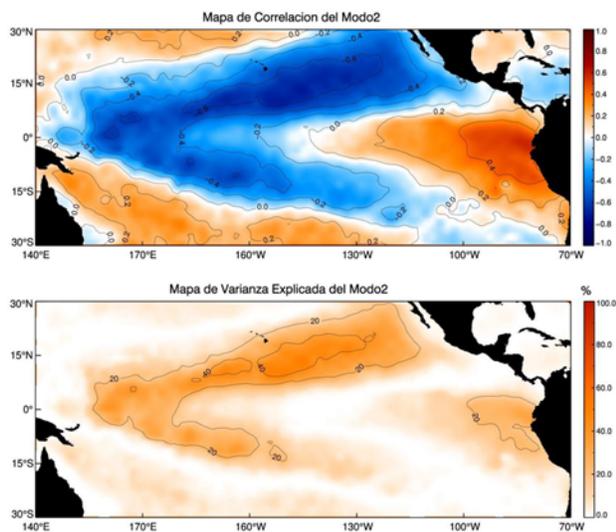


Figura 18: Distribuciones espaciales de la Correlación y Varianza explicada asociadas al segundo modo.

Similar a la Figura 14, se muestran las distribuciones de varianza explicada y correlación para el modo 2. Un aspecto destacado es la configuración dipolar que se manifiesta entre el Pacífico Oriental y el Pacífico central, donde ambas áreas aportan a la varianza explicada del

segundo modo, que asciende al 10.19% (ver Figura 18).

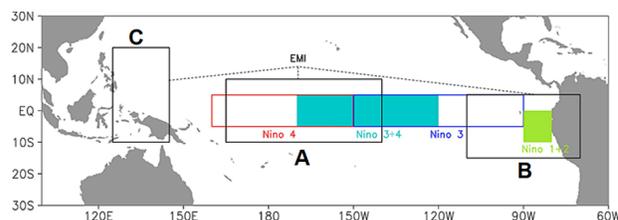


Figura 19: Regiones definidas por [10] para calcular el índice EMI

A partir del patrón temporal y espacial del modo 2, [10] desarrolló un índice conocido como el índice EMI, el cual se define de la siguiente manera:

$$EMI = [ATSM]_A - 0.5 * [ATSM]_B - 0.5 * [ATSM]_C$$

Las anomalías utilizadas en la ecuación previa están definidas sobre las siguientes áreas: A(10°S-10°N, 165°E-140°W), B(15°S-5°N, 110°W-70°W) y C(10°S-20°N, 125°E-145°E) respectivamente (ver Figura 19). La serie de tiempo del EMI se presenta en la Figura 20.

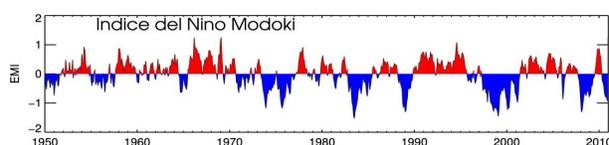


Figura 20: Serie de tiempo del índice EMI

Los coeficientes de determinación entre esta serie temporal y la PC2 es de 0.64 (ver Figura 21). Esta relación es inversa porque se ha tomado en cuenta el signo opuesto al calcular el segundo modo (tener en mente que el signo del modo espacial es arbitrario, dado que el producto entre el segundo modo y la PC2 recupera la polaridad adecuada). Además, considerar esta elección facilita la interpretación de las relaciones no lineales. Esta correlación indica que el índice presentado aquí es adecuado para representar los eventos EMI en el Pacífico Tropical.

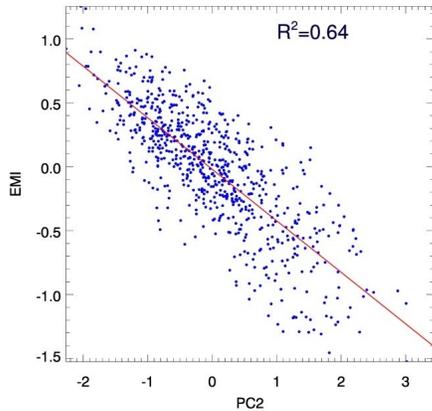


Figura 21: Diagrama de dispersión entre PC2 y el índice El Niño Modoki-EMI

Los primeros 2 modos representan más del 50% de la varianza total en los datos. ¿Por qué no considerar los modos subsiguientes? La respuesta se ilustra en el gráfico siguiente, que algunos textos de análisis multivariado llaman "gráfico de sedimentación" (ver Figura 22). Este gráfico es valioso porque facilita la identificación del número de modos relevantes para el análisis. A simple vista, es evidente que estos dos primeros modos se encuentran significativamente separados del resto. En cada punto del gráfico se agregó una barra que refleja el error en el cálculo de los autovalores.

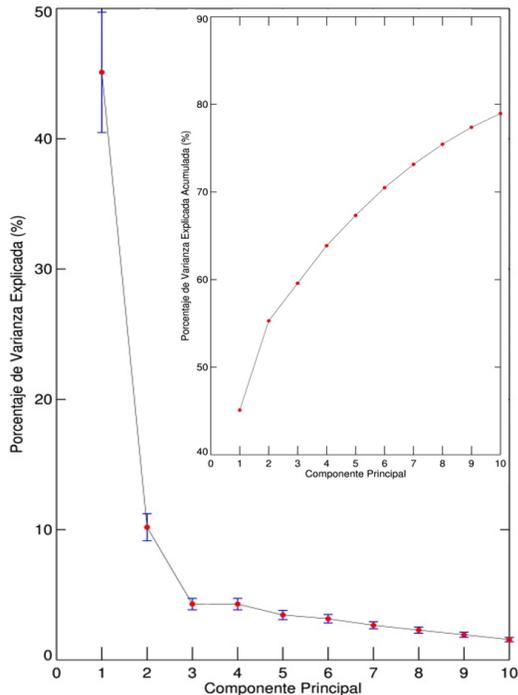


Figura 22: Varianza explicada y Varianza acumulada para cada modo. Datos HadISST

4.4 Estimación de los patrones asociados al Pacífico Central y Oriental

Dado que las varianzas explicadas de los 2 primeros modos están claramente diferenciadas, es razonable suponer que estos modos reflejan distintos tipos de variabilidad en el clima. No obstante, como se discutió en la sección 1.3, uno de los principales desafíos del análisis es la obligación de que los modos espaciales sean ortogonales y que los modos temporales no estén correlacionados. Es posible que el primer modo evite este tipo de complicaciones; sin embargo, los otros modos estarán expuestos a configuraciones artificiales en sus distribuciones espaciales. A continuación, se presentará un enfoque analítico para identificar el índice que mejor represente los modos de variabilidad en la TSM.

En cuanto a las discrepancias entre ENOS y EMI, se han realizado investigaciones significativas para identificar, comprender y describir estas clases de El Niño, lo que ha suscitado discusiones sobre la existencia real de dos modos diferentes de variabilidad. Las características distintivas asociadas con EMI, o los eventos cálidos en el Pacífico Central que no se repiten en los eventos ENOS en la parte oriental del Pacífico (Niño canónico), se documentan en [10], [32] y [34]. Por otro lado, algunos investigadores plantean que EMI es un artefacto estadístico que emerge como un modo dominante debido a la no linealidad entre La Niña y El Niño [9]. Esta hipótesis se apoya en el patrón similar al índice EMI que se obtiene al sumar los promedios de ATSM de El Niño y La Niña. Se realizaron los cálculos de manera similar a lo realizado en [44].

Se puede inicialmente resumir la no linealidad del modo ENOS dividiendo la ATSM en dos categorías, las cuales se definen por 2 características: la intensidad y el signo. Los promedios de estos patrones correspondientes a cada una de estas dos categorías se ilustran en las Figuras 23 y 24. Para identificar cada uno de los dos patrones a partir de los datos de ATSM de HadISST (con la tendencia lineal removida) correspondientes al período de 1950 a 2010, se aplicarán los siguientes criterios:

- a) Eventos El Niño Fuertes: Para los datos con ATSM del Niño3+4 $> 1^{\circ}\text{C}$
- b) Eventos La Niña Fuertes: Para los datos con ATSM del Niño3+4 $< -1^{\circ}\text{C}$

Es importante considerar que las ATSM más intensas suelen concentrarse en el Pacífico Central durante los episodios intensos de La Niña, mientras que se ubican en el Pacífico Oriental durante los eventos más intensos de El Niño. Esta asimetría de los promedios fue observada anteriormente por [44].

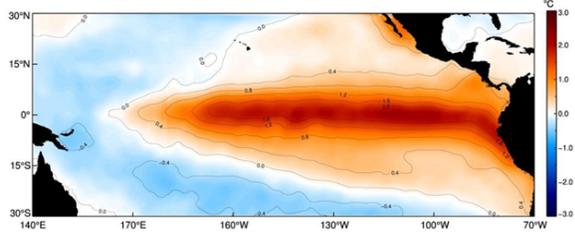


Figura 23: Promedio de los eventos El Niño fuertes usando el criterio dado en a).

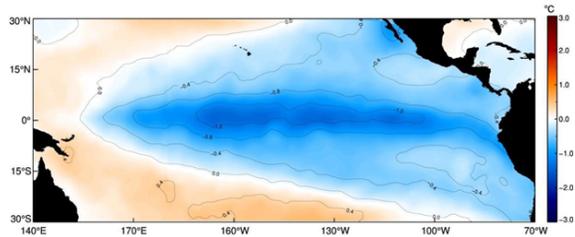


Figura 24: Promedio de los eventos La Niña fuertes usando el criterio dado en b).

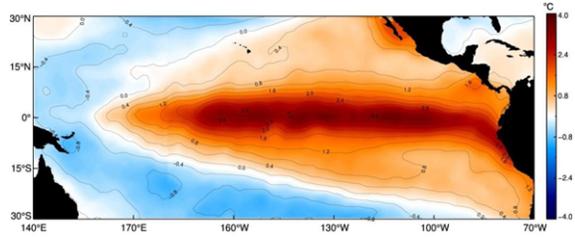


Figura 25: Modo espacial FOE1

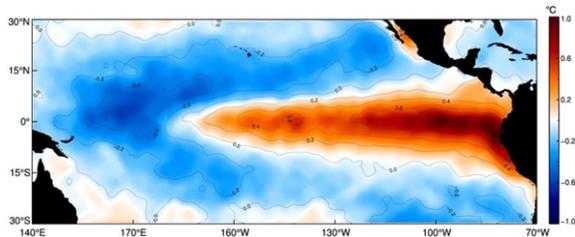


Figura 26: Modo espacial FOE2

Al comparar las Figuras 25 y 26 con las Figuras 11 y 15, notamos que la variación entre eventos intensos de La Niña y El Niño reproduce aproximadamente el primer modo espacial FOE1. Además, su combinación da lugar al mismo dipolo que se observa en el modo espacial FOE2. Esto sugiere que se podrían formular las siguientes ecuaciones:

$$FOE1 \approx \sum [ElNiño]_{fuerte} - \sum [LaNiña]_{fuerte}$$

$$FOE2 \approx \sum [ElNiño]_{fuerte} + \sum [LaNiña]_{fuerte}$$

Al considerar el modo del Pacífico Este como el resultado de la superposición de eventos El Niño fuertes, y el modo del Pacífico Central como la combinación de eventos La Niña fuertes, es posible reformular las relaciones establecidas anteriormente y definir nuevas ecuaciones:

$$Patrón\ Este \approx FOE1 + FOE2$$

$$Patrón\ Centro \approx -FOE1 + FOE2$$

De forma similar, se pueden emplear los modos temporales (PCs) para establecer lo siguiente:

$$PC_Este \approx PC1 + PC2$$

$$PC_Centro \approx -PC1 + PC2$$

Donde las PC's del Pacífico Central (PC_Centro) y del Pacífico Este (PC_Este) pueden ser calculadas mediante operaciones sobre las dos primeras componentes principales.

4.4.1 Índices C y E

De acuerdo con [11] y [12], los primeros FOEs de la ATSM en el Pacífico Tropical no representan fenómenos distintos (es decir ENOS y EMI), sino que reflejan el desarrollo no lineal del ENOS. Los patrones C y E representan eventos cálidos extremos en la parte este y eventos cálidos moderados o fríos en la parte central del Pacífico Ecuatorial, los cuales corresponden a diferentes evoluciones de los regímenes.

Se llevó a cabo un análisis de las series de tiempo PC1 y PC2. La gráfica de dispersión muestra la relación entre PC1 y PC2 (Figura 27) y, posteriormente, se examinó la evolución en el tiempo de los eventos extremos en estas series. Las series temporales de estos dos patrones temporales no están correlacionadas, pero evidentemente no son independientes (no forman una nube de puntos isotrópica, como se esperaría en interacciones lineales). En otras palabras la dispersión de los puntos debería conformar una distribución circular (en el origen de coordenadas). Sin embargo, se observa una distribución asimétrica, lo que indica que no hay evidencia que sugiera que estos patrones representan fenómenos nuevos y separados del ENOS, como afirma [10].

La Figura 27 muestra que los valores extremos, tanto positivos como negativos, de la PC1 están relacionados con los valores de la PC2 positivos. Físicamente, esto significa que las anomalías positivas más intensas, que ocurren durante un evento típico de El Niño, se localizan más cerca del límite este del Pacífico en comparación con las anomalías más frías que se presentan durante un evento típico de La Niña.

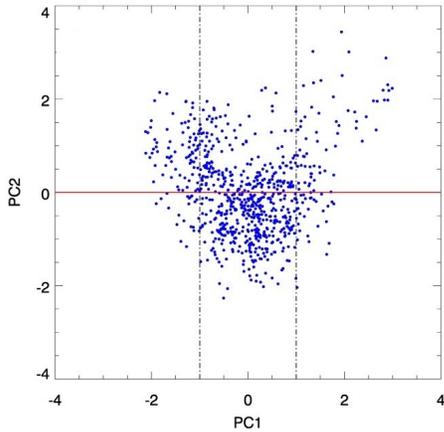


Figura 27: Diagrama de dispersión de las series de tiempo PC1 y PC2.

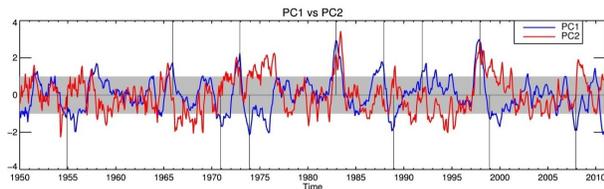


Figura 28: Series de tiempo de las primeras dos componentes principales (de manera similar a lo que presentó [12]).

Esto también puede observarse siguiendo los procedimientos descritos en [12]. Se grafican las dos series de tiempo, PC1 y PC2 (ver Figura 28). La franja gris marca el umbral que distingue los eventos intensos de los moderados. Las líneas verticales señalan seis eventos intensos de La Niña y El Niño. Se elaboró un promedio de la evolución en el tiempo de cada uno de estos eventos utilizando esos años (ver Figura 29), considerando diciembre como el mes central, el cual se representa con el número 0 en el eje de las abscisas. La Figura 29 muestra que PC2 y PC1 presentan una covariabilidad no lineal. Cada vez que la PC1 se halla en una fase extrema (ya sea un evento fuerte de La Niña o El Niño), la PC2 es positiva. Por otro lado, la PC2 tiende a exhibir valores negativos cuando la PC1 se aproxima a niveles neutros ($-1 < PC1 < 1$).

Tras identificar este comportamiento no lineal, se establece una relación entre las componentes principales y los índices C y E. En la Figura 27, los puntos de dispersión parecen mostrar dos direcciones principales perpendiculares (formando un ángulo de 45° con respecto a los ejes PC2 y PC1), según lo indicado en [11] y [12]. Estos ejes se presentan en la Figura 30.

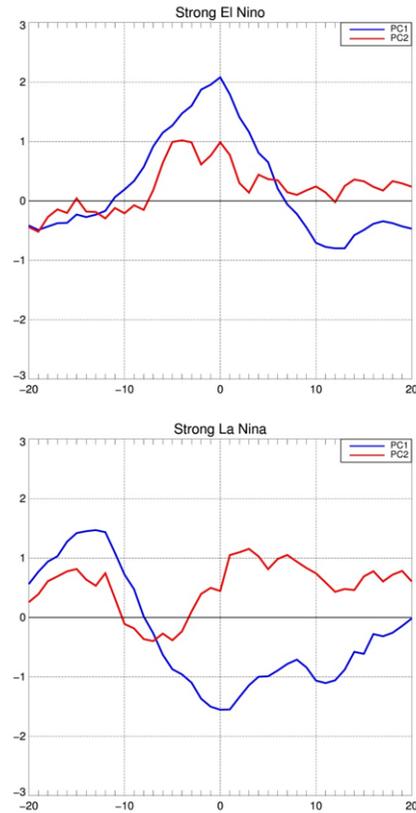


Figura 29: Promedio de los eventos El Niño y La Niña fuertes (evolución temporal). Similar a lo encontrado por [12]).

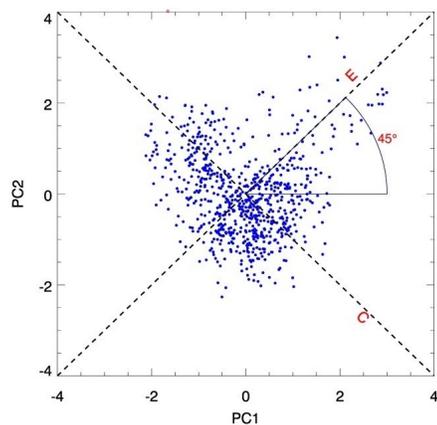


Figura 30: Ejes principales superpuestos sobre el diagrama de dispersión de PC2 vs PC1.

Para rotar las componentes principales PC2 y PC1 un ángulo de 45° , solo es necesario aplicar la ecuación:

$$E = PC1 \cos 45^\circ + PC2 \sen 45^\circ$$

$$C = -PC1 \sen 45^\circ + PC2 \cos 45^\circ$$

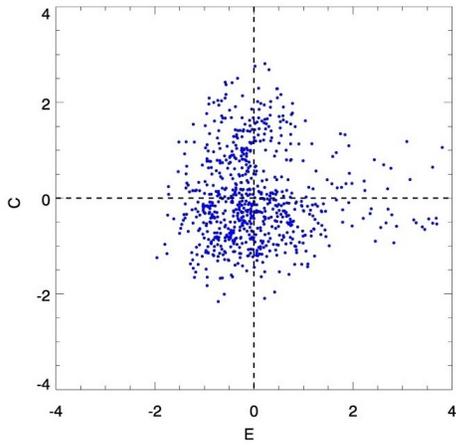


Figura 31: Diagrama de dispersión de los nuevos índices luego de la rotación (índices C y E).

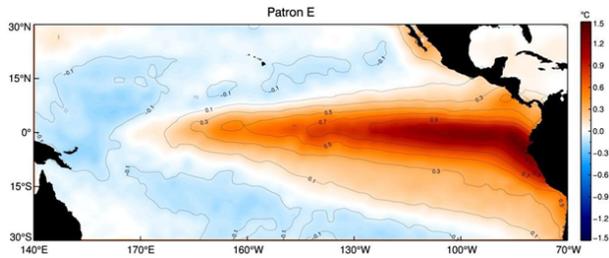


Figura 32: Patrón E (distribución espacial)

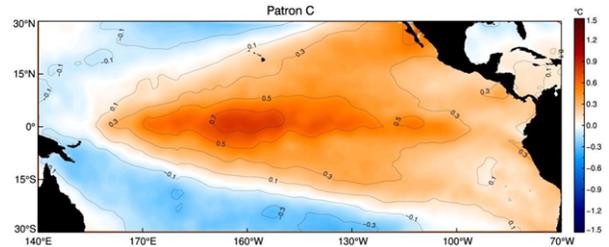


Figura 33: Patrón C (Distribución Espacial)

Una vez que se han generado estas nuevas series de tiempo, se lleva a cabo una regresión de los índices C y E con los datos de ATSM para derivar los modos espaciales correspondientes a C y E.

El modo E (ver Figura 32) se encuentra claramente desplazado hacia el este y más concentrado en el Ecuador en comparación con el FOE1, lo que es coherente con la discusión sobre la no linealidad de los promedios presentados en las Figuras 25 y 26. Por otro lado, el modo C (ver Figura 33), que se muestra con signo opuesto, está notablemente desplazado hacia el oeste y tiene una extensión meridional más amplia que el FOE1. Estos modos o patrones representan, en esencia, la distribución espacial no lineal de ENOS en una forma óptima.

Para verificar la validez del algoritmo utilizado para estimar los índices C y E, se efectúa una comparación entre ambos índices y los que fueron generados por el Instituto Geofísico del Perú (IGP), obteniendo correlaciones altas (0.9 y 0.86) para ambos índices (ver Figuras 34 y 35). Los datos se adquirieron a través del siguiente enlace: http://met.igp.gob.pe/datos/ecindex_ersstv5.txt

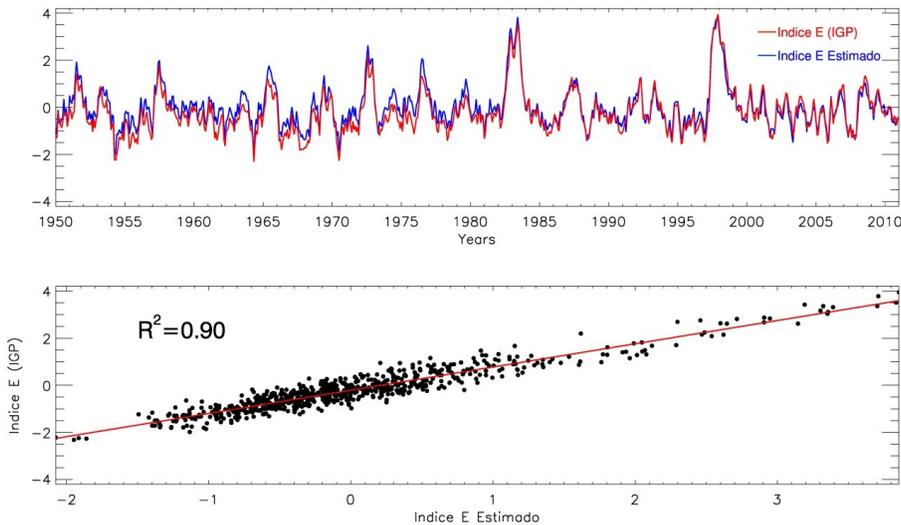


Figura 34: Comparativa entre el índice E del IGP y el índice E estimado.

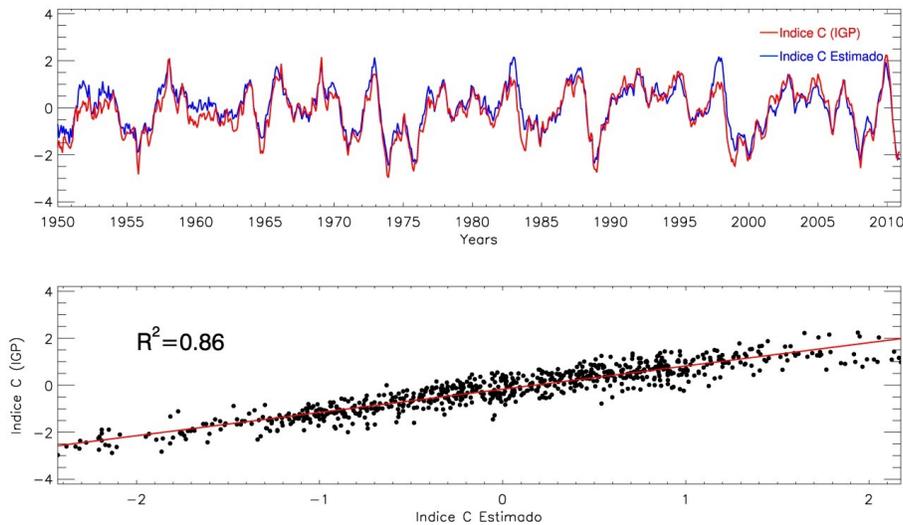


Figura 35: Comparativa entre el índice C del IGP y el índice C estimado.

4.4.2 Índices PC y PE

Se emplean datos HadISST y se comparan dos clases de ENOS en el Pacífico Tropical: uno correspondiente al Pacífico Oriental o Este (PE) y otro de la parte central del Pacífico (PC). El ENOS-PE presenta las anomalías cerca de la costa de Sudamérica. Este tipo de ENOS se relaciona con las variaciones en la termoclina y en los vientos superficiales, mostrando una conexión significativa con el Océano Índico Tropical. En contraste, en el ENOS-PC la ATSM y los vientos de superficie se ubican en el Pacífico Central. Esta clase de ENOS muestra una menor correlación con la dinámica de la termoclina y puede verse más afectado por forzantes de la atmósfera. Además, presenta una intensa teleconexión con la parte Sur del Océano Índico [32]. Para obtener los índices PC y PE, se utiliza un enfoque que combina las funciones ortogonales empíricas (FOE) con el análisis de regresión, para distinguir entre estas dos clases de ENOS.

La clase de ENOS-PE se define por anomalías que se distribuyen desde la costa sudamericana en el Pacífico Central a lo largo de la línea ecuatorial. Estas anomalías se concentran principalmente en las zonas oceánicas Niño3 y Niño1+2. La estructura resultante es comparable a la de la función ortogonal empírica (FOE) principal derivada de las anomalías originales de TSM (ver Figura 11), aunque es menos amplia en la parte Central del Pacífico (ver Figura 36).

La Figura 37 ilustra la configuración de la clase ENOS-PC, cuyas anomalías están mayormente restringidas al Pacífico Central, abarcando el área entre 120°W y 160°E, y abarcando las zonas Niño 4 y Niño 3+4.

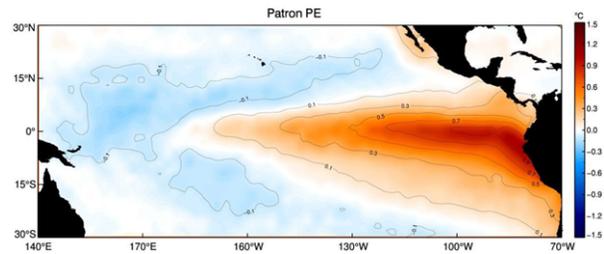


Figura 36: Patrón espacial PE.

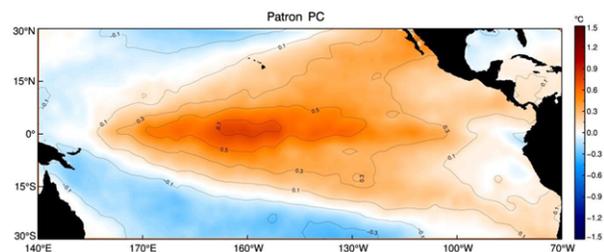


Figura 37: Patrón espacial PC.

Las series temporales (componentes principales) de los patrones ENOS-PC y ENOS-PE son representadas en la Figura 38. Se observa que ambas series presentan una falta de simetría en relación al valor cero. Los coeficientes de asimetría para las componentes principales de las clases PC y PE son -0.47 Y 1.3, respectivamente. Los histogramas que aparecen en la Figura 39 sugieren que la clase ENOS-PE se relaciona más con eventos intensos de El Niño, mientras que el ENOS-PC está más vinculado a eventos fuertes de La Niña. Esta asimetría en el espacio ha sido documentada en diversos estudios, como los

de [45] y [46]. Sin embargo, es importante tener en cuenta que también ocurren con frecuencia eventos cálidos de ENOS-PC y eventos fríos de ENOS-PE.

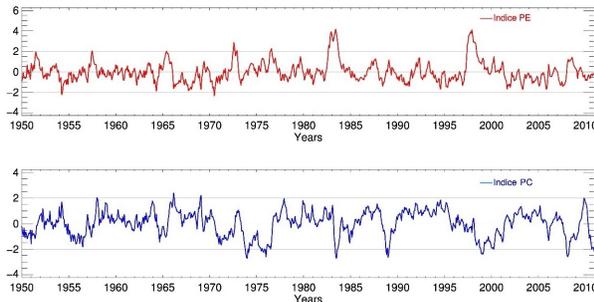


Figura 38: Series de Tiempo de los índices PC y PE.

Aparte de las variaciones en su configuración espacial, las categorías ENOS-PC y ENOS-PE también presentan discrepancias en sus evoluciones a lo largo del tiempo. En las Figuras 40 y 41 se ilustran los coeficientes de correlación desfasada (lag correlation) entre las componentes principales de las clases de ENOS (PC y PE) y la ATSM del Pacífico Tropical.

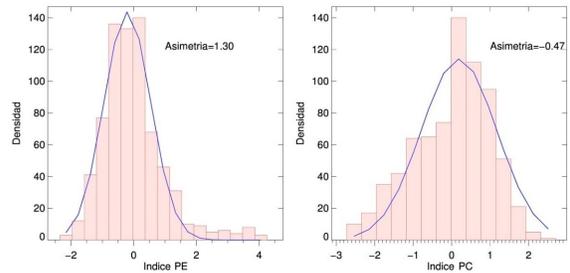


Figura 39: Histogramas de los índices PC y PE.

En el caso del ENOS-PE (ver Figura 40), las anomalías comienzan a manifestarse en las costas de Sudamérica, se extienden hacia el oeste, alcanzando el Pacífico Central, y su intensidad disminuye al alejarse del ecuador. Las anomalías más altas de esta categoría de ENOS se registran cerca de las costas, mientras que en el Pacífico Ecuatorial Oeste se observan anomalías débiles con signo contrario. En el caso del ENOS-PC (ver Figura 41), las anomalías se manifiestan alrededor de los 180°W, propagándose hacia los subtrópicos en los dos hemisferios, y posteriormente su intensidad se reduce en el centro del Pacífico Ecuatorial. Este tipo de ENOS presenta una conexión más fuerte con las anomalías en los subtrópicos en comparación con el ENOS-PE.

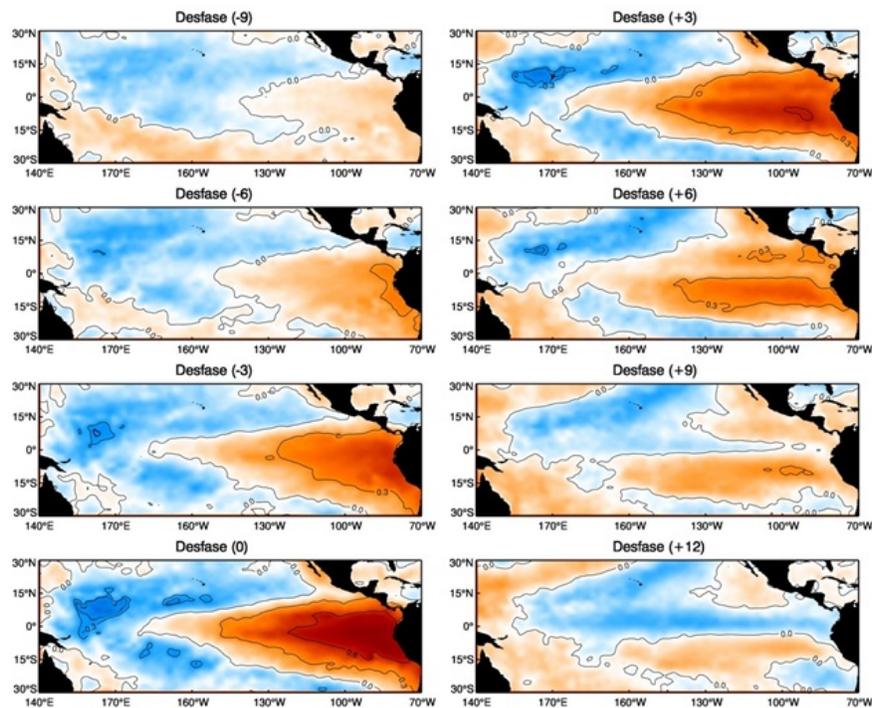


Figura 40: Correlación desfasada (lag correlation) entre la ATSM del Pacífico Tropical y el índice PE (parecido a lo presentado por [32]).

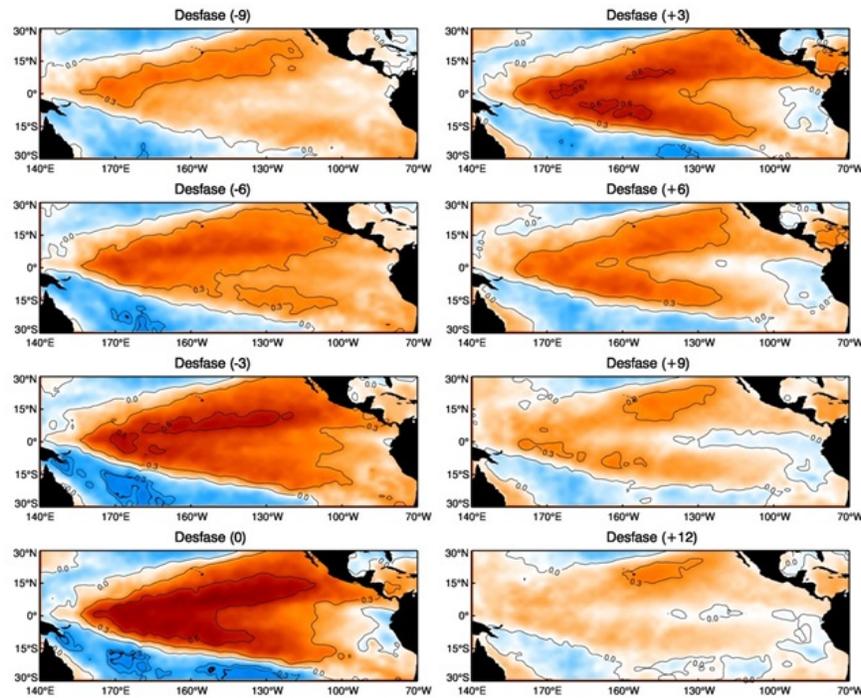


Figura 41: Correlación desfasada (lag correlation) entre la ATSM del Pacífico Tropical y el índice PC (parecido a lo presentado por [32]).

Frente a la costa de Perú, en el contexto de un fuerte ENOS-PE, se transportan aguas cálidas dentro del sistema de afloramiento, lo que provoca un aumento en la temperatura de la superficie del mar. En contraste, durante el ENOS-PC, no se produce dicho transporte y las temperaturas experimentan cambios leves frente a Perú [47]. Se realizó una comparación entre los ín-

dices estimados de PE y PC y los índices calculados por [32] (<https://www.ess.uci.edu/~yu/20SC/>), obteniendo una vez más altas correlaciones (0.83 y 0.88) para los índices de PC y PE (Figuras 42 y 43). Esto sugiere que los índices obtenidos reflejan adecuadamente ambas clases de ENOS (del Pacífico Este y del Pacífico Central).

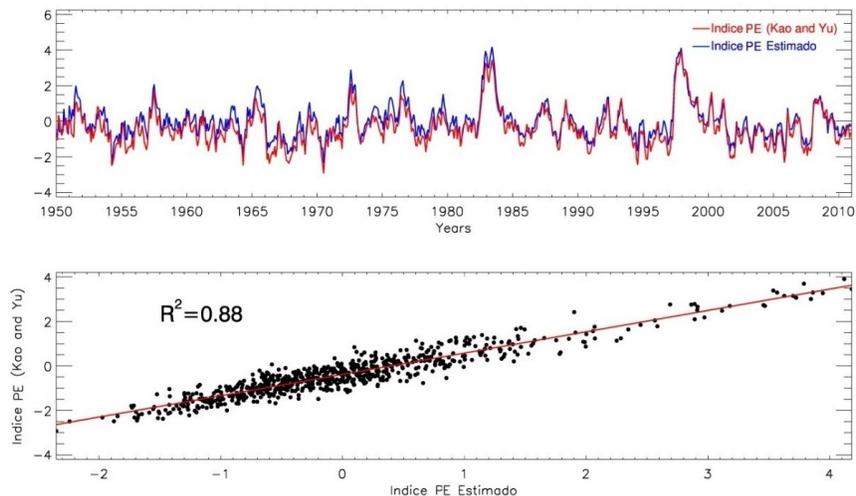


Figura 42: Comparativa entre el índice PE obtenido por [32] y el índice PE estimado.

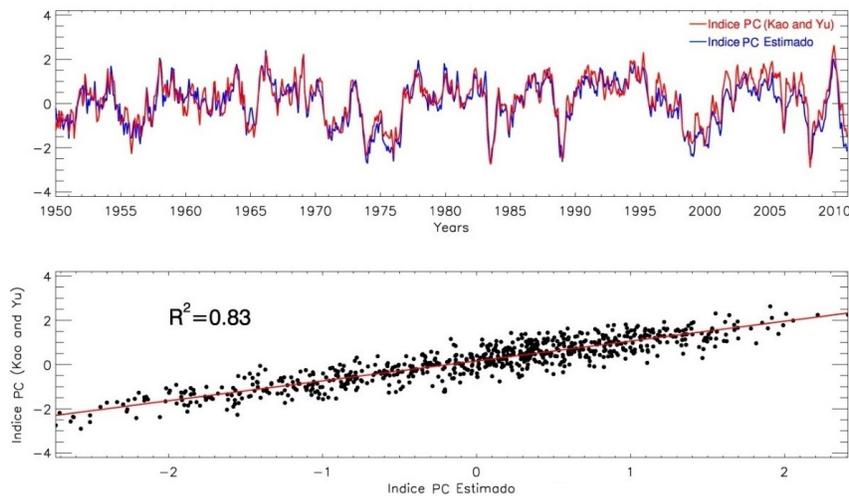


Figura 43: Comparativa entre el índice PC obtenido por [32] y el índice PC estimado.

Conclusiones y Recomendaciones

- El análisis de las Funciones Ortogonales Empíricas (FOE) permite reducir la dimensionalidad de los datos si es que estos aportan poca información y permite identificar fenómenos globales como EL NIÑO, LA NIÑA y EL NIÑO MODOKI, los cuales están representados en los primeros dos modos de variabilidad interanual.
- Aunque útil, el análisis de FOE no es infalible. Las FOE dependen del dominio espacial considerado, están sujetos a restricciones de ortogonalidad, y pueden no ser separables si representan porcentajes similares de la varianza.
- Los índices E y C son por construcción independientes (no correlacionados) y describen los regímenes que representan a) Eventos cálidos extremos y b) Eventos fríos ordinarios y eventos moderadamente cálidos.
- La magnitud de El Niño PC suele ser inferior a la de El Niño PE, mientras que La Niña PC tiende a ser más intensa que La Niña PE. Las causas de esta variación en el comportamiento entre las fases de El Niño y La Niña en estos dos tipos de ENOS siguen siendo objeto de estudio por parte de la comunidad científica.
- La posibilidad de que El Niño se manifieste en ocasiones en el Pacífico oriental, en otras en el Pacífico central, e incluso simultáneamente en ambas regiones, ha llevado a la hipótesis de que podrían existir múltiples clases de ENOS, cuyos mecanismos de formación y fenómenos climáticos extremos relacionados podrían ser distintos.
- Es fundamental registrar los cambios en la dinámica de los dos clases de El Niño para poder entender la tendencia a largo plazo de la variabilidad en el Pacífico tropical.

Referencias

- [1] R. Preisendorfer, and C. Mobley, Principal Component Analysis in Meteorology and Oceanography. Elsevier Amsterdam (1988).
- [2] K. Pearson, On lines and planes of closest fit to systems of points in space. Philosophical Magazine, Ser. 6, 2:559–72 (1901).
- [3] A. Fukuoka, A study of 10-day forecast (a synthetic report). Geophys. Mag., 22, 177-208 (1951).
- [4] E. Lorenz, Empirical orthogonal functions and statistical weather prediction. Sci. rep. no. 1, Statistical Forecasting Project, Dept. of Meteorology, MIT (1956).
- [5] K. Trenberth. A quasi-biennial standing wave in the southern hemisphere and interrelations with sea surface temperature. Quart. J. Roy. Meteor. Soc. 101, 55-74 (1975).
- [6] B. Weare, A. Navato, and R. Newell. Empirical Orthogonal Analysis of Pacific Sea Surface Tempe-

- ratures, *Journal of Physical Oceanography*: Vol. 6, No. 5, pp. 671–678 (1976).
- [7] G. Lagerloef, and R. Bernstein. Empirical orthogonal function analysis of advance very high resolution radiometer surface temperature patterns in Santa Barbara channel. *J. Geophys. Res.*, 93, 6863–6873 (1988).
- [8] T. C. Gallaudet, and J. J. Simpson. An empirical orthogonal function analysis of remotely sensed sea surface temperature variability and its relation to interior oceanic processes off Baja California. *Remote Sens. Environ.* 47: 375-389 (1994).
- [9] A. H. Monahan and A. Dai. The spatial and temporal structure of ENSO nonlinearity. *Journal Of Climate*, 17, 3026-3036 (2004).
- [10] K. Ashok, S. Behera, S. Rao, H. Weng, T. Yamagata. El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, 112, C11007 (2007).
- [11] K. Takahashi, A. Montecinos, K. Goubanova and B. Dewitte. ENSO regimes: Reinterpreting the canonical and Modoki El Niño. *Geophys. Res. Lett.*, 38, L10704 (2011).
- [12] D. Dommenges, T. Bayr, and C. Frauen. Analysis of the non-linearity in the pattern and time evolution of El Niño Southern Oscillation. *Climate Dyn.*, 40, 2825–2847 (2013).
- [13] R. E. Davis. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, 6, 249 (1976).
- [14] J. W. Kidson. Eigenvector Analysis of Monthly Mean Surface Data. *Monthly Weather Review*, 103(3):177-186 (1975).
- [15] C. Murray, L. Mueller, and H. Zwally. Matrix partitioning and EOF/principal component analysis of Antarctic sea ice brightness temperatures. National Aeronautics and Space Administration. Technical Memorandum 83916. Goddard Space Flight Center (1984).
- [16] B. C. Weare, J. S. Nasstrom. Examples of extended empirical orthogonal function analyses. *Mon. Weath. Rev.* 110, 481-485 (1982).
- [17] A. Navarra. *A Guide to Empirical Orthogonal Functions for Climate Data Analysis*. Springer (2010).
- [18] R. Thomson and W. Emery. *Data Analysis Methods in Physical Oceanography*. Elsevier Science (2014).
- [19] A. Shlens. Tutorial on Principal Component Analysis. Copy retrieved from: <https://arxiv.org/pdf/1404.1100>. (2014).
- [20] G. R. North. Empirical orthogonal functions and normal modes. *J. Atmos. Sci.*, 41, 879–887 (1984).
- [21] D. Wilks. *Statistical Methods in the Atmospheric Sciences*. Elsevier Science (2019).
- [22] J. P. Peixoto, and A. H. Oort. *Physics of Climate*, New York. American Inst. Physics, 520 pp (1992).
- [23] F. Husson, S. Lê, J. Pagès. *Exploratory Multivariate Analysis by Example Using R*. 2nd edition. Chapman and Hall/CRC (2017).
- [24] W. Hsieh. *Introduction to Environmental Data Science*. Cambridge University Press (2023).
- [25] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–41 (1933).
- [26] G. R. North, T. L. Bell, R. F. Cahalan and F. J. Moeng. Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review*, 110:699–706 (1982).
- [27] K. Kelly. Comment on ‘Empirical orthogonal function analysis of advanced very high-resolution radiometer surface temperature patterns in Santa Barbara Channel’ by G.S.E. Lagerloef and R.L. Bernstein. *Journal of Geophysical Research*, 93(C12):15743–54 (1988).
- [28] G. Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press (2023).
- [29] X. Kong, C. Hu, Z. Duan. *Principal Component Analysis Networks and Algorithms*. Springer (2017).
- [30] J. E. Overland, and R. Preisendorfer. A significance test for principal components applied to a cyclone climatology. *Mon. Wea. Rev.*, 110, 1–4 (1981).
- [31] S. G. Philander. *El Niño, La Niña, and the Southern Oscillation*. San Diego: Academic Press (1990).
- [32] H. Kao and J. Yu. Contrasting Eastern Pacific and Central Pacific types of ENSO. *J. Clim.* 22:615–632 (2009).
- [33] J. Yu, and S. Kim. Three evolution, patterns of central Pacific El Niño, *Geophys. Res. Lett.*, 37, L08706 (2010).
- [34] J. Kug, F. Jin, and S. An. Two types of El Niño events: Cold tongue El Niño and warm pool El Niño, *J. Clim.*, 22, 1499–1515 (2009).

- [35] H. Weng, S. Behera and T. Yamagata. Anomalous winter climate conditions in the Pacific Rim during recent El Niño Modoki and El Niño events, *Clim. Dyn.*, 32, 663–674 (2009).
- [36] H. Kim, P. Webster, and J. Curry. Impact of shifting patterns of Pacific Ocean warming on North Atlantic tropical cyclones, *Science*, 325,77–80 (2009).
- [37] K. Ashok and T. Yamagata. Climate change: The El Niño with a difference, *Nature*, 461, 481–484 (2009).
- [38] S. Yeh, J. Kug, B. Dewitte, M. Kwon, B. Kirtman and F. Jin. El Niño in a changing climate, *Nature*, 461, 511–515 (2009).
- [39] A. Capotondi, and Coauthors. Understanding ENSO diversity, *Bull. Amer. Meteor. Soc.* (2015).
- [40] E. M. Rasmusson and T. H. Carpenter. Variation in tropical sea surface temperature and surface wind fields associated with Southern Oscillation/El Niño, *Mon. Wea. Rev.*, 110, 354–384 (1982).
- [41] T. Lee and M. McPhaden. Increasing intensity of El Niño in the central-equatorial Pacific, *Geophys. Res. Lett.*, 37, L14603 (2010).
- [42] D. Vimont, J. Wallace and D. Battisti. The seasonal footprinting mechanism in the Pacific: implications for ENSO. *J. Clim* 16(16):2668–2675 (2003).
- [43] P. Chang, L. Zhang, R. Saravanan, D. J. Vimont, J. Chiang, L. Ji, H. Seidel and M. K. Tippett. Pacific meridional mode and El Niño–Southern Oscillation. *Geophys Res Lett* 34:L16608 (2007).
- [44] M. P. Hoerling, A. Kumar and M. Zhong. El Niño, La Niña, and the nonlinearity of their teleconnections. *J. Climate*, 10, 1769–1786 (1997).
- [45] A. Monahan. Nonlinear principal component analysis: Tropical Indo–Pacific Sea surface temperature and sea level pressure. *J. Climate*, 14, 219–233 (2001).
- [46] HW. Hsieh. Nonlinear multivariate and time series analysis by neural network methods. *Rev. Geophys.*, 42, RG1003 (2004).
- [47] B. Dewitte, J. Vazquez-Cuervo, K. Goubanova, S. Illig, K. Takahashi, G. Cambon, S. Purca, D. Correa, D. Gutierrez, A. Sifeddine and L. Ortlieb. Change in El Niño flavours over 1958–2008: Implications for the long-term trend of the upwelling off Peru, *Deep Sea Res., Part II*, 77–80, 143–156 (2012).