TECNICAS DE PRONOSTICOS

M.Sc. Christian W. Barreto Schuler

CONTENIDO

VII. FUNCIONES ORTOGONALES EMPÍRICAS/ANALISIS DE COMPONENTES PRINCIPALES

COVARIANZA

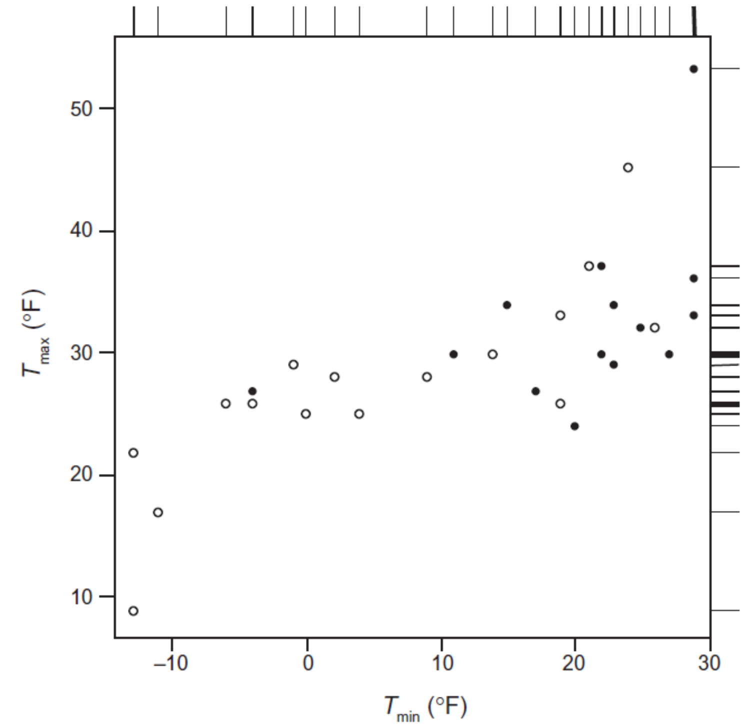


FIGURE 3.17 Scatterplot for daily maximum and minimum temperatures during January 1987 at Ithaca, New York. "Fringes" along the margins separately indicate the individual data distributions, with repeated data represented by heavier lines. Closed circles represent days with at least 0.01 in. of precipitation (liquid equivalent).

$$S_{X,Y} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{(n-1)}$$

Wilks, 2019

CORRELACION

Correlación de Pearson

Una forma de ver la correlación de Pearson es como la relación entre la covarianza muestral entre las dos variables, y el producto de las dos desviaciones estándar; donde los numeradores denotan anomalías, o resta de valores medios.

$$r_{x,y} = \frac{\mathrm{Cov}(x,y)}{s_x s_y} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}[(x_i-\overline{x})(y_i-\overline{y})]}{\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2\right]^{1/2}\left[\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})^2\right]^{1/2}},$$

$$= \frac{\sum_{i=1}^{n}(x_i' y_i')}{\left[\sum_{i=1}^{n}(x_i')^2\right]^{1/2}\left[\sum_{i=1}^{n}(y_i')^2\right]^{1/2}}$$

Varianza (S^2)	$s^2 = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}$
Desviación estándar (S)	$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}}$

$$r_{x,y} = \frac{1}{n-1}\sum_{i=1}^{n}\left[\frac{(x_i-\overline{x})}{s_x}\frac{(y_i-\overline{y})}{s_y}\right] = \frac{1}{n-1}\sum_{i=1}^{n} z_{x_i} z_{y_i},$$

Wilks, 2019

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

Statistical Methods for Signal Detection in Climate

Silvia A. Venegas

Danish Center for Earth System Science (DCESS),
Niels Bohr Institute for Astronomy, Physics and Geophysics,
University of Copenhagen, Denmark.

DCESS Report # 2

January, 2001

PREPARACION DE DATOS

$$\text{Time} \longrightarrow$$

$$\mathbf{F} = \begin{bmatrix} F_1(1) & F_1(2) & \dots & F_1(N) \\ F_2(1) & F_2(2) & \dots & F_2(N) \\ \dots & \dots & \dots & \dots \\ F_M(1) & F_M(2) & \dots & F_M(N) \end{bmatrix} \Big\downarrow \text{Location}$$

$$F_m(t) = \frac{\psi_m(t) - \mu_m}{\sigma_m} \tag{2.1}$$

where μ_m is the record mean:

$$\mu_m = \frac{1}{N} \sum_{t=1}^{N} \psi_m(t) \tag{2.2}$$

and σ_m is the record standard deviation:

$$\sigma_m = \left[\frac{1}{N-1} \sum_{t=1}^{N} \psi_m^2(t) \right]^{1/2} \tag{2.3}$$

Matriz de F de anomalías

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

The Covariance Matrix Approach

The Singular Value Decomposition Approach

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

The Covariance Matrix Approach

Matriz de Covarianza $\longrightarrow$ $\mathbf{R_{FF}} = \mathbf{F} * \mathbf{F}^{\dagger}$

$$\mathbf{R_{FF}} = \begin{bmatrix} \langle F_1 F_1 \rangle & \langle F_1 F_2 \rangle & \dots & \langle F_1 F_M \rangle \\ \langle F_2 F_1 \rangle & \langle F_2 F_2 \rangle & \dots & \langle F_2 F_M \rangle \\ \dots & \dots & \dots & \dots \\ \langle F_M F_1 \rangle & \langle F_M F_2 \rangle & \dots & \langle F_M F_M \rangle \end{bmatrix}$$

$$\langle F_i F_j \rangle = \langle F_j F_i \rangle = \frac{1}{N-1} \sum_{t=1}^{N} F_i(t)\, F_j(t)$$

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

The Covariance Matrix Approach

$$\mathbf{R_{FF}} * \mathbf{E} = \mathbf{E} * \boldsymbol{\Lambda}$$

MxM MxM

Eigen Value Problem

loadings

$$\mathbf{E} = \begin{bmatrix} E_1^1 & E_1^2 & \dots & E_1^M \\ E_2^1 & E_2^2 & \dots & E_2^M \\ \dots & \dots & \dots & \dots \\ E_M^1 & E_M^2 & \dots & E_M^M \end{bmatrix}$$

$$\begin{array}{cccc} \downarrow & \downarrow & & \downarrow \\ E^1 & E^2 & & E^M \end{array} \longrightarrow \text{Eigenvectors } E^k$$

Eigenvectors

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

Eigenvalues

Componentes Principales

$$A^k(t) = \sum_{m=1}^{M} E_m^k \, F_m(t)$$

$$\mathbf{A} = \mathbf{E}^\dagger * \mathbf{F}$$

MxN

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

The Covariance Matrix Approach

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

The original field F can be totally reconstructed by multiplying each EOF pattern E^k by its corresponding principal component A^k and adding the products over all K modes:

$$F_m(t) = \sum_{k=1}^{K} E_m^k A^k(t)$$

In matrix notation:

$$\mathbf{F} = \mathbf{E} * \mathbf{A}$$

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

The Singular Value Decomposition Approach

$$\mathbf{F} = \mathbf{U} * \boldsymbol{\Gamma} * \mathbf{V}^{\dagger}$$

$\mathbf{F}$	$\mathbf{U}$	$\boldsymbol{\Gamma}$	$\mathbf{V}^{\dagger}$
MxN	MxM	MxN	NxN

Matriz de anomalias

Proporcional a componentes Principales

Singular vectors, equivalentes a eigenvectors

Singular values, proporcional a eigenvalues

Componentes Principales

$$\mathbf{A} = \boldsymbol{\Gamma} * \mathbf{V}^{\dagger}$$

$$A^k(t) = \gamma_k \, V^{\dagger k}(t)$$

we can reconstruct field F adding all K modes of the decomposition as:

$$F_m(t) = \sum_{k=1}^{K} U_m^k \, \gamma_k \, V^{\dagger k}(t)$$

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

UNIDADES Y PRESENTACION

$$A^{k*}(t) = \frac{A^k(t)}{\sqrt{\lambda_k}}$$

$$E_m^{k*} = E_m^k \sqrt{\lambda_k}$$

$$\mathbf{r}_m^k = [A^k(t),\ F_m(t)]$$

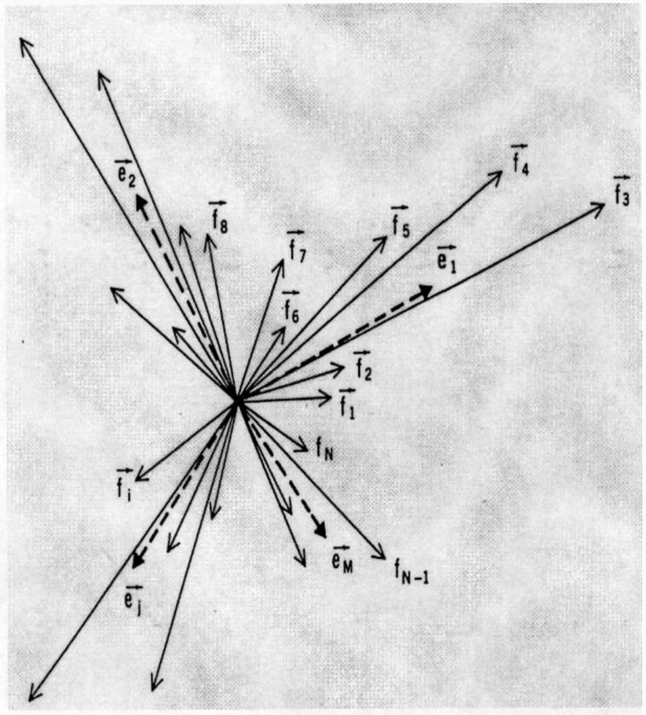


Figure 2.1: Example of a possible configuration of the data vectors f_n ($n = 1 \ldots N$ denote the time steps) and the empirical orthogonal vectors e_m, $m = 1 \ldots M$. From Peixoto and Oort (1992)

PRACTICA

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

Combined Empirical Orthogonal Functions

$$\mathbf{F} = \begin{bmatrix} S_1(1) & S_1(2) & \dots & S_1(N) \\ S_2(1) & S_2(2) & \dots & S_2(N) \\ \dots & \dots & \dots & \dots \\ S_{M_s}(1) & S_{M_s}(2) & \dots & S_{M_s}(N) \\ P_1(1) & P_1(2) & \dots & P_1(N) \\ P_2(1) & P_2(2) & \dots & P_2(N) \\ \dots & \dots & \dots & \dots \\ P_{M_p}(1) & P_{M_p}(2) & \dots & P_{M_p}(N) \end{bmatrix}$$

$$\mathbf{R_{FF}} = \begin{bmatrix} \langle S_1S_1\rangle & \dots & \langle S_1S_{M_s}\rangle & \langle S_1P_1\rangle & \dots & \langle S_1P_{M_p}\rangle \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \langle S_{M_s}S_1\rangle & \dots & \langle S_{M_s}S_{M_s}\rangle & \langle S_{M_s}P_1\rangle & \dots & \langle S_{M_s}P_{M_p}\rangle \\ \langle P_1S_1\rangle & \dots & \langle P_1S_{M_s}\rangle & \langle P_1P_1\rangle & \dots & \langle P_1P_{M_p}\rangle \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \langle P_{M_p}S_1\rangle & \dots & \langle P_{M_p}S_{M_s}\rangle & \langle P_{M_p}P_1\rangle & \dots & \langle P_{M_p}P_{M_p}\rangle \end{bmatrix}$$

$$\mathbf{E} * \mathbf{R_{FF}} = \mathbf{\Lambda} * \mathbf{E}$$

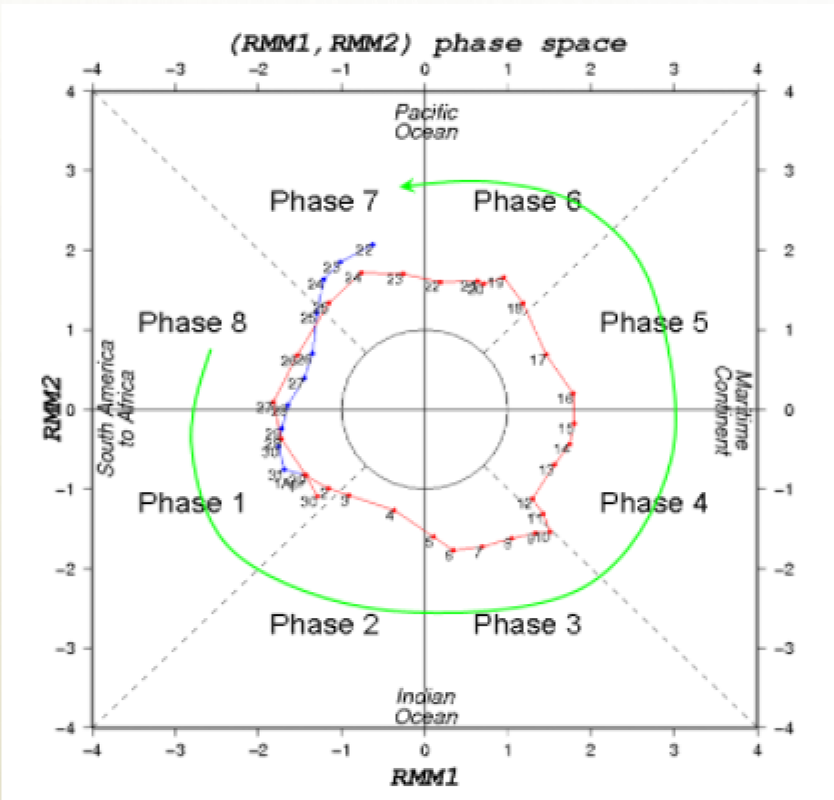
$$\mathbf{R_{SP}} = \begin{bmatrix} \langle S_1P_1\rangle & \langle S_1P_2\rangle & \dots & \langle S_1P_{M_p}\rangle \\ \langle S_2P_1\rangle & \langle S_2P_2\rangle & \dots & \langle S_2P_{M_p}\rangle \\ \dots & \dots & \dots & \dots \\ \langle S_{M_s}P_1\rangle & \langle S_{M_s}P_2\rangle & \dots & \langle S_{M_s}P_{M_p}\rangle \end{bmatrix}$$

$$\mathbf{R_{SP}} = \mathbf{U} * \mathbf{\Gamma} * \mathbf{V}^{\dagger}$$

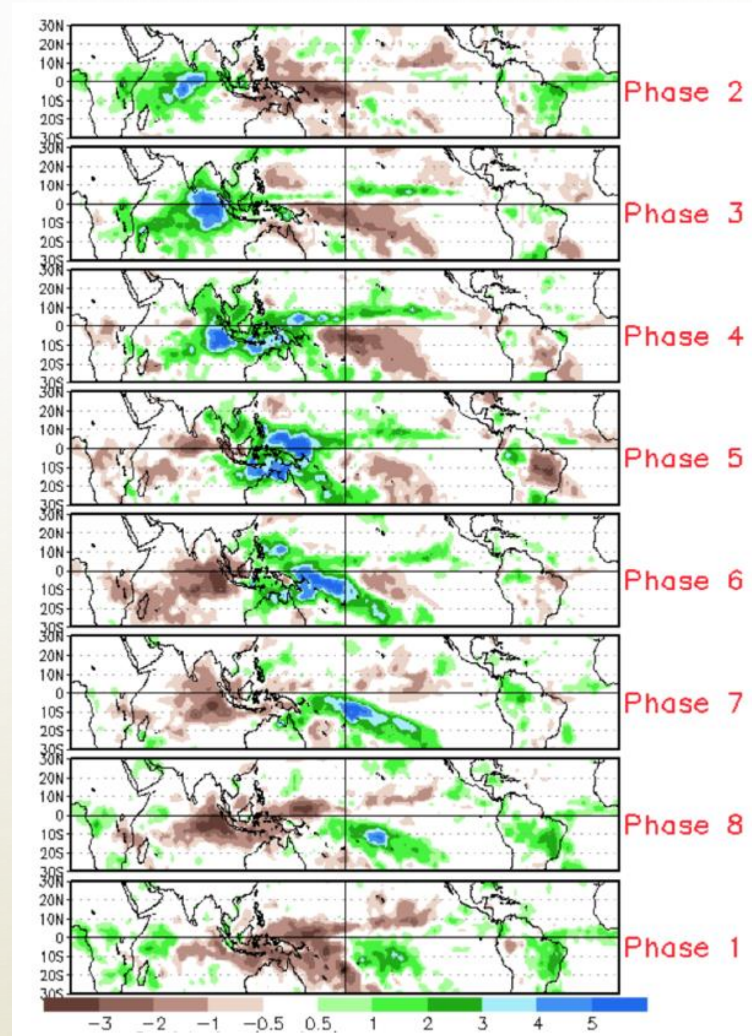
Combined Empirical Orthogonal Functions

EJEMPLO: MJO INDEX

Campos: WIND850 ,WIND200, OLR



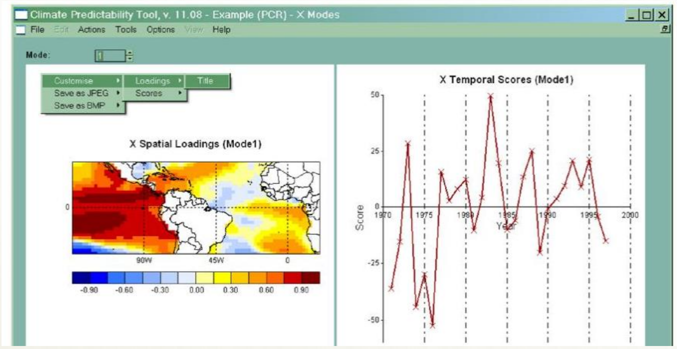
https://ds.data.jma.go.jp/tcc/tcc/products/clisys/mjo/explanation_mjo.html



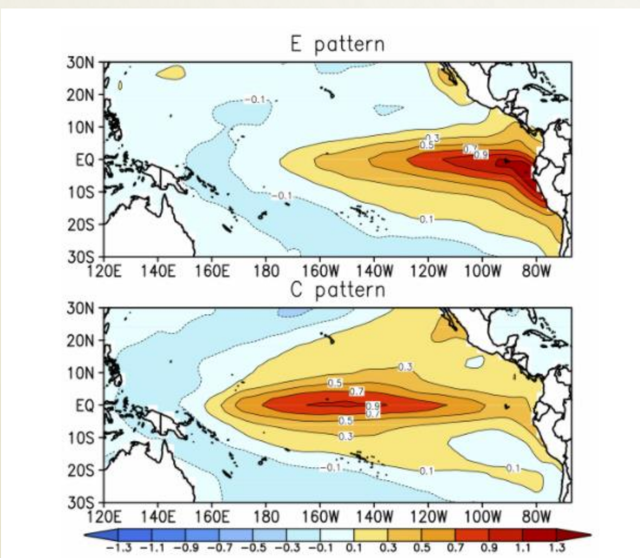
https://www.climate.gov/news-features/blogs/enso/what-mjo-and-why-do-we-care

ANALISIS DE FUNCIONES ORTOGONALES EMPÍRICAS / COMPONENTES PRINCIPALES

OTRAS APLICACIONES



https://iri.columbia.edu/wp-content/uploads/2013/07/CPT_Tutorial.pdf



http://met.igp.gob.pe/elnino/indices.html

An improved methodology for filling missing values in spatiotemporal climate data set

Application to Tanganyika Lake data set

Antti Sorjamaa · Amaury Lendasse · Yves Cornet · Eric Deleersnijder

Received: 20 April 2008 / Accepted: 22 January 2009 / Published online: 12 February 2009

Fig. 7 Southern part of the Tanganyika Lake data set. The pictures are the same examples filled by the EOF pruning as in Fig. 3