



Análisis de Componentes Principales Funcionales (FPCA) de Precipitaciones en el Lago de Valencia

Lugo U. Andrés A.^{1,2*}, Marcano José G.^{1,2}

¹*Departamento de Matemáticas, Facultad de Ciencias y Tecnología, Universidad de Carabobo. Naganagua, Edo. Carabobo-Venezuela*

²*Centro Multidisciplinario de Visualización y Cómputo Científico (CEMVICC)*

***Autor de correspondencia:** adlugo.ven@gmail.com

Recibido: 29/01/2025, Revisado: 22/02/2025, Aceptado: 08/05/2025

Resumen

En este trabajo se presenta un análisis de los datos mensuales de precipitaciones provenientes de 11 estaciones meteorológicas ubicadas en el Lago de Valencia, Venezuela utilizando el método de componentes principales funcionales (FCP). Este proceso consiste en transformar los datos de cada estación en curvas funcionales mediante técnicas de suavizado usando bases de Fourier debido a la naturaleza periódica presente en los datos. Al realizar el análisis de componentes principales se obtuvo que sólo dos componentes contribuyen con más del 85% de variabilidad y de las cuales se identifican dos periodos bien definidos de lluvia y sequía que contribuyen a comprender mejor las variaciones climáticas en esta región.

Palabras Claves: Precipitaciones, componentes principales funcionales, Lago de Valencia

Functional Principal Component Analysis (FPCA) of Rainfall in Lake Valencia

Abstract

This paper presents an analysis of monthly rainfall data from 11 meteorological stations located in Lake Valencia, Venezuela, using the functional principal component analysis (FPCA) method. This process consisted of transforming the data from each station into functional curves by means of smoothing techniques using Fourier bases, due to the periodic nature present in the data. By performing the principal component analysis, it was found that only two components contribute more than 85% of the variability, and from which two well-defined periods of rain and drought are identified, which contributes to a better understanding of the climatic variations in this region.

Keywords: Precipitation, Functional Principal Components, Lake Valencia

1. Introducción

El análisis de las precipitaciones en el Lago de Valencia ha sido abordado desde diversas perspectivas. Estudios previos han empleado metodologías variadas para comprender el régimen climático y los patrones de lluvia en la región.

En [3], se realizó un estudio del régimen climático y el patrón espacial de lluvias, utilizando “Análisis de correlación espacial” y “los criterios de error de la lluvia media y la interpolación espacial del método de Kagan”.

En [6] y [10], se analizaron datos de ocho estaciones ubicadas en la cuenca del Lago de Valencia, aplicando técnicas de análisis de series de tiempo.

En [13], se utilizaron series temporales para examinar diversas variables meteorológicas en el Lago de Valencia, incluyendo las precipitaciones.

Además de los estudios específicos del Lago de Valencia, se han realizado investigaciones sobre el análisis de datos funcionales (FDA) y, en particular, el análisis de componentes principales funcionales (FPCA) en el contexto de estudios de precipitaciones.

En [11], se estudiaron las precipitaciones en Canadá utilizando FPCA, con el apoyo de un paquete estadístico en el software **R**. Este estudio también se presenta como ejemplo en [1], donde se encuentran fundamentos teóricos más detallados. Estas serán las principales fuentes para el presente trabajo.

En [1], [2], [4] y [5], se aplicaron diversos métodos de análisis de datos funcionales en estudios relacionados con el clima y las precipitaciones.

En este trabajo, se presenta un análisis de los datos mensuales de precipitaciones de 11 estaciones meteorológicas en el Lago de Valencia, Venezuela, utilizando el método de componentes principales funcionales (FPCA). Este proceso transforma los datos de cada estación en curvas funcionales mediante técnicas de suavizado con bases de Fourier, dada la periodicidad de los datos. El análisis de componentes principales reveló que dos componentes explican más del 85 % de la variabilidad, identificando dos períodos definidos, de lluvia y sequía, lo que contribuye a una mejor comprensión de las variaciones climáticas en la región.

La aplicación del FPCA permitirá reducir la dimensionalidad de los datos, identificar las principales fuentes de variabilidad en las precipitaciones y detectar patrones espaciales y temporales significativos. Los resultados de este estudio tendrán implicaciones importantes para la gestión de recursos hídricos, optimizando el uso del agua y mitigando los efectos de eventos extremos.

2. Bases teóricas

En esta primera sección vemos un breve resumen de la teoría que sustenta este trabajo. Se sigue aquí a [8] debido a ser más conciso en lo que nos ocupa. En [14] se puede encontrar una gran variedad de citas que muestran el desarrollo histórico del análisis de datos funcionales (FDA) y del análisis de componentes principales funcionales (FPCA). Para un estudio más completo se tienen también [9] y [12].

Variables aleatorias en L^2

El espacio L^2 es el conjunto de todas las funciones x definidas en $[0, T]$ que satisfacen que $\int_0^T x^2(t)dt < \infty$. El espacio L^2 es un espacio separable de Hilbert con producto interno,

$$\langle x, y \rangle = \int_0^T x(t)y(t)dt.$$

Una variable (o curva) aleatoria $X = \{X(t), t \in [0, T]\}$ es un elemento aleatorio de L^2 equipado con una σ -álgebra de Borel.

Si X y Y son variables aleatorias, estas son independientes si

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Decimos que X es integrable si

$$E\|X\|^2 = E \int_0^T X^2(t)dt < \infty$$

Se debe notar que: $\int_0^T X^2(t)dt$ es una variable aleatoria real.

Si X es integrable entonces existe $\mu \in L^2$ tal que,

$$\langle \mu, x \rangle = E\langle X, x \rangle,$$

para cada $x \in L^2$. μ se denomina «esperanza de X », y se denota EX . EX conmuta con operadores acotados, es decir, si $\ell \in \mathcal{L}$ y X son integrables, entonces

$$E(\ell(X)) = \ell(EX).$$

Si $EX = 0$, el operador de covarianza de X se define como,

$$C(x) = E[\langle X, x \rangle X], \quad \forall x \in L^2.$$

El operador C es simétrico y positivo definido, es decir,

$$\begin{aligned} \langle C(x), y \rangle &= \langle x, C(y) \rangle, \forall x, y \in L^2. \text{ Simétrico.} \\ \langle C(x), x \rangle &\geq 0, \forall x \in L^2. \text{ Definido positivo.} \end{aligned}$$

El operador de covarianza es además compacto, es decir, dadas dos bases ortonormales de L^2 , $\{u_j\}$ y $\{v_j\}$ se puede escribir como

$$C(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, v_j \rangle u_j.$$

El operador de covarianza es de Hilbert-Schmidt, es decir, se verifica la condición

$$\sum_{j=1}^{\infty} \lambda_j^2 < \infty.$$

Si se observa una muestra de N curvas $X_1, \dots, X_N$, estas pueden verse como las realizaciones de una función aleatoria X o como un elemento aleatorio de L^2 con la misma distribución de X .

Se asume a menudo que $X_1, \dots, X_N$ son iid en L^2 , y se distribuyen igual a X , que asumimos es cuadrado integrable. Así definimos los siguientes parámetros:

$$\begin{aligned} \mu(t) &= E[X(t)] \quad (\text{función media}), \\ c(t, s) &= E[(X(t) - \mu(t))(X(s) - \mu(s))] \\ &\quad (\text{función de covarianza}), \\ C(\cdot) &= E[(X - \mu), \cdot](X - \mu) \\ &\quad (\text{operador de covarianza}). \end{aligned}$$

La función media μ se estima por la función de media muestral

$$\hat{\mu}(t) = \frac{1}{N} \sum_{i=1}^N X_i(t) = \frac{X_1(t) + \dots + X_N(t)}{N}, \quad (1)$$

y la función de covarianza,

$$\hat{c}(t, s) = \frac{1}{N} \sum_{i=1}^N (X_i(t) - \hat{\mu}(t))(X_i(s) - \hat{\mu}(s)).$$

El operador de covarianza muestral se define como

$$\hat{C}(x) = \frac{1}{N} \sum_{i=1}^N \langle X_i - \hat{\mu}, x \rangle (X_i - \mu), \quad x \in L^2. \quad (2)$$

El operador $\hat{C}$ cumple las mismas propiedades que el operador C , de esta manera, el operador $\hat{C}$ puede escribirse como,

$$\hat{C}(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, v_j \rangle v_j$$

donde λ_j son los valores propios de $\hat{C}$ y v_j sus vectores propios, que forman una base ortonormal de L^2 .

Los valores propios de $\hat{C}$ se asumen positivos y monótonamente decrecientes, es decir,

$$\lambda_1 > \lambda_2 > \dots$$

Los vectores propios de $\hat{C}$ se denominan «componentes principales funcionales empíricos» (EFPC) de los datos $X_1, \dots, X_N$.

Componentes principales

En la subsección anterior se vio que los EFPC se pueden interpretar como una base óptima ortonormal con respecto a la cual se pueden expandir los datos.

El producto interno $\langle X_i, \hat{v}_j \rangle = \int X_i(t) \hat{v}_j(t) dt$ se denomina el j -ésimo resultado de X_i y se puede interpretar como el «peso de la contribución» de los FPC a la curva X_i .

Un método para elegir el número de componentes principales es cuando la suma de los valores propios supera el 85 % de la suma total. En caso de que la primera componente supere el 85 %, se hace una rotación en los vectores propios (se elige otra base), de modo que haya por lo menos dos componentes principales.

3. Aplicación e interpretación del modelo a los datos del lago

En esta sección aplicamos el modelo descrito en [11] a los datos de precipitaciones del Lago de Valencia.

Las curvas asociadas a los datos

Se cuenta con datos mensuales de precipitaciones en los alrededores del Lago de Valencia medidos en once estaciones meteorológicas (al lado de cada estación se colocó el tiempo durante el cual se tomaron datos de acuerdo a la estación), a saber:

E417: Santa Cruz, 1966 a 1999.
 E444: «Sin información», 1954 a 1966.
 E452: Mariara, 1949 a 1993.
 E466: Maracay, 1934 a 1992.
 E488: El Trompillo, 1960 a 2005.
 E489: Agua Blanca, 1934 a 2005.
 E491: Las Dos Bocas, 1949 a 2005.
 E497: La Ceniza, 1960 a 2003.
 E1483: «Sin información», 1944 a 1994.
 E1494: Embalse Taiguaguay, 1951 a 1999.
 E9317: «Sin información», 1974 a 1993.

Como se puede ver, desconocemos la ubicación de las estaciones 444, 1483 y 9317. En la Figura 1 se puede observar dónde están ubicadas aquellas estaciones de las que se conoce su ubicación. Para aplicar el modelo se le calculó la media mensual a cada estación, dejando una tabla de datos con dos variables: mes y estación. Para reducir la escala se le aplicó logaritmo en base 10 a cada uno de los datos. La tabla resultante se guardó en un archivo.txt, para luego ser cargada en el software estadístico **R**. Luego se cargó el paquete `fda` de **R** para el tratamiento de datos funcionales.

Los datos vienen en forma discreta. Por eso se deben transformar a datos continuos o funciones. Es decir, por cada estación se debe encontrar una función en L^2 que interpole los datos discretos. En la literatura se recomienda usar series de Fourier cuando se trata de datos de precipitaciones, ya que son fenómenos periódicos. Se debe encontrar una serie de modo que cada función aleatoria se pueda escribir como,

$$X_j(t) = \sum_{k=1}^K a_{kj} \mathbf{v}_k(t), \quad (3)$$

donde j indica la estación. Se escogió $K = 13$ (el

número debe ser impar) ya que un número más pequeño daba errores en la estimación y uno más grande no modificaba la estimación. Existen métodos para estimar K , pero no se usaron en este estudio. Dado que es una serie de Fourier, los $\mathbf{v}_k$ vienen dados de la siguiente manera:

$$\mathbf{v}_1 = 1, \mathbf{v}_{2k} = \sin(k\omega t), \mathbf{v}_{2k+1} = \cos(k\omega t),$$

para $k = 1, \dots, K-1$. La constante ω viene dada por la relación $\omega = 2\pi/T$, donde T , en este caso, es 12, por el número de los meses del año, por lo tanto $\omega = \pi/6$. Para obtener las curvas asociadas a los datos se creó una base de Fourier, usando el comando en **R**, `create.fourier.basis`, que requiere tres entradas:

1. `rango`, en nuestro caso 0 a 12, y lo definimos como el vector $(0, 12)$.
2. $K = 13$, el número de vectores de la base.
3. El periodo $T = 12$, que hay que añadirlo, ya que por defecto es 365, por los días del año.

La base creada se muestra en la Figura 2.

Ahora bien, para que cada curva pueda escribirse en términos de la base de Fourier, se deben estimar los coeficientes de (3) utilizando la función `fdPar` del paquete `fda` de **R**. Esta función usa los siguientes parámetros de suavizado para la interpolación.

Operador de aceleración armónica

El suavizado se admite mediante la clase `Lfd` que expresa el concepto de un operador diferencial lineal. Un operador diferencial lineal L es un operador lineal de la forma,

$$L(x(t)) = a_0 x(t) + \sum_{k=1}^m a_k D^k x(t).$$

Un caso especial importante es el «operador de aceleración armónica»,

$$L(X(t)) = \omega^2 DX(t) + D^3 X(t), \quad (4)$$

que se usa ampliamente con funciones de base de Fourier para suavizar datos periódicos. Este operador de aceleración armónica se usa para penalizar los términos más altos en la serie de Fourier (3). A los operadores lineales que se usan para este propósito se les denomina «operadores de penalización por rugosidad».

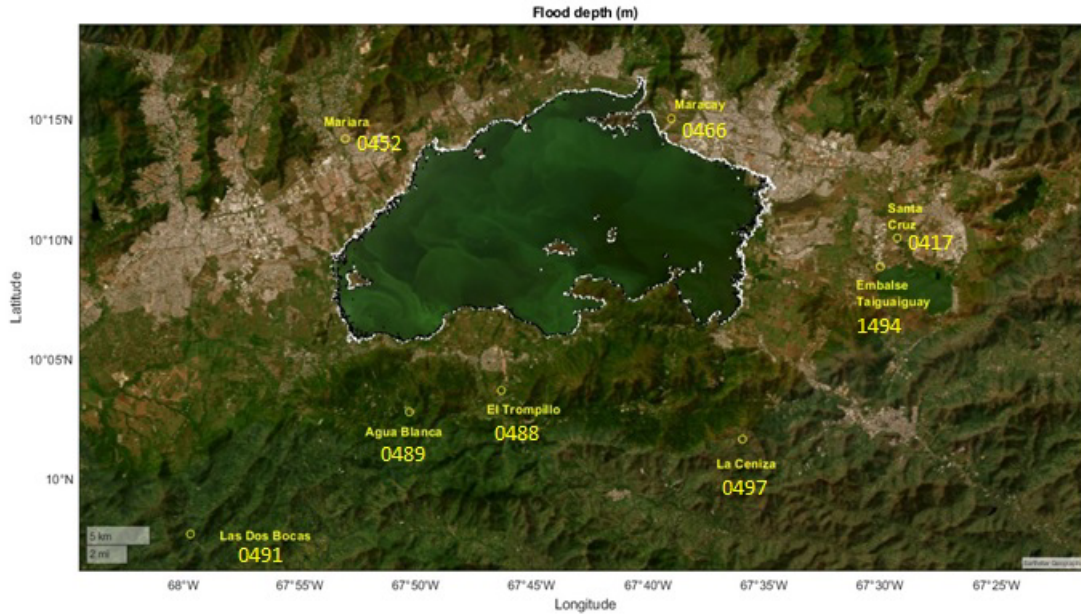


Figura 1. Estaciones conocidas

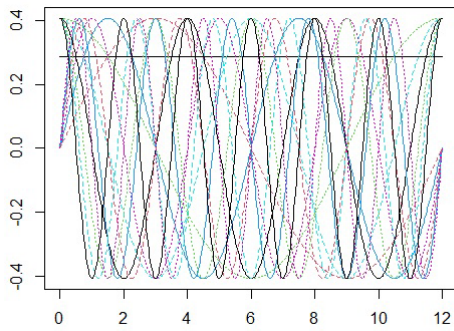


Figura 2. Base de Fourier. Fuente: propia

Matriz de rugosidad y Matriz de suavizado

La ecuación (3) puede escribirse de manera matricial como, $X(t) = \mathbf{c}^T \mathbf{v}(t)$.

Así, se define el «criterio de ajuste penalizado por rugosidad» como,

$$F(\mathbf{c}) = (\mathbf{y} - \mathbf{V}\mathbf{c})^T (\mathbf{y} - \mathbf{V}\mathbf{c}) + \lambda \int [L(X(t))]^2 dt. \quad (5)$$

donde j es la estación, λ es llamado «parámetro de suavizado» y V es la matriz de los datos. La matriz de rugosidad es

$$R = \int [L(X(t))]^2 dt = \int \mathbf{v}(t) \mathbf{v}^T(t) dt,$$

y la matriz de suavizado como,

$$H(\lambda) = V(V^T V + \lambda R)^{-1} V^T.$$

Parámetro de suavizado

La medida de validación cruzada generalizada, GCV, desarrollada por Craven y Wahba (1979), está diseñada para localizar el mejor valor para el parámetro de suavizado λ . Otras medidas se muestran en [7]. Esta medida se define como,

$$GCV(\lambda) = \left(\frac{n}{n - df(\lambda)} \right) \left(\frac{SSE}{n - df(\lambda)} \right) \quad (6)$$

donde $df(\lambda) = \text{Traza}(H(\lambda))$ y

$$SSE = \sum_{j=1}^n [y_j - X(t_j)]^2.$$

El mejor parámetro de suavizado se obtiene minimizando (6). En la Figura 3 se observa que el mínimo de $GCV(\lambda)$ se obtiene cuando $\log(\lambda) = -1$, es decir, $\lambda = 0,1$.

Finalmente, usando la base de Fourier, la penalización y el parámetro de suavizado se utiliza el comando `smooth.basis` para calcular los coeficientes para cada una de las curvas asociadas a los datos de las estaciones. En la Figura 4 se muestran las curvas asociadas a cada una de las estaciones. Para un desarrollo más completo de esta subsección ver [11] y [7].

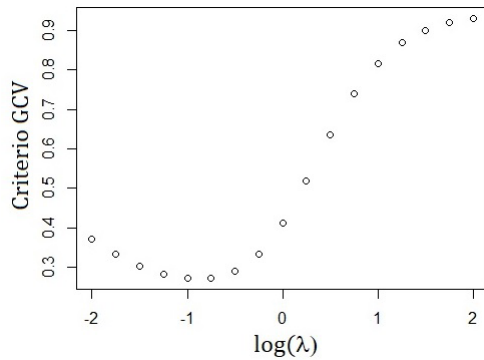


Figura 3. Criterio GCV. Fuente: propia

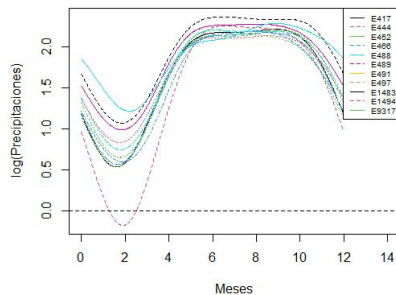


Figura 4. Datos funcionales por estación. Fuente: propia

Componentes principales

El análisis de componentes principales se implementa con la función `fda.fd` en **R**. Esta función tiene los siguientes parámetros de entrada:

```
pca.fd(datos funcionales, nharm =
      3, harmfdPar = fdPar(datos
      funcionales), centerrfns=TRUE)
```

donde «datos funcionales» son los que se generaron con el comando `smooth.basis`, «nharm = 3» indica el número de componentes principales que se desean (esto se modifica una vez obtenidos los datos), «fdPar» es la función de penalidad.

La función `fda.fd` devuelve un objeto funcional con los siguientes componentes:

harmonics: Un objeto funcional que contiene los vectores de las componentes principales.

values: El conjunto completo de valores propios.

scores: Muestra la influencia de cada estación en cada componente principal.

varprop: Un vector con la proporción de la varian-

za explicada por cada vector propio.

meanfd: Un objeto funcional que da la función media de X .

El criterio para elegir el número de componentes principales depende de la magnitud de los valores propios. Esto lo vemos en el vector con la proporción de la varianza explicada por cada vector propio, `varprop`:

(0.88, 0.06, 0.04).

Esto indica que el 88 % de la variabilidad es explicada por la primera componente principal. La sugerencia en la literatura es que haya más de una componente principal. Se aplica una matriz de rotación para hallar una nueva base en la cual al menos dos componentes principales expliquen la variabilidad de los datos. Con el comando `varmx.pca.fd` se crean otras componentes principales a través de un proceso de rotación. El comando arroja, junto con los mismos resultados del comando `pca.fd` la matriz de rotación usada. Igual que con el caso anterior, revisamos el vector de proporción, que en este caso es:

(0.66, 0.28).

Este cubre el 94 % de la variabilidad de los datos, y ninguna de sus componentes cubre más del 85 %, que es lo que buscábamos con la rotación. Los coeficientes de las componentes principales en la base de Fourier (armónicos) vienen dados por,

$$A = \begin{bmatrix} 3.93 \times 10^{-1} & 7.97 \times 10^{-1} \\ 7.68 \times 10^{-1} & -4.23 \times 10^{-1} \\ 3.76 \times 10^{-1} & 2.69 \times 10^{-1} \\ 2.00 \times 10^{-1} & 1.09 \times 10^{-2} \\ -2.65 \times 10^{-1} & 3.30 \times 10^{-1} \\ -1.34 \times 10^{-2} & 4.52 \times 10^{-2} \\ -4.31 \times 10^{-2} & 4.93 \times 10^{-2} \\ -1.83 \times 10^{-2} & 2.43 \times 10^{-2} \\ -2.06 \times 10^{-2} & 1.03 \times 10^{-2} \\ 4.91 \times 10^{-3} & -6.48 \times 10^{-3} \\ 2.40 \times 10^{-6} & 1.31 \times 10^{-7} \\ 5.88 \times 10^{-3} & -5.91 \times 10^{-3} \end{bmatrix} \quad (7)$$

y las componentes principales vienen dadas por,

$$PC = \mathbf{v}^T A,$$

donde $PC = [PC_1(t) \ PC_2(t)]^T$.

La gráfica de las componentes principales viene dada en la Figura 5.

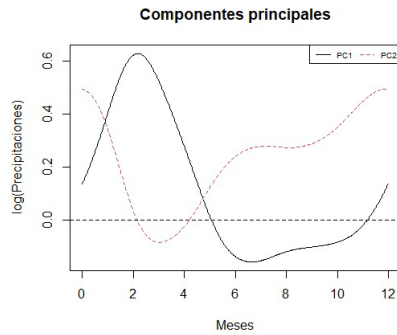


Figura 5. Las componentes principales. Fuente: propia

Los valores propios asociados a las componentes principales son:

$$\lambda_1 = 0.2942269 \quad \text{y} \quad \lambda_2 = 0.1237606.$$

Discusión de resultados

Aquí se analizan los «scores» y el «gráfico de media perturbada».

Los scores

Los «scores», o resultados, son los coeficientes de cada una de las curvas aleatorias asociadas a las estaciones con respecto a la base formada por las componentes principales, quedando cada curva X_j como,

$$\begin{aligned} X_{417}(t) &= -0.1656PC_1(t) - 0.1894PC_2(t), \\ X_{444}(t) &= -1.3683PC_1(t) - 0.4572PC_2(t), \\ X_{452}(t) &= -0.1702PC_1(t) - 0.3378PC_2(t), \\ X_{466}(t) &= -0.2363PC_1(t) - 0.2324PC_2(t), \\ X_{488}(t) &= 0.0791PC_1(t) + 0.0445PC_2(t), \\ X_{489}(t) &= 0.4268PC_1(t) + 0.3139PC_2(t), \\ X_{491}(t) &= 0.5458PC_1(t) + 0.6481PC_2(t), \\ X_{497}(t) &= 0.2446PC_1(t) - 0.0875PC_2(t), \\ X_{1483}(t) &= -0.0453PC_1(t) - 0.0588PC_2(t), \\ X_{1494}(t) &= -0.1382PC_1(t) - 0.2418PC_2(t), \\ X_{9317}(t) &= 0.8275PC_1(t) + 0.5986PC_2(t). \end{aligned}$$

Los «scores» muestran el peso o la influencia que tiene la estación en la componente principal respectiva. Mientras más distante de cero sea el score, más peso tiene esa estación en la componente principal. La Figura 6 muestra los «scores» de las estaciones, cada punto representa una estación, en el eje de las

abscisas aparece la primera componente principal y en el de las ordenadas el segundo. Las coordenadas de cada punto corresponden a los coeficientes de las dos componentes principales respectivamente. Los puntos más alejados del origen representan estaciones con mayor peso, mientras que los más cercanos al origen representan estaciones con menor peso. También se pueden ver las estaciones que están más aisladas en la gráfica como datos atípicos. En este caso, los valores atípicos serían las estaciones 444, 9317, 491 (Las Dos Bocas), 489 (Agua Blanca).

Adicionalmente, en la Figura 6 se observan dos cluster, uno en el centro que al revisar la ubicación de las estaciones en la Figura 1 se observa que justamente estas estaciones son las que están más cercanas al lago y en un terreno más llano. El segundo cluster se observa en la esquina superior a la derecha tres estaciones que de acuerdo a la Figura 6 son justamente las que están ubicadas en la parte alta que es zona de montaña. Note también que las estaciones más elevadas tienen un score positivo respecto a la primera componente principal, mientras que las más bajas tienen un score negativo. Esto se repite en la segunda componente principal, en una escala más pequeña. Las estaciones más elevadas tienden a tener lluvias más abundantes, lo que hace que el peso en cada componente esté relacionado con la cantidad de lluvia. La ubicación de las estaciones de las que se desconoce su ubicación tienen datos interesantes: 444 está más a la izquierda que todas, 9317 está más a la derecha de todas y 1483 está más al centro que todas.

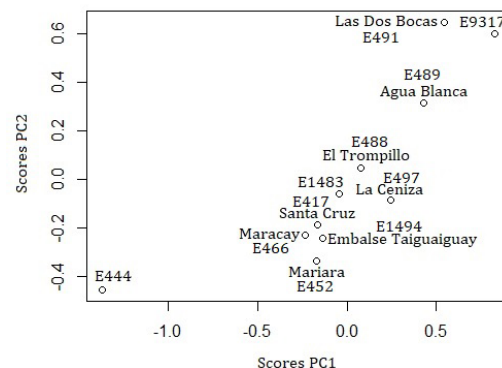


Figura 6. Mapa de scores con las estaciones. Fuente: propia

Media perturbada

El gráfico de «media perturbada», Figura 7, viene dado por tres curvas, $\hat{\mu}(t)$ (ver (1)) es la curva representada con una línea continua, $\hat{\mu}(t) + 2\sqrt{\lambda_i}PC_i$ es la curva representada con símbolos de «+» y $\hat{\mu}(t) - 2\sqrt{\lambda_i}PC_i$ es la curva representada por símbolos de «-». Estas fórmulas se encuentran en [7], pág. 48 a 51. La primera componente principal muestra una variación mayor en la época de sequía, aproximadamente de noviembre a abril, y la segunda componente principal muestra una variación mayor en la época de lluvias, aproximadamente de mayo a octubre. La alta variación de octubre a febrero es posible que se deba a lluvias atemporales en las estaciones 9317 y 491 (Las Dos Bocas), que tienen un microclima de montaña.

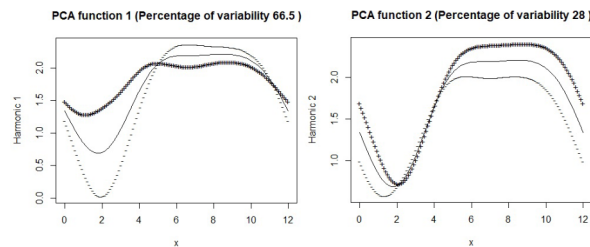


Figura 7. Media perturbada. Fuente: propia

Agradecimientos

Esta investigación ha sido financiada por el FONACIT bajo la subvención 2024PGP142.

Referencias

- [1] Aoki, P. M. (2017). Functional data analysis for rain rate statistics. *Proyect report, CS 229, Stanford, University*.
- [2] Ghumman, A. R., Haider, H., and Shafiqzaman, M. (2020). Functional data analysis of models for predicting temperature and precipitation under climate change scenarios. *Journal of Water and Climate Change*, 11(4):1748–1765.
- [3] Guevara Pérez, E., Guevara Bello, J. E., and García, E. (2008). Régimen climático y patrón espacial de las lluvias en la cuenca del Lago de Valencia, Venezuela. *Revista de Climatología*, 8, 29-40.
- [4] Hael, M. A., Yongsheng, Y., and Saleh, B. I. (2020). Visualization of rainfall data using functional data analysis. *SN Applied Sciences*, 2(3):461.
- [5] Hamdan, M. F., Jemain, A. A., and Jamaludin, S. S. S. (2018). Estimation of rainfall curve by using functional data analysis and ordinary kriging approach. *Matematika*, pages 167–177.
- [6] Hernández, J. and Rodríguez, L.-Á. (2020). Is ENSO associated with precipitation patterns in Lake Valencia Venezuela? *Revista de Climatología*, 20:75–87.
- [7] Hooker, G. (2010). Functional data analysis: A short course. international workshop on statistical modeling.
- [8] Horváth, L. and Kokoszka, P. (2012). *Inference for functional data with applications*, volume 200. Springer Science & Business Media.
- [9] Hsing, T. and Eubank, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*, volume 997. John Wiley & Sons.
- [10] Olivares, B. O. (2018). Condiciones tropicales de la lluvia estacional en la agricultura de secano de Carabobo, Venezuela. *LA GRANJA. Revista de Ciencias de la Vida*, 27(1):86–102.
- [11] Ramsay, J. O., Hooker, G., and Graves, S. (2009). *Functional Data Analysis with R and MATLAB*. Springer Publishing Company, Incorporated, 1st edition.
- [12] Ramsay, J. O. and Silverman, B. O. (2005). *Functional data analysis*. Springer-Verlag.
- [13] Tinoco Mantilla, J. (2015). *Estudio de algunas variables meteorológicas en el Lago de Valencia usando series temporales*. PhD thesis, Universidad Central de Venezuela.
- [14] Wang, J.-L., Chiou, J.-M., and Müller, H.-G. (2016). Functional data analysis. *Annual Review of Statistics and its Application*, 3(1):257–295.