

NOMBRE DEL TRABAJO

**Artículo aprobado ok.pdf**

RECUENTO DE PALABRAS

**4198 Words**

RECUENTO DE PÁGINAS

**14 Pages**

FECHA DE ENTREGA

**Jul 9, 2024 4:13 PM GMT-5**

RECUENTO DE CARACTERES

**21855 Characters**

TAMAÑO DEL ARCHIVO

**489.1KB**

FECHA DEL INFORME

**Jul 9, 2024 4:13 PM GMT-5****● 14% de similitud general**

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos.

- 12% Base de datos de Internet
- Base de datos de Crossref
- 7% Base de datos de trabajos entregados
- 5% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

**● Excluir del Reporte de Similitud**

- Fuentes excluidas manualmente

# 47 Modelo de regresión funcional para la predicción de PM2.5 en función de PM10 en Lima - Perú

Harold A. Choquetico Apaza – Universidad Nacional Intercultural de la Amazonía - ORCID 0000-0003-3889-7266

Edith Silva Rubio – Universidad Peruana Unión - ORCID. 0000-0001-7411-2708

## Abstract

This pioneering study leverages functional data analysis within environmental metrology to examine the intricate relationships between critical environmental variables such as PM2.5 particle concentrations in the Metropolitan Area of Lima-Callao. Addressing the link between air quality and these particles from a functional standpoint has yielded valuable insights that surpass conventional analysis. By examining temporal variations and trends, this research has discerned not only the magnitude of pollution but also its seasonal and daily patterns. The functional data model stands out for its ability to fully utilize historical data, integrating information over time to provide a more comprehensive perspective. This advanced approach has also paved the way for incorporating multiple environmental datasets, such as temperature, humidity, and other pollutants, to offer a broader, more multifaceted view of air quality. 24 The goal of this research is to develop an advanced predictive model that estimates PM2.5 levels based on the concentrations of PM10, utilizing daily air quality records from 2023. Through an integrated framework, the study aims to capture the complex interactions and temporal dynamics of these atmospheric components, highlighting the significant impact of PM2.5 on public health and environmental quality.

## Keywords

Functional Data Analysis, PM2.5, PM10, Air Quality, Concurrent model.

## 1. Introducción

Numerosos estudios, como los realizados por Londoño y Cañón en 2018, han revelado de manera concluyente que los contaminantes atmosféricos, y en particular las partículas finas de menos de 2.5 micras (PM2.5), tienen un impacto significativo en la salud de los habitantes de las áreas metropolitanas. Estas partículas, debido a su pequeño tamaño, pueden penetrar profundamente en el sistema respiratorio, causando desde problemas pulmonares hasta enfermedades cardiovasculares. Estos efectos se observan tanto en términos de un aumento en la mortalidad como en la prevalencia de diversas enfermedades crónicas. 20 El impacto de la contaminación atmosférica en la salud humana ha sido objeto de una investigación intensiva debido a su vínculo con el incremento de la mortalidad y la prevalencia de enfermedades respiratorias y cardiovasculares. Dichos efectos son observables tanto a corto como a largo plazo, con diversas implicaciones y consecuencias en todas las etapas de la vida, desde la concepción hasta la vejez. (Baena, Jiménez, Zapata & Ramírez, 2019) 52 25 12

A medida que las economías urbanas se expanden, con un aumento en la industrialización y la urbanización, la calidad del aire se convierte en una preocupación creciente. Esta preocupación no es infundada, ya que el deterioro ambiental, un subproducto frecuente del desarrollo económico acelerado, puede, si no se aborda de manera efectiva y oportuna, limitar el futuro crecimiento económico. Este fenómeno crea un círculo vicioso donde la contaminación atmosférica deteriora la salud pública, lo que a su vez reduce la productividad laboral y aumenta los costos sanitarios, restringiendo así el potencial económico de la región. (Azqueta y Delacámara, 2004)

Además, la Secretaría de Medio Ambiente y Recursos Naturales, en su informe de 2011, ha identificado una correlación entre la presencia de contaminantes atmosféricos y eventos climáticos extremos, como sequías prolongadas, lluvias intensas y nevadas severas. Estos eventos no solo perturban los ecosistemas naturales y la agricultura, sino que también tienen repercusiones directas en la vida cotidiana y la infraestructura urbana.

En el caso específico del Área Metropolitana de Lima-Callao, el aumento demográfico ha exacerbado la problemática de la contaminación del aire. La densidad poblacional, combinada con un alto número de vehículos y una industrialización creciente, ha llevado a que esta zona sea una de las más contaminadas en términos de partículas suspendidas en el aire en toda América Latina

En el campo de la metrología ambiental, el uso del análisis de datos funcionales aún no es muy extendido, a pesar de su potencial significativo. Este estudio pionero adopta el modelo de regresión funcional para investigar las complejas interrelaciones entre variables medioambientales críticas, como las concentraciones de partículas PM2.5. Al abordar la conexión entre la calidad del aire y estas partículas desde una perspectiva funcional, logramos obtener insights valiosos que van más allá de los análisis convencionales. (Zhu, Zhao, Li, Tan y Dai, 2023).

Este enfoque funcional permite una comprensión más profunda de cómo las variaciones temporales y las tendencias en las concentraciones de PM2.5 afectan la calidad del aire. Por ejemplo, a través del análisis de series temporales y patrones funcionales, es posible discernir no solo la magnitud de la contaminación, sino también sus patrones estacionales y diarios. Esta comprensión detallada es crucial para la formulación de estrategias efectivas de prevención y tratamiento de la contaminación atmosférica. (Sanjana, DeSouza, Machida, Mora, Duarte y Ratti, 2023).

Además, el modelo de datos funcionales se distingue por su capacidad para aprovechar al máximo la información histórica, integrando datos a lo largo del tiempo para proporcionar una perspectiva más completa. A diferencia de los métodos tradicionales que pueden depender en gran medida de juicios subjetivos o de análisis estáticos puntuales, el modelo funcional captura la dinámica y la evolución de los factores ambientales, permitiendo así predicciones y análisis más precisos y confiables.

Este enfoque avanzado también abre la puerta a la integración de múltiples conjuntos de datos ambientales, como la temperatura, la humedad y otros contaminantes, para obtener un

panorama más completo y multidimensional de la calidad del aire. Al correlacionar estas variables ambientales con los niveles de PM2.5, el modelo puede revelar patrones ocultos y relaciones complejas que de otro modo serían difíciles de detectar. (Xiao y Hu, 2018).

En el marco de este estudio, nuestro objetivo es desarrollar un modelo predictivo avanzado que estime los niveles de PM2.5 basándose en las concentraciones de PM10, O3, NO2 y SO2. Para lograr esto, implementamos un enfoque de regresión funcional múltiple que aprovecha los registros diarios de la calidad del aire recopilados en el período comprendido entre 2021 y 2023. La relevancia de modelar PM2.5 radica en su impacto significativo en la salud pública y la calidad ambiental, ya que partículas de este tamaño pueden penetrar profundamente en el sistema respiratorio y causar efectos adversos. Mediante el análisis de estos contaminantes en un marco integrado, buscamos capturar las interacciones complejas y la dinámica temporal de estos componentes atmosféricos.

## 2. Datos de estudio y metodología

### 2.1. Descripción del conjunto de datos

Se emplearon los datos provenientes de la Plataforma de datos históricos acerca de la calidad del aire. El grupo de trabajo encargado del desarrollo del Índice Mundial de la Calidad del Aire (WAQI) y del Sistema de Gestión de Información de la Calidad del Aire en Perú (SENAMHI) suministra información clara sobre la calidad del aire en los distritos de Lima (San Juan de Lurigancho, Puente Piedra, Carabayllo, Santa Anita, San Martín de Porras, Campo de Marte, San Borja, Ate, Villa María El Triunfo y Huachipa). Esta información abarca datos diarios de PM2.5, PM10.

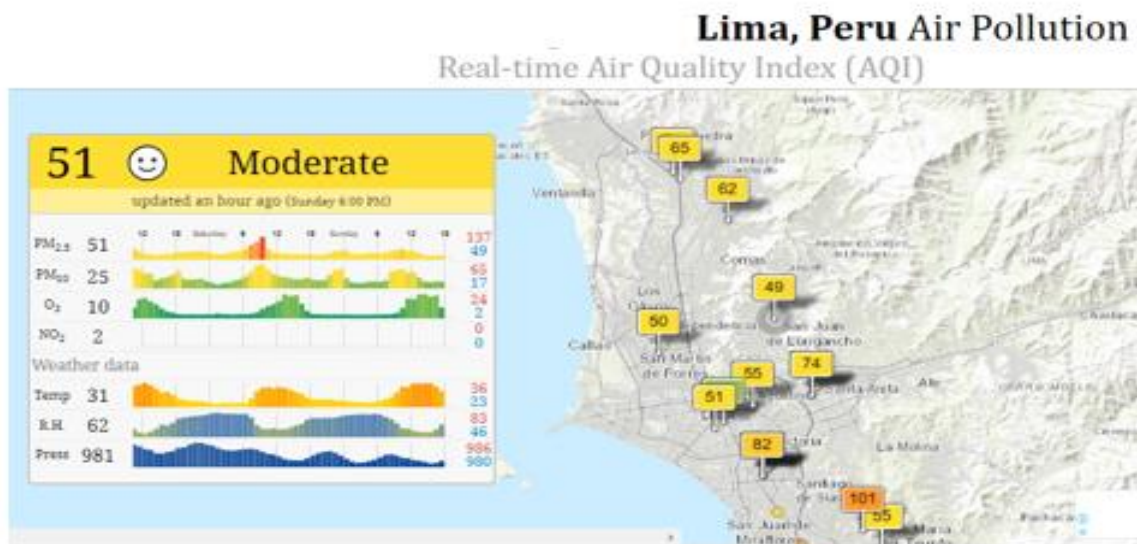
En la fase inicial de nuestra investigación, se realizó un minucioso examen de 348 variables con el fin de determinar su idoneidad para el análisis propuesto. Este procedimiento incluyó un diagnóstico detallado para identificar las variables que cumplían con los criterios de selección establecidos. Una parte crucial de este proceso fue la evaluación exhaustiva de la integridad de los datos, llevada a cabo durante la etapa de depuración. En consecuencia, se decidió centrar el estudio en un conjunto de 6 compuestos por 5 variables explicativas denominadas 'X' y una variable dependiente etiquetada como 'Y', utilizada para representar la variable de interés, que en este caso es la 'CALIDAD DEL AIRE'.

La población son todas las concentraciones de PM2.5 obtenidas de las estaciones fijas de monitoreo, así como el índice normalizado los valores obtenidos del PM10. La Concentración de PM2.5 (media en 24 h) obtenida de las estaciones fijas de monitoreo del SENAMHI durante el año 2023. Los datos provienen de siete estaciones fijas que se muestran en la tabla 1.

**Tabla 1.** Ubicación geográfica de estaciones fijas de monitoreo de calidad del aire en Lima Metropolitana

Código	Distrito	Latitud	Longitud
E1	San Juan de Lurigancho	12° 01' "S"	76° 59' "W"
E2	Puente Piedra	11° 55' "S"	77° 05' "W"
E3	Carabaylo	11° 54' "S"	77° 02' "W"
E4	Santa Anita	12° 02' "S"	76° 58' "W"
E5	San Martín de Porras	11° 58' "S"	77° 04' "W"
E6	Campo de Marte	12° 04' "S"	77° 02' "W"
E7	San Borja	12° 06' "S"	77° 00' "W"
E8	Ate	12° 01' "S"	76° 55' "W"
E9	Villa María El Triunfo	12° 13' "S"	76° 56' "W"
E10	Huachipa	12° 01' "S"	76° 56' "W"

**Figura 1.** Estaciones de medición del Índice de calidad del aire (ICA) en tiempo real.



## 2.2. Análisis de Datos Funcional (ADF)

A continuación, se presentan los conceptos preliminares del ADF. Ésta es una rama de la estadística, que estudia y analiza la información contenida en curvas, superficies, o cualquier elemento que varía sobre un continuo, generalmente el tiempo (Ramsay & Silverman, 2005). Las particularidades de estos objetos, las funciones, hacen que surja todo un aparato de definiciones, teoremas y herramientas para poder estudiarlas correctamente (Manteiga & Vieu, 2007). Por ello, el primer paso será definir y comentar los conceptos básicos, las características y peculiaridades de este tipo de datos.

La explosión tecnológica ha hecho del ADF una disciplina emergente dentro de la estadística con multitud de publicaciones y estudios. Prueba de este interés es la creciente presencia en revistas especializadas y la creación de un grupo de trabajo específico dentro del ERCIM (European Research Consortium for Informatics and Mathematics). Entre la literatura disponible sobresalen como referencias básicas los libros de (Ramsay & Silverman, 2005) y (Ferraty & Vieu, 2006). En el primer caso, tratan muchos de los problemas básicos de la estadística funcional, en el segundo caso abordan la metodología desde el punto de vista no paramétrico. Ramsay y Silverman también tienen un segundo libro de carácter más aplicado en el que se estudian soluciones a problemas sobre conjuntos de datos concretos (Ramsay & Silverman, 2002).

En la práctica, los datos funcionales son obtenidos de forma discreta, siendo necesario convertirlos a funciones sobre las cuales se pueda trabajar, para esto pueden emplearse diferentes tipos de ajustes. Teniendo en cuenta lo anterior, la representación de estos datos como datos funcionales a partir de una base ortonormal brinda una ventaja tanto teórica como práctica, esta base se utiliza para transformar datos discretos en datos funcionales.

### 2.3. Construcción de datos funcionales

En Análisis de Datos Funcionales (ADF) los datos en lugar de ser un conjunto de vectores, como en análisis multivariado clásico, son un conjunto de curvas. En la mayoría de las aplicaciones, las curvas observadas proceden de la observación de un proceso estocástico en tiempo continuo (Aguilera, 2009).

El dato funcional se define a partir de una base de funciones  $\{\phi_k\}_{k \in N}$ , tal que cualquier función puede representarse mediante una combinación lineal de ella, es decir  $x(t) = \sum_{k=1}^K C_k \phi_k(t)$ . La expresión de forma matricial queda expresada como,

$$X(t) = C^T \Phi = \Phi^T C \quad (2.1)$$

donde los  $C_k$  son los coeficientes en la nueva base,  $\Phi$  es una matriz que contiene los valores de  $\phi_k(t)$  y  $K$  será el número de bases. En este contexto es importante en la selección del valor de  $K$ , dado que la expresión de la función seguirá siendo infinita o demasiado grande, por lo que habrá que delimitar la serie de datos. De aquí la importancia de la elección del número de bases  $K$ , este se comporta como otro parámetro del modelo midiendo el grado de suavización de la función (Ramsay & Silverman, 2005).

De la expresión 2.1 se denota que el problema consiste en estimar el vector de coeficientes  $C$ , este vector pondera cada función de la base sobre la función de interés  $x(t)$ , (Ramsay & Silverman, 2005) estimaron el vector  $C$  a partir de mínimos cuadrados ordinarios. A continuación, se presenta el proceso de estimación propuesto:

Sea  $y_i$  el valor observado de la variable  $Y$  en el tiempo  $t_j \forall j = 1, \dots, n$ , tal que puede ser expresados como:

$$y_i = x(t_j) + \epsilon_j = \phi^T(t_j)C + \epsilon_j \quad (2.2)$$

Donde  $x(t_j)$  es el valor de la función ajustada con (2.1) en el tiempo  $t_j$  y  $\epsilon_j$  es el error de estimación en el momento  $t_j$ . Suponiendo que los errores se distribuyen normal, tienen media cero, varianza constante e independientes. La expresión por mínimos cuadrados ordinarios a minimizar es [Ramsay & Silverman, 2005]:

$$SSE_{(C)} = \sum_j^n [y_j - \sum_k^K C_k \phi_k(t_j)]^2 = \sum_j^n [y_j - \Phi^T(t_j)C]^2 \quad (2.3)$$

Sea  $\mathbf{Y}$  un vector conformado por todos los valores ajustados  $y_j$ ,  $\Phi^T$  una matriz de dimensión  $n \times K$  compuesta por los valores de las  $K$  funciones base  $\phi^T(t_j)$  y  $\epsilon$  el vector de residuales de tamaño  $n \times 1$ , por lo tanto, expresión (2.2) es expresada de forma matricial como (Ramsay & Silverman, 2005):

$$Y = \Phi^T C + \epsilon \quad (2.4)$$

de esta manera, la expresión (2.3) se puede expresar como:

$$SSE(c) = (Y - \Phi^T C)^T (Y - \Phi^T C) \quad (2.5)$$

resolviendo el sistema para estimar  $C$ , se obtiene el estimador  $\hat{C}$  que minimiza la suma de cuadrados de (2.5), y está dado por

$$\hat{C} = (\Phi \Phi^T)^{-1} \Phi Y \quad (2.6)$$

Finalmente, los valores ajustados son expresados como:

$$\hat{Y} = \Phi^T (\Phi \Phi^T)^{-1} \Phi Y \quad (2.7)$$

### 3. Resultados

#### 3.1. Análisis exploratorio de datos

Para el estudio de la calidad del aire en el año 2023, se llevó a cabo un análisis exhaustivo utilizando registros diarios de contaminantes atmosféricos. Durante la fase preliminar del análisis, se identificaron y abordaron los datos faltantes, a través de un proceso de imputación múltiple, lo que permitió la reconstrucción de un conjunto de datos íntegro y fiable para su

posterior evaluación. A continuación, se presenta el análisis descriptivos de los contaminantes atmosféricos por sitio.

**Tabla 1.** Resúmenes descriptivos de los contaminantes analizados, en el año 2023.

Contaminante/Sitio	Promedio	Mediana	Mínimo	Máximo	Desv. Estándar	Coef. de variación (%)
<b>PM2.5</b>						
Campo_de_Marte	61.3	61.0	1.0	118.0	17.7	29.0
Carabayllo	84.4	83.0	47.0	130.0	14.6	17.0
Puente_Piedra	61.6	60.0	1.0	118.0	18.6	30.0
San_Borja	64.5	62.0	13.0	110.0	11.8	18.0
San_Juan_Lurigancho	79.4	79.0	6.0	126.0	19.0	24.0
San_Martin_de_Porras	64.1	62.0	1.0	130.0	19.5	30.0
Santa_Anita	71.4	70.0	1.0	119.0	15.9	22.0
Villa_Maria_Triunfo	71.8	70.0	33.0	122.0	16.6	23.0
<b>PM10</b>						
Campo_de_Marte	28.6	27.0	1.0	84.0	13.4	47.0
Carabayllo	39.3	37.0	15.0	87.0	14.1	36.0
Puente_Piedra	28.9	28.0	1.0	84.0	13.8	48.0
San_Borja	44.3	45.0	1.0	84.0	11.2	25.0
San_Juan_Lurigancho	39.5	41.0	1.0	85.0	14.4	36.0
San_Martin_de_Porras	27.7	26.0	1.0	68.0	11.8	43.0
Santa_Anita	31.5	31.0	9.0	62.0	8.1	26.0
Villa_Maria_Triunfo	35.1	33.0	10.0	89.0	13.6	39.0

31 La **Tabla 1** muestra el resumen de las estadísticas descriptivas de la calidad del aire para varias zonas de Perú en el año 2023, específicamente los niveles de PM2.5, PM10, O3, NO2 y SO2. Los datos reflejan la variabilidad y concentración de estos contaminantes atmosféricos, los cuales tienen implicaciones directas en la salud pública y el medio ambiente.

45 Para PM2.5, conocido por su capacidad de penetrar profundamente en los pulmones y pasar al torrente sanguíneo, los valores promedio varían entre 61.3  $\mu\text{g}/\text{m}^3$  en Campo de Marte y 79.4  $\mu\text{g}/\text{m}^3$  en San Juan, Lurigancho, indicando que esta última zona podría estar experimentando niveles más preocupantes de partículas finas. La mediana, cercana al promedio en todos los sitios, sugiere una distribución relativamente simétrica de los datos. Sin embargo, el amplio rango (mínimo y máximo) y el coeficiente de variación destacan la

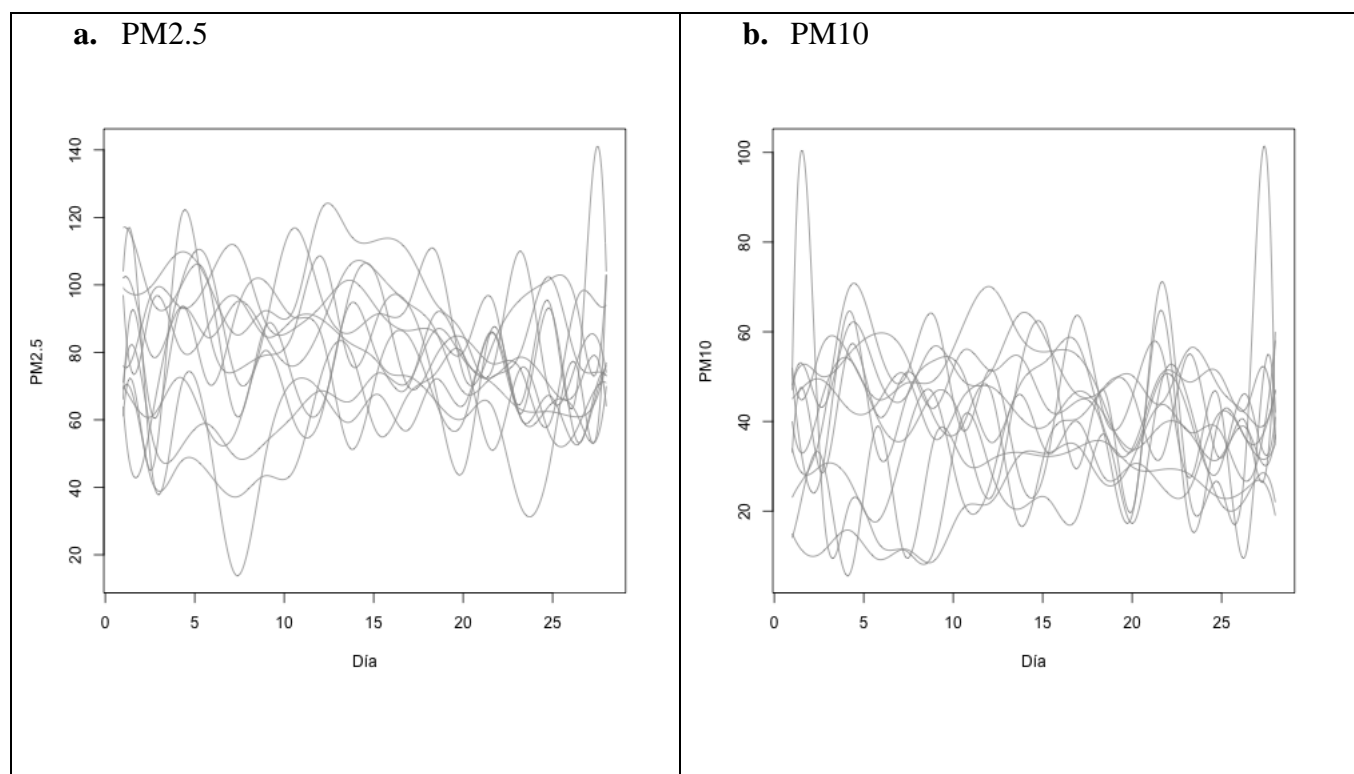
presencia de días con alta contaminación, especialmente en Villa María, Triunfo, donde el coeficiente de variación alcanza el 33%.

Para PM10, las concentraciones promedio oscilan entre 28.6 y 44.3  $\mu\text{g}/\text{m}^3$ , siendo San Juan, Lurigancho la zona con la concentración promedio más alta. El rango de estas partículas es menos variable que el de PM2.5, aunque el sitio de Villa María, Triunfo muestra nuevamente una considerable fluctuación con un coeficiente de variación del 39%.

### 3.2. Construcción de datos funcional

Para la construcción de los datos funcional se utilizaron los datos del día 1 hasta el día 28 para que todos los meses tuvieran la misma cantidad de datos diarios. Se utilizaron bases B-Spline y validación cruzada generalizada para seleccionar el número óptimo de bases.

**Figura 2.** Datos funcionales de contaminantes atmosféricos.



La **Figura 2** muestra los datos funcionales para los contaminantes PM2.5 y PM10, reflejados a través de curvas que representan las concentraciones diarias a lo largo del año 2023. Cada curva individual puede corresponder a un mes.

En el gráfico de PM2.5 (Subfigura a), las curvas muestran una considerable fluctuación, con algunos días presentando picos significativamente altos que podrían estar vinculados a eventos específicos que influyen en la calidad del aire, como condiciones climáticas adversas, incendios o incrementos en las emisiones locales. La variabilidad diaria es notable y sugiere una influencia de factores dinámicos sobre las concentraciones de PM2.5.

Por su parte, el gráfico de PM10 (Subfigura b) también exhibe variabilidad, pero parece tener menos fluctuaciones extremas en comparación con PM2.5. A pesar de esto, hay días en los que las concentraciones alcanzan valores máximos similares a los de PM2.5, lo que indica episodios de alta contaminación que afectan tanto a partículas finas como más gruesas.

Ambos gráficos ilustran que los niveles de contaminantes no son constantes y refuerzan la importancia de considerar la variabilidad temporal al evaluar la exposición a la contaminación atmosférica. La interpretación de estos datos funcionales es crucial para entender el comportamiento de los contaminantes en el aire y puede informar las estrategias de mitigación y políticas de salud pública para abordar la calidad del aire en entornos urbanos y reducir los efectos adversos de la contaminación en la población.

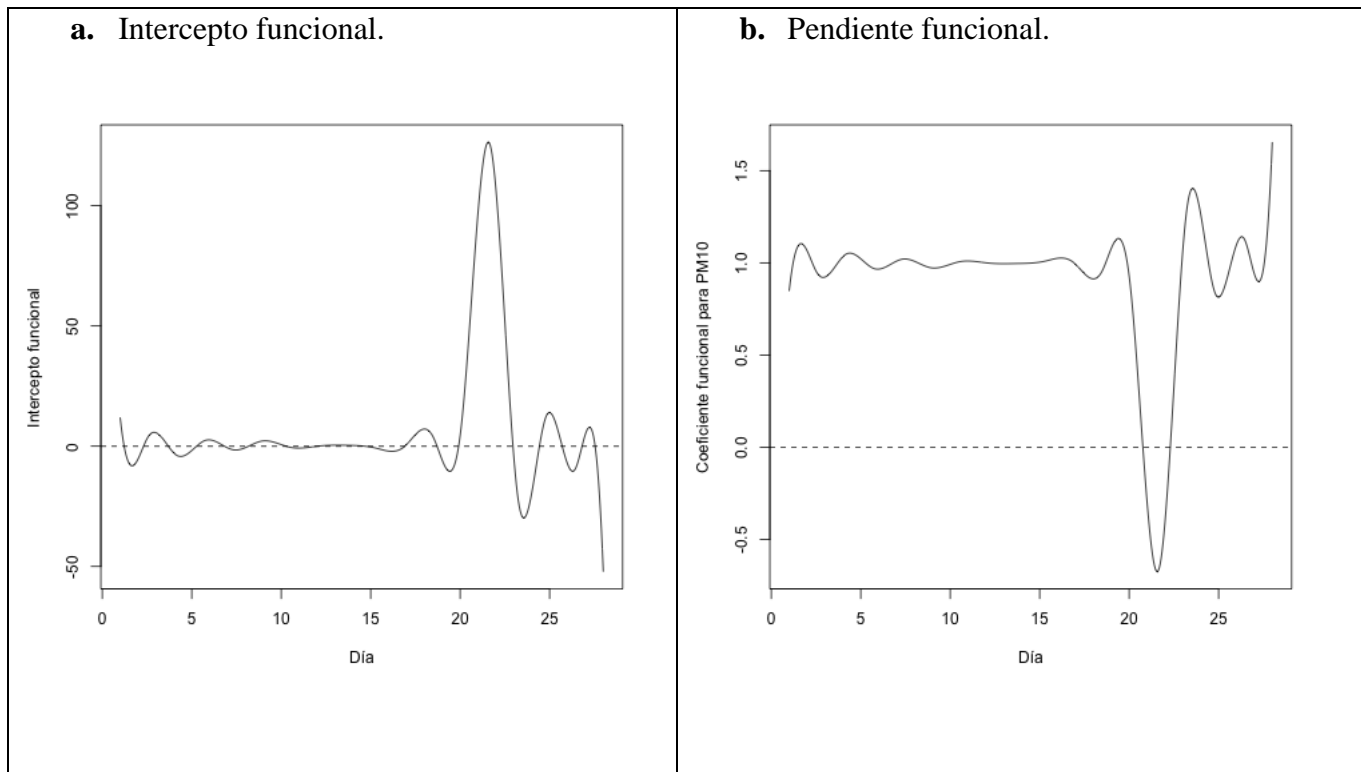
### 3.3 Modelo de regresión funcional

Para evaluar el aporte del PM10 en el ajuste del PM2.5. Se ajustó un modelo de regresión funcional concurrente. A continuación, se expresa el modelo previamente descrito:

$$PM2.5_i(t) = \beta_0(t) + \beta_1(t)PM10_i(t) + \epsilon(t)_i$$

donde,  $PM2.5_i(t)$  es la variable respuesta funcional que corresponde a los datos

funcionales de PM2.5,  $PM10_i(t)$  es la variable explicativa funcional que corresponde a los datos funcionales PM10  $\epsilon(t)_i$  es la componente aleatoria del modelo correspondiente al error funcional,  $i = 1, 2, \dots, 12$  correspondiente a 12 meses y  $t$  representa el tiempo en días que está en un rango 28 días.



El análisis presentado en la Figura 2 se centra en un modelo de regresión funcional concurrente que tiene como objetivo comprender la relación entre las concentraciones diarias de PM2.5 y PM10 durante un período de tiempo específico. El componente del intercepto funcional ilustra la línea de base de PM2.5 en ausencia de PM10, revelando fluctuaciones significativas en las concentraciones de PM2.5 que no son explicadas por PM10. Estos cambios en el intercepto podrían atribuirse a influencias externas o variaciones en fuentes de contaminación distintas de las partículas PM10.

La presencia de picos y valles prominentes en el intercepto sugiere días específicos con niveles inusualmente altos o bajos de PM2.5, lo que resalta la influencia de factores como las actividades industriales, condiciones meteorológicas o episodios de incendios forestales que pueden contribuir a las concentraciones de PM2.5 independientemente de las de PM10.

Por otro lado, la pendiente funcional ofrece una visión dinámica de cómo las concentraciones de PM10 están asociadas con las de PM2.5 a lo largo del tiempo. La variabilidad de la pendiente a lo largo del mes refleja la naturaleza compleja de esta relación, mostrando días en los que la correlación entre PM10 y PM2.5 es más fuerte, ya sea positiva o negativamente. Esto implica que la influencia de PM10 en PM2.5 no es constante, sino que puede ser modulada por factores como la dispersión atmosférica, las reacciones químicas en la atmósfera o la contribución de diferentes tipos de fuentes de contaminación.

La interpretación detallada de la pendiente funcional es esencial para modelar con precisión la calidad del aire, ya que permite a los investigadores y responsables de políticas públicas identificar patrones y establecer estrategias más efectivas para la gestión de la contaminación atmosférica y la protección de la salud pública.

#### 4. Conclusiones

La incorporación del modelo de regresión funcional concurrente en este estudio ha sido instrumental para dilucidar la relación compleja y evolutiva entre los contaminantes PM2.5 y PM10. Este modelo ha destacado la existencia de fluctuaciones en las concentraciones de PM2.5 que no son directamente atribuibles a variaciones en PM10, lo cual sugiere que existen factores independientes y potencialmente ocultos que afectan los niveles de PM2.5. De manera crucial, el intercepto funcional ha revelado una serie de fluctuaciones que parecen ser autónomas de las variaciones de PM10, lo que podría indicar la presencia de fuentes de emisión adicionales o cambios en las condiciones ambientales que afectan específicamente a las partículas más finas.

Por otro lado, la pendiente funcional ha proporcionado una descripción detallada de cómo las concentraciones de PM10 ejercen una influencia sobre los niveles de PM2.5, con una relación que varía a lo largo del tiempo y que es sensible a una gama de factores ambientales dinámicos, incluyendo la dispersión atmosférica y las complejas reacciones químicas que pueden ocurrir en la atmósfera urbana.

Este análisis avanzado, al desglosar las interacciones entre estos dos contaminantes clave, no solo aporta una comprensión más profunda de la dinámica de la contaminación del aire, sino que también se convierte en un recurso valioso para los responsables de la toma de decisiones y los profesionales de la salud pública. Con estos resultados, se pueden identificar patrones críticos que informan las estrategias de intervención y mitigación, apoyando esfuerzos para mejorar la calidad del aire y proteger la salud pública.

La investigación ha evidenciado que los modelos funcionales son esenciales en la cartografía y análisis de la contaminación atmosférica, permitiendo predecir y responder a los desafíos ambientales con mayor eficacia. Al establecer una base sólida para la comprensión de las complejidades del aire urbano, este estudio abre un camino hacia futuras investigaciones que podrían explorar aún más los factores que influyen en la calidad del aire y cómo estos se interrelacionan dentro del entorno urbano, lo que es de suma importancia en la lucha continua contra la contaminación atmosférica y sus impactos en la salud global.

#### Referencias

Aguilera, M. d. C. (2009). Estimación penalizada de datos funcionales. Tesis de Maestría - Universidad de Granada.

- Azqueta, D. & Delacámara, G. (2004). El papel de las variables ambientales en la nueva geografía económica. *Investigaciones Regionales*, núm. 4, primavera, 2004, pp. 145-175. Asociación Española de Ciencia Regional. Madrid, España. <https://www.redalyc.org/pdf/289/28900408.pdf>
- Baena, D.; Jiménez, J.; Zapata, C. & Ramírez, A. (2019) Red neuronal artificial aplicado para el pronóstico de eventos críticos de PM2.5 en el Valle de Aburrá. (Laboratorio de Calidad del Aire - CALAIRE, Facultad de Minas, Universidad Nacional de Colombia Sede Medellín). <http://www.scielo.org.co/pdf/dyna/v86n209/0012-7353-dyna-86-209-347.pdf>
- Digesa. (2011). Estudio de Saturación Lima Metropolitana y Callao, Año 2011. Lima, Perú. <https://www.yumpu.com/es/document/view/13237199/ii-estudio-de-saturacion-dela-calidad-del-aire-en-lima->
- Ferraty, F. & Vieu, P. (2006). *Nonparametric functional data analysis: theory and practice*. New York, USA: Springer. 260p
- García, C. (2015). Concentración de Material Particulado PM en Función de la Humedad y 2,5 Reflectancia Atmosférica Usando Imágenes Landsat-8 en Lima Metropolitana, 2015 – 2016
- Manteiga, W. G. & Vieu, P. (2007). *Statistics for functional data. Computational Statistics & Data Analysis*, 51(10), 4788–4792.
- Ramsay, J. & Silverman, B. (2005). *Functional data analysis*. New York, USA: Springer. 428p.
- Ramsay, J. O. & Silverman, B. W. (2002). *Applied functional data analysis: methods and case studies*, volume 77. New York, USA: Springer. 191p.
- Sanjana Paul, Priyanka deSouza, Yuki Machida, Simone Mora\*, Fábio Duarte y Carlo Ratti (2023). Key themes, trends, and drivers of mobile ambient air quality monitoring: a systematic review and meta-analysis. *environmental science & technology*, 57(26), 9427-9444. <https://doi.org/10.1021/acs.est.2c06310>
- Xiao, W y Hu, Yan (2018). *Functional data analysis of air pollution in six major cities*. College of Sciences, North China University of Technology, Beijing 100144, China
- Zhu, Y.; Zhao, F.; Li, J., Tan, T. y Dai, L. (2023). The quantitative research on atmospheric environmental corrosion of aluminum alloy products. *e3s web of conferences*, 575, 01007. <https://doi.org/10.1051/e3sconf/202337501007>

## EVIDENCIAS DE SUMISIÓN

Revista Científica: BIOTECH AND ENGINEERING Tareas 0 Español (España) Ver el sitio harold\_choquetico

OJS  
OPEN JOURNAL SYSTEMS

Enviar un artículo

1. Inicio 2. Cargar el envío 3. Introducir los metadatos 4. Confirmación 5. Sigüientes pasos

Archivos de envío [Q Buscar](#) [Subir archivo](#)

591-1	harold_choquetico, Artículo aprobado.docx	febrero 28, 2024	Texto del artículo
-------	---	------------------	--------------------

[Guardar y continuar](#) [Cancelar](#)

Platform & workflow by OJS / PKP

https://revistas.untel.edu.pe/index.php/files/submissions

Windows taskbar: 25°C, 09:24, 28/02/2024

Revista Científica: BIOTECH AND ENGINEERING Tareas 0 Español (España) Ver el sitio harold\_choquetico

OJS  
OPEN JOURNAL SYSTEMS

Enviar un artículo

1. Inicio 2. Cargar el envío 3. Introducir los metadatos 4. Confirmación 5. Sigüientes pasos

Se ha subido su envío y ya está listo para ser enviado. Puede volver atrás para revisar y ajustar la información que desee antes de continuar. Cuando esté listo haga clic en "Finalizar envío".

[Finalizar envío](#) [Cancelar](#)

Platform & workflow by OJS / PKP

Windows taskbar: 25°C, 09:30, 28/02/2024

## Envío completo

Gracias por su interés por publicar con Revista Científica: BIOTECH AND ENGINEERING .

### ¿Y ahora qué?

La revista ha sido notificada acerca de su envío y se le enviará un correo electrónico de confirmación para sus registros. Cuando el editor haya revisado el envío, se contactará con usted.

Por ahora, usted puede:

- [Revisar este envío](#)
- [Crear un nuevo envío](#)
- [Volver al escritorio](#)

## ● 14% de similitud general

Principales fuentes encontradas en las siguientes bases de datos:

- 12% Base de datos de Internet
- Base de datos de Crossref
- 7% Base de datos de trabajos entregados
- 5% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

### FUENTES PRINCIPALES

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	<b>univ-sba.dz</b> Internet	<1%
2	<b>e3s-conferences.org</b> Internet	<1%
3	<b>r-users.gal</b> Internet	<1%
4	<b>scielo.br</b> Internet	<1%
5	<b>Leibnitz Rojas-Bustamante, Crayla Alfaro, Ivan Molero, Dennis Aparicio...</b> Crossref	<1%
6	<b>Santiago Santa, Diego Velez, Gustavo Patino. "Analysis and prediction ...</b> Crossref	<1%
7	<b>aprenderly.com</b> Internet	<1%
8	<b>repositorioacademico.upc.edu.pe</b> Internet	<1%

9	<b>uvadoc.uva.es</b> Internet	<1%
10	<b>Markham College on 2020-11-14</b> Submitted works	<1%
11	<b>repositorio.unap.edu.pe</b> Internet	<1%
12	<b>lahey.org</b> Internet	<1%
13	<b>mobilisyourcity.net</b> Internet	<1%
14	<b>theibfr.com</b> Internet	<1%
15	<b>Cuadros Falcon, Zaida Nieves Rodriguez Riojas, Ivonne Miriam. "Plane...</b> Publication	<1%
16	<b>Universidad Continental on 2017-06-30</b> Submitted works	<1%
17	<b>decide.madrid.es</b> Internet	<1%
18	<b>descubridor.umad.edu.mx</b> Internet	<1%
19	<b>pdfslide.net</b> Internet	<1%
20	<b>china-files.com</b> Internet	<1%

21	<b><a href="http://garrigues.com">garrigues.com</a></b>	Internet	<1%
22	<b><a href="http://inaut.unsj.edu.ar">inaut.unsj.edu.ar</a></b>	Internet	<1%
23	<b>London School of Hygiene and Tropical Medicine on 2007-10-11</b>	Submitted works	<1%
24	<b><a href="http://cris.iucc.ac.il">cris.iucc.ac.il</a></b>	Internet	<1%
25	<b><a href="http://entrenos.eafit.edu.co">entrenos.eafit.edu.co</a></b>	Internet	<1%
26	<b><a href="http://prezi.com">prezi.com</a></b>	Internet	<1%
27	<b><a href="http://technodocbox.com">technodocbox.com</a></b>	Internet	<1%
28	<b><a href="http://agendalocal21huelva.com">agendalocal21huelva.com</a></b>	Internet	<1%
29	<b><a href="http://amazon.com">amazon.com</a></b>	Internet	<1%
30	<b><a href="http://avizora.com">avizora.com</a></b>	Internet	<1%
31	<b><a href="http://ecap.uab.es">ecap.uab.es</a></b>	Internet	<1%
32	<b><a href="http://mindmeister.com">mindmeister.com</a></b>	Internet	<1%

33	<b>progeophys.cn</b> Internet	<1%
34	<b>worldbank.org</b> Internet	<1%
35	<b>ELISA ADELL SALES. "Material particulado y bioaerosoles en el aire de ...</b> Crossref posted content	<1%
36	<b>Universidad TecMilenio on 2024-01-22</b> Submitted works	<1%
37	<b>Yuqin Zhu, Fangchao Zhao, Jiameng Li, Tiantian Tan, Lu Dai. "The Qua...</b> Crossref	<1%
38	<b>breatheclean.ca</b> Internet	<1%
39	<b>ddd.uab.cat</b> Internet	<1%
40	<b>de.slideshare.net</b> Internet	<1%
41	<b>webs.ucm.es</b> Internet	<1%
42	<b>worldwidescience.org</b> Internet	<1%
43	<b>prnewswire.com</b> Internet	<1%
44	<b>slideshare.net</b> Internet	<1%

45	<b>transportenvironment.org</b> Internet	<1%
46	<b>"Humanitarian Logistics from the Disaster Risk Reduction Perspective"...</b> Crossref	<1%
47	<b>Miguel Moya, Joaquín Alcañiz-Colomer. " The different effects of inco...</b> Crossref	<1%
48	<b>Zulema Conto Quispe. "La movilidad urbana en Lima. Análisis desde la...</b> Crossref posted content	<1%
49	<b>agustinianosvii.blogspot.com</b> Internet	<1%
50	<b>cofepris.salud.gob.mx</b> Internet	<1%
51	<b>community.esri.com</b> Internet	<1%
52	<b>ideas.repec.org</b> Internet	<1%
53	<b>issuu.com</b> Internet	<1%
54	<b>repositorioinstitucional.uabc.mx</b> Internet	<1%
55	<b>cepis.ops-oms.org</b> Internet	<1%
56	<b>kienyke.com</b> Internet	<1%

57	<b>rug.nl</b> Internet	<1%
58	<b>Pontificia Universidad Catolica del Peru on 2017-07-08</b> Submitted works	<1%
59	<b>Universidad Cientifica del Sur on 2018-07-24</b> Submitted works	<1%
60	<b>slidehtml5.com</b> Internet	<1%

**● Excluir del Reporte de Similitud**

- Fuentes excluidas manualmente

## FUENTES EXCLUIDAS

<b>hdl.handle.net</b>	<b>18%</b>
Internet	
<b>cimat.repositorioinstitucional.mx</b>	<b>6%</b>
Internet	
<b>docplayer.es</b>	<b>6%</b>
Internet	
<b>eps.uam.es</b>	<b>6%</b>
Internet	
<b>repositorio.uam.es</b>	<b>6%</b>
Internet	
<b>probayestadistica.cimat.mx</b>	<b>6%</b>
Internet	
<b>1library.co</b>	<b>5%</b>
Internet	
<b>repositorio.unal.edu.co</b>	<b>4%</b>
Internet	
<b>revistas.unj.edu.pe</b>	<b>3%</b>
Internet	
<b>gredos.usal.es</b>	<b>3%</b>
Internet	

<b>Pontificia Universidad Catolica del Peru on 2021-03-03</b>	<b>3%</b>
Submitted works	
<b>Pontificia Universidad Catolica del Peru on 2020-07-29</b>	<b>3%</b>
Submitted works	
<b>coursehero.com</b>	<b>2%</b>
Internet	
<b>bdigital.unal.edu.co</b>	<b>2%</b>
Internet	
<b>ouci.dntb.gov.ua</b>	<b>2%</b>
Internet	
<b>researchgate.net</b>	<b>2%</b>
Internet	
<b>dokumen.pub</b>	<b>2%</b>
Internet	
<b>Universidad Cesar Vallejo on 2023-12-27</b>	<b>2%</b>
Submitted works	
<b>repositorio.unfv.edu.pe</b>	<b>2%</b>
Internet	
<b>Universidad Peruana de Ciencias Aplicadas on 2024-07-04</b>	<b>2%</b>
Submitted works	
<b>Universidad Nacional de Colombia on 2024-01-30</b>	<b>2%</b>
Submitted works	
<b>doaj.org</b>	<b>2%</b>
Internet	

<b>metropol.gov.co</b>	2%
Internet	
<b>scielo.org.co</b>	1%
Internet	
<b>123dok.net</b>	1%
Internet	
<b>tandfonline.com</b>	1%
Internet	
<b>aqicn.org</b>	1%
Internet	
<b>Universidad Nacional de Colombia on 2019-01-15</b>	1%
Submitted works	
<b>repositorio.uigv.edu.pe</b>	1%
Internet	
<b>core.ac.uk</b>	1%
Internet	
<b>Universidad Nacional de Colombia on 2017-08-11</b>	1%
Submitted works	
<b>thesesups.ups-tlse.fr</b>	1%
Internet	
<b>rdoc.univ-sba.dz</b>	1%
Internet	
<b>Universidad Cesar Vallejo on 2023-07-17</b>	1%
Submitted works	

<b>Universidad Peruana Cayetano Heredia on 2019-12-03</b>	1%
Submitted works	
<b>Ahmed Ait Saidi, Mecheri Kheira. "The conditional cumulative distribution fun...</b>	1%
Crossref	
<b>export.arxiv.org</b>	1%
Internet	
<b>link.springer.com</b>	1%
Internet	
<b>docksci.com</b>	1%
Internet	
<b>Universidad Europea de Madrid on 2023-10-25</b>	1%
Submitted works	
<b>Universidad Europea de Madrid on 2023-10-20</b>	1%
Submitted works	
<b>Universidad Europea de Madrid on 2023-10-20</b>	1%
Submitted works	
<b>Universidad Catolica De Cuenca on 2023-11-15</b>	1%
Submitted works	
<b>monash.edu</b>	1%
Internet	
<b>idus.us.es</b>	1%
Internet	
<b>coek.info</b>	1%
Internet	

<b>Alessia Pini, Simone Vantini. "The Interval Testing Procedure: A General Fram...</b>	1%
Crossref	
<b>Universidad Cesar Vallejo on 2023-07-04</b>	1%
Submitted works	
<b>Sarıçam, Semanur. "Fonksiyonel regresyon modellerinde kısmi en küçük karel...</b>	1%
Publication	
<b>Hiba Nassar. "A consistent estimator of the smoothing operator in the functio...</b>	1%
Crossref	
<b>univ-sba.dz</b>	1%
Internet	
<b>repositorio.unbosque.edu.co</b>	1%
Internet	
<b>acikbilim.yok.gov.tr</b>	1%
Internet	
<b>Ronan Baron, John Saffell. "Amperometric Gas Sensors as a Low Cost Emergi...</b>	1%
Crossref	
<b>Universidad Europea de Madrid on 2023-10-25</b>	1%
Submitted works	
<b>repositorio.usfq.edu.ec</b>	1%
Internet	
<b>c.coek.info</b>	<1%
Internet	
<b>Lydia Kara-Zaitri, Ali Laksaci, Mustapha Rachdi, Philippe Vieu. "Uniform in ban...</b>	<1%
Crossref	

**Florent Burba, Frédéric Ferraty, Philippe Vieu. " -Nearest Neighbour method in... <1%**

Crossref

---

**Ahmed Ait-Saïdi, Frédéric Ferraty, Rabah Kassa, Philippe Vieu. "Cross-validate... <1%**

Crossref

---

**Unviersidad de Granada on 2016-09-05 <1%**

Submitted works

---

**repositorio.senamhi.gob.pe <1%**

Internet

---

**revistas.unal.edu.co <1%**

Internet

---

**53250 on 2015-06-26 <1%**

Submitted works

---

**grafiati.com <1%**

Internet

---

**ntnuopen.ntnu.no <1%**

Internet

---

**ism.ac.jp <1%**

Internet

---

**slideplayer.fr <1%**

Internet

---

**acp.copernicus.org <1%**

Internet

---

**yumpu.com <1%**

Internet

---

<b>mdpi.com</b>	<1%
Internet	
<b>Politécnico Colombiano Jaime Isaza Cadavid on 2020-12-01</b>	<1%
Submitted works	
<b>birs.ca</b>	<1%
Internet	
<b>The University of Manchester on 2023-12-04</b>	<1%
Submitted works	
<b>arxiv-vanity.com</b>	<1%
Internet	
<b>mate.polimi.it</b>	<1%
Internet	
<b>renati.sunedu.gob.pe</b>	<1%
Internet	
<b>edoc.ub.uni-muenchen.de</b>	<1%
Internet	
<b>I M ALMANJAHIE. "MODERN STATISTICAL ANALYSIS OF FORAGE QUALITY ...</b>	<1%
Crossref	
<b>Amit Kumar Sharma, Munish Sharma, Arvind Kumar Sharma, Munit Sharma, ...</b>	<1%
Crossref	
<b>funcas.es</b>	<1%
Internet	
<b>aloki.hu</b>	<1%
Internet	

<b>projecteuclid.org</b>	<1%
Internet	
<b>epa.oszk.hu</b>	<1%
Internet	
<b>arxiv.org</b>	<1%
Internet	
<b>Weiwei Xiao, Yan Hu. "Functional data analysis of air pollution in six major citi..."</b>	<1%
Crossref	
<b>stat.wisc.edu</b>	<1%
Internet	
<b>redalyc.org</b>	<1%
Internet	
<b>rua.ua.es</b>	<1%
Internet	
<b>studylibfr.com</b>	<1%
Internet	
<b>iopscience.iop.org</b>	<1%
Internet	
<b>Larabi, Samah. "Nouvelles Approches de Calibration Automatique Appliquées ..."</b>	<1%
Publication	
<b>lib.kobe-u.ac.jp</b>	<1%
Internet	
<b>quijote.biblio.iteso.mx</b>	<1%
Internet	

<b>espace.inrs.ca</b>	<1%
Internet	
<b>Chirag Manchanda, Robert Harley, Julian Marshall, Alexander Turner, Joshua ...</b>	<1%
Crossref posted content	
<b>d.documentop.com</b>	<1%
Internet	
<b>utn on 2023-05-04</b>	<1%
Submitted works	
<b>pepite-depot.univ-lille.fr</b>	<1%
Internet	
<b>Peijun Sang, Liangliang Wang, Jiguo Cao. "Parametric Functional Principal Co...</b>	<1%
Crossref	
<b>Universidad San Ignacio de Loyola on 2022-10-31</b>	<1%
Submitted works	
<b>Pontificia Universidad Catolica del Peru on 2018-11-14</b>	<1%
Submitted works	
<b>Pontificia Universidad Catolica del Peru on 2018-12-10</b>	<1%
Submitted works	
<b>scribd.com</b>	<1%
Internet	
<b>limacomovamos.org</b>	<1%
Internet	
<b>repository.usta.edu.co</b>	<1%
Internet	

<b>repositorio.ug.edu.ec</b>	<1%
Internet	
<b>fdocuments.ec</b>	<1%
Internet	
<b>fatcat.wiki</b>	<1%
Internet	
<b>es.scribd.com</b>	<1%
Internet	
<b>arantxa.ii.uam.es</b>	<1%
Internet	
<b>Nathalia Celis, Alejandro Casallas, Ellie Anne López-Barrera, Hermes Martínez...</b>	<1%
Crossref	
<b>An Wang, Sanjana Paul, Priyanka deSouza, Yuki Machida, Simone Mora, Fábio...</b>	<1%
Crossref	
<b>Rosario Roman, Dennis. "A tool for functional data analysis and experimentati...</b>	<1%
Publication	
<b>Peter Hall. "On properties of functional principal components analysis", Journ...</b>	<1%
Crossref	
<b>D. Billheimer. "Functional Data Analysis, 2nd edition edited by J. O. Ramsay a...</b>	<1%
Crossref	
<b>André Rupp. "Quantifying Subpopulation Differences for a Lack of Invariance ...</b>	<1%
Crossref	
<b>A.J. Harrison, W. Ryan, K. Hayes. "Functional data analysis of joint coordinatio...</b>	<1%
Crossref	

<b>Caroline Mogollón-Sotelo, Alejandro Casallas, Sergio Vidal, Nathalia Celis, Ca...</b>	<1%
Crossref	
<b>proempresa.com.pe</b>	<1%
Internet	
<b>repositorio.usil.edu.pe</b>	<1%
Internet	
<b>etda.libraries.psu.edu</b>	<1%
Internet	
<b>Ludwig Fahrmeir, Thomas Kneib, Stefan Lang. "Regression", Springer Science ...</b>	<1%
Crossref	
<b>tdx.cat</b>	<1%
Internet	
<b>Chavez Vargas, Giovanna Paola. "Estudio de la gestion ambiental para la prev...</b>	<1%
Publication	
<b>Liverpool John Moores University on 2024-04-14</b>	<1%
Submitted works	
<b>Liverpool John Moores University on 2024-04-13</b>	<1%
Submitted works	
<b>Markham College on 2020-10-23</b>	<1%
Submitted works	
<b>Yunlong Nie, LiangLiang Wang, Jiguo Cao. "Estimating Time-Varying Directed ...</b>	<1%
Crossref	
<b>Garate Robles, Richard. "Planeamiento Estrategico del Distrito de Punta Negr...</b>	<1%
Publication	

<b>riunet.upv.es</b>	<1%
Internet	
<b>medworm.com</b>	<1%
Internet	
<b>cdr.lib.unc.edu</b>	<1%
Internet	
<b>Martinez Morales, Rodolfo. "Differentiation of forest stands, assessment of fo...</b>	<1%
Publication	
<b>Cavan Reilly. "Using Image and Curve Registration for Measuring the Goodnes...</b>	<1%
Crossref	
<b>Westcliff University on 2022-01-08</b>	<1%
Submitted works	
<b>Universidad de Salamanca on 2020-05-22</b>	<1%
Submitted works	
<b>Universidad de San Martín de Porres on 2019-11-12</b>	<1%
Submitted works	
<b>Universitat Politècnica de València on 2018-09-25</b>	<1%
Submitted works	
<b>sidalc.net</b>	<1%
Internet	
<b>revista.cincel.com.co</b>	<1%
Internet	
<b>mjfas.utm.my</b>	<1%
Internet	

<b>koha.ideam.gov.co</b>	<1%
Internet	
<b>facminas-unalmed.demo.elsevierpure.com</b>	<1%
Internet	
<b>era.ed.ac.uk</b>	<1%
Internet	
<b>Jin-Ting Zhang, Xuehua Liang, Shengning Xiao. "On the Two-Sample Behrens-...</b>	<1%
Crossref	
<b>Universidad Del Magdalena on 2017-11-18</b>	<1%
Submitted works	
<b>Universidad Del Magdalena on 2017-02-08</b>	<1%
Submitted works	
<b>Northcentral on 2022-08-17</b>	<1%
Submitted works	
<b>Universidad Tecnologica del Peru on 2022-07-28</b>	<1%
Submitted works	
<b>Westcliff University on 2022-01-07</b>	<1%
Submitted works	
<b>Corporación Universitaria Iberoamericana on 2024-04-23</b>	<1%
Submitted works	
<b>Corporación Universitaria Iberoamericana on 2024-04-23</b>	<1%
Submitted works	
<b>Corporación Universitaria Iberoamericana on 2024-04-22</b>	<1%
Submitted works	

<b>unboundmedicine.com</b>	<1%
Internet	
<b>seio.es</b>	<1%
Internet	
<b>pesquisa.bvsalud.org</b>	<1%
Internet	
<b>sciencedocbox.com</b>	<1%
Internet	
<b>Universidad Rey Juan Carlos on 2024-06-29</b>	<1%
Submitted works	
<b>Universidad Tecnica De Ambato- Direccion de Investigacion y Desarrollo , DID...</b>	<1%
Submitted works	
<b>Universidad Nacional Abierta y a Distancia, UNAD,UNAD on 2020-11-28</b>	<1%
Submitted works	
<b>Universidad Nacional Abierta y a Distancia, UNAD,UNAD on 2020-07-27</b>	<1%
Submitted works	
<b>Universidad Estatal a Distancia on 2016-08-21</b>	<1%
Submitted works	
<b>Universidad Estatal a Distancia on 2016-08-18</b>	<1%
Submitted works	
<b>Universidad Pedagogica y Tecnologica de Colombia on 2020-09-28</b>	<1%
Submitted works	
<b>Universidad Pedagogica y Tecnologica de Colombia on 2020-09-24</b>	<1%
Submitted works	