

FACTOR ANALYSIS Y CLUSTER ANALYSIS PARA CONCENTRACIONES DE PM_{10} EN MINAS DE CARBÓN A CIELO ABIERTO: CERREJÓN, COLOMBIA

Roberto Rojano, Heli Arregoces, Luis C. Angulo, Gloria Restrepo y Juan M. Marín

RESUMEN

La mina de carbón a cielo abierto del Cerrejón, ubicada al norte de Colombia, es una de las más grandes de Suramérica. Tiene una red de calidad de aire con más de 30 años de registro de niveles de material particulado. Entre enero 2012 y noviembre 2013 se registraron las concentraciones medias diarias de PM_{10} cada tres días en 8 estaciones de la red de calidad de aire, utilizando muestreadores de alto volumen. De igual manera, se estudiaron las mediciones de dirección y velocidad del viento en ese periodo de tiempo. El promedio de las concentraciones de PM_{10} en todas las estaciones variaron entre 17,49 y $42,01\mu g\cdot m^{-3}$. Estos niveles no excedieron el nivel máximo diario permisible según lo establecido en la guía de la Organiza-

ción Mundial de la Salud. Las técnicas de factor analysis (FA), cluster analysis (CA) y dirección y velocidad del viento fueron utilizados para inferir sobre las principales fuentes de contaminación y para determinar los patrones especiales de las concentraciones de PM_{10} . El FA identificó dos grupos con tres estaciones cada uno con el mismo comportamiento de concentraciones. Estos dos grupos tienen tajos (o canteras) de extracción viento arriba de su ubicación. Los resultados de CA mostraron tres clúster diferenciables con las mismas estaciones obtenidas en el FA. Los resultados de FA, CA y dirección y velocidad identificaron como fuentes potenciales de emisión los tajos o canteras de carbón ubicados viento arriba de las estaciones.

Introducción

Colombia es el primer productor de carbón de América Latina y el décimo en el mundo, con más de 74×10^6 toneladas métricas de producción anual. La mina de carbón del Cerrejón, ubicada en el departamento de La Guajira, al norte de Colombia, representa una de las minas más grandes a cielo abierto del mundo, con aporte del 38% en la producción y más del 50% en la exportación, convirtiéndose en la principal compañía de obtención de carbón térmico en el país (UPME, 2006). Esta capacidad de producción hace que las operaciones mineras realizadas a cielo abierto generen

material particulado en cantidades significativas. El material particulado con diámetro aerodinámico $<10\mu m$ (PM_{10}) es uno de los parámetros más importantes que se utilizan para evaluar la calidad del aire. Es sabido que la contaminación atmosférica en áreas mineras con estas características está ampliamente relacionada a la presencia significativa de material particulado en el ambiente (Huertas *et al.*, 2012, Hinwood *et al.*, 2014). El aumento de las concentraciones de PM_{10} en cercanías a zonas mineras está asociado con el aumento de síntomas respiratorios, agravamiento del asma, ingresos hospitalarios y muerte prematura; el riesgo es más alto para los

ancianos, niños y personas con asma o enfermedades del corazón (Qu *et al.*, 2007; Hendryx, 2008; Hendryx y Zullig, 2009).

Con el fin de proteger la salud humana, el Estado colombiano establece los niveles de inmisión mediante la Resolución 610 de 2010 (MAVDT, 2010) para todo el territorio nacional en condiciones de referencia. El nivel máximo permisible para un tiempo de exposición de 24h es $100\mu g\cdot m^{-3}$. Por ende, el monitoreo de la calidad de aire en la zona minera del Cerrejón debe ser gestionado adecuadamente. La ubicación de los puntos óptimos de muestreo, así como determinar el comportamiento y las

tendencias de los niveles de concentraciones, permite aumentar la eficacia para alcanzar tales objetivos respecto a la calidad del aire.

Esta eficacia se alcanza con un análisis sistemático de los datos disponibles. Hay varias técnicas de análisis de datos de contaminación atmosférica; entre ellas, la técnica de estadísticas de *factor analysis* (FA) y el *cluster analysis* (CA) han sido utilizado en diversas áreas de investigación. El objetivo de cualquier técnica de agrupación es maximizar la homogeneidad de elementos dentro de los grupos y maximizar la heterogeneidad entre el clúster (Makra *et al.*, 2011). En el área ambiental, FA se ha

PALABRAS CLAVE / Análisis Multivariado / Cluster Analysis / Factor Analysis / Minas de Carbón / PM_{10} /

Recibido: 06/11/2014. Modificado: 20/12/2016. Aceptado: 21/12/2016.

Roberto Rojano. Ingeniero Industrial, Universidad de la Guajira (Uniguajira), Colombia. Maestría en Ingeniería Ambiental, Universidad del Zulia, Venezuela. Candidato a Doctor en Ingeniería Ambiental, Universidad de Antioquia (UdeA), Colombia. Profesor, Uniguajira, Colombia. Dirección: Grupo de Investigación GISA, Facultad de Inge-

nería, Uniguajira. Km 5 Vía a Maicao, Riohacha, Colombia. E-mail: rrojano@uniguajira.edu.co

Heli Arregoces. Ingeniero Ambiental, Uniguajira, Colombia. Maestría en Ingeniería, UdeA, Colombia. Profesor, Uniguajira, Colombia.

Luis C. Angulo. Ingeniero de Minas, Universidad Pedagógica

y Tecnológica de Colombia. Maestría en Ciencias Ambientales, Sistema Universitario Estatal del Caribe Colombiano. Doctor en Ingeniería Ambiental, UdeA, Colombia. Profesor, Uniguajira y Universidad Popular del César, Colombia.

Gloria Restrepo. Ingeniera Química, Universidad Pontificia Bolivariana, Colombia. Doctora

en Ciencias Químicas, Universidad de Sevilla, España. Profesora, UdeA, Colombia.

Juan M. Marín. Ingeniero Químico, Maestría en Ingeniería Ambiental y Doctor en Ciencias Químicas, UdeA, Colombia. Profesor, UdeA, Colombia.

FACTOR AND CLUSTER ANALYSIS FOR PM₁₀ CONCENTRATIONS IN AN OPEN PIT COAL MINE: CERREJÓN, COLOMBIA

Roberto Rojano, Heli Arregoces, Luis C. Angulo, Gloria Restrepo and Juan M. Marín

SUMMARY

The Cerrejon coal mine in the north of Colombia is one of the largest open pits in South America. The Cerrejon has an air quality monitoring network with over 30 years of records of the levels of particulate material. Between January 2012 and November 2013, the mean daily concentrations of PM₁₀ were registered in eight stations of the air quality monitoring network using high volume samplers. Similarly, measurements of wind speed and direction were studied during this period. The average concentrations of PM₁₀ in the eight stations varied in the range 17,49-42,01µg·m⁻³. These levels do not exceed the maximum

permissible daily level as laid down in the guide of the World Health Organization. The techniques of factor analysis (FA) and cluster analysis (CA) as well as the wind direction and speed were used to infer the main sources of pollution and to determine the special patterns of PM₁₀ concentrations. The FA identified two groups with the same concentration behavior; these two groups have extraction pits located upwind. The results of the CA showed three differential clusters with the same stations identified in the FA. These techniques identified the coal pits located upwind of the stations as potential sources of PM₁₀.

FACTOR ANALYSIS E CLUSTER ANALYSIS PARA CONCENTRAÇÕES DE PM₁₀ EM MINAS DE CARVÃO A CÉU ABERTO: CERREJON, COLÔMBIA

Roberto Rojano, Heli Arregoces, Luis C. Angulo, Gloria Restrepo e Juan M. Marín

RESUMO

A mina de carvão a céu aberto do Cerrejon, localizada ao norte da Colômbia, é uma das maiores da América do Sul. Têm uma rede de qualidade de ar com mais de 30 anos de registro de níveis de material particulado. Entre janeiro de 2012 e novembro de 2013 se registraram as concentrações médias diárias de PM₁₀ cada três dias em 8 estações da rede de qualidade de ar, utilizando mostradores de alto volume. De mesma forma, se estudaram as medições de direção e velocidade do vento nesse período de tempo. A média das concentrações de PM₁₀ em todas as estações variaram entre 17,49 e 42,01µg·m⁻³. Estes níveis não excederam os máximos valores diários permitíveis segundo o estabelecido na guia da Organização

Mundial da Saúde. As técnicas de fator analysis (FA), cluster analysis (CA) e de direção e velocidade do vento foram utilizados para inferir sobre as principais fontes de contaminação e para determinar os padrões especiais das concentrações de PM₁₀. O FA identificou dois grupos com três estações cada um com o mesmo comportamento de concentrações. Estes dois grupos têm minas (ou pedreiras) de extração vento acima de sua localização. Os resultados de CA mostraram três clusters diferenciáveis com as mesmas estações obtidas na FA. Os resultados de FA, CA, e de direção e velocidade identificaram, como fontes potenciais de emissão, as minas ou pedreiras de carvão situadas vento acima das estações.

utilizado para identificar fuentes de contaminación de recursos ambientales (Yaouti *et al.*, 2009; Masoud, 2014), para identificar perfiles típicos en concentración de material particulado PM₁₀ durante el período de un año (Zibert y Praznikar, 2012), y para identificar y caracterizar las fuentes principales en un área costera afectada por las emisiones naturales y antropogénicas (Masiol *et al.*, 2010). Por otra parte, el CA ha sido utilizado como estrategia en la identificación y evaluación de los mecanismos de contaminación de cuerpos de agua (Güler *et al.*, 2002; Güler y Thyne, 2004; Cloutier *et al.*, 2008; Monjerezi *et al.*, 2011; Hynds *et al.*, 2014). En el estudio de la calidad del aire, estas técnicas han permitido determinar tendencias temporales, distribución espacial de contaminantes y comportamientos disímiles de

contaminación del aire (Gramsch *et al.*, 2006; Pires *et al.*, 2008, 2009; Dominick *et al.*, 2012). Sus aplicaciones han sido eficaces para la toma de decisiones en zonas con grandes grupos de poblaciones y problemas de contaminación atmosférica (Lu *et al.*, 2011).

El objetivo de esta investigación es evaluar la utilización de FA y CA para la gestión de calidad de aire a partir del monitoreo de PM₁₀ en el área del Cerrejón, buscando identificar áreas en la zona con compartimientos de concentraciones similares, factores comunes de emisión y los mecanismo de contaminación.

Metodología

Área de estudio

La región de explotación de carbón a cielo abierto más grande de Sur América se

encuentra al norte de Colombia, cerca de la frontera con Venezuela. En el departamento de La Guajira, península colombiana, está ubicada la mina el Cerrejón, cerca al municipio de Barrancas (Figura 1). Es una de las minas a cielo abierto más grandes del mundo. Está ubicada en las coordenadas 11°5'2''N y 72°40'31''O, y

limita con los municipios de Maicao, Albania, Hato nuevo, Barrancas y Fonseca.

Los depósitos de carbón se encuentran en una extensión de 69 000ha, dentro de la cual, hay cinco áreas de explotación (canteras) denominadas: tajo patilla, tajo 100, tajo Oreganal, tajo Tabaco y tajo La Puente.

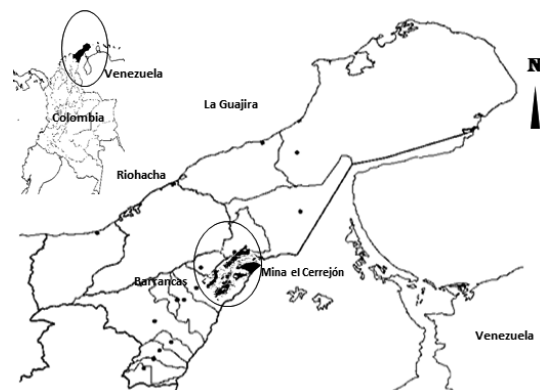


Figura 1. Localización del área de estudio. Tomado de Rojano *et al.* (2013).

Las concentraciones de PM₁₀ fueron registradas en la Red de Vigilancia de calidad de aire del Cerrejón desde enero 2012 hasta noviembre 2013. La Red de Vigilancia de calidad de aire es administrada por la misma empresa, con reporte periódico a la corporación autónoma ambiental de la región. Esta red cuenta con información de 18 estaciones, registrando mediciones de partículas sedimentables, TSP, PM₁₀ y PM_{2,5}. Para este estudio se tomaron las estaciones con equipos PM₁₀ y que estuvieran cercas a centros urbanos. Según estos criterios, se estudiaron las concentraciones de PM₁₀ de ocho sitios: Barrancas, Casitas 3, Patilla, Provincial, Roche, Sol y Sombra, Vivienda y Nuevo Roche. La Tabla I muestra las coordenadas y altitudes de estas estaciones, mientras que la Figura 2 muestra su localización geográfica en la zona.

La concentración de PM₁₀ fue determinada gravimétricamente según lo establecido en el método de referencia del Código de Regulaciones Federales de los EEUU (CFR, por sus siglas en inglés) por el cual se determina material particulado en la atmosfera <10µm (Appendix J to Part 50, 1997). Se empleó un muestreador de alto volumen con cabezal PM₁₀. El equipo Hi-Vol se operó durante las 24h, desde las 00:00 hasta las 24:00. Se utilizó filtros de fibra de cuarzo. El muestreo se desarrolló con un flujo constante de 1,13m³.min⁻¹. Los filtros se colocaron antes y después del muestreo en un desecador por un periodo de

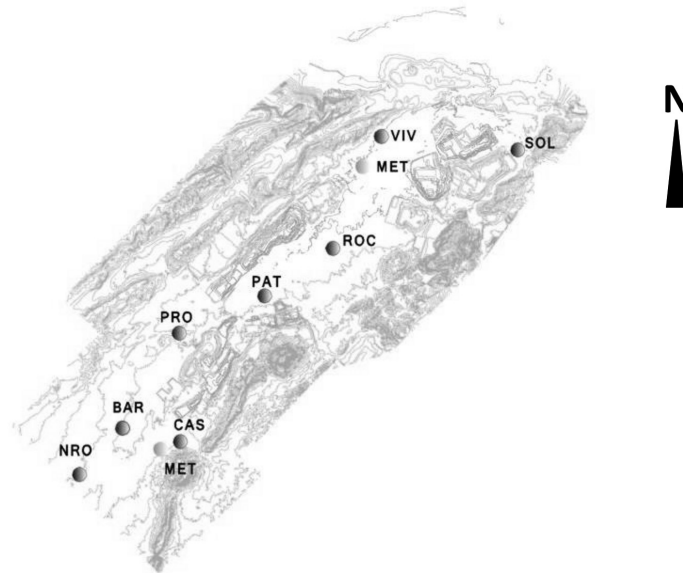


Figura 2. Localización de las estaciones de monitoreo en el área del Cerrejón.

24h y de igual manera, se pesaban antes y después del muestreo en una balanza analítica con cámara de pesaje con precisión de 1µg. Los filtros fueron codificados con el nombre de la estación, la numeración del equipo y la fecha. El equipo se calibró siguiendo el estándar de calibración recomendado por el fabricante. En cada punto se tomaron 120 muestras en el periodo comprendido desde el 01/01/2012 hasta el 30/11/2013.

Meteorología

Debido a que la característica topográfica de la zona donde están ubicadas las ocho estaciones es plana, la meteorología de los región fue caracterizada por los datos recogidos de una estación representativa del complejo carbonífero del Ce-

rejón. La estación meteorológica del Centro de Desarrollo Ambiental (CDA), está ubicada en las coordenadas 11°7'59,39"N y 72°37'49,88"O. Datos horarios de los años 2012 y 2013 fueron procesados para los parámetros velocidad y dirección del viento, temperatura, y precipitaciones. La rosa de viento se calculó utilizando el programa WindRose para los periodos de tiempo efectivo de muestreo.

Análisis estadístico

El *factor analysis* (FA) se utiliza para condensar un conjunto de variables medibles en un número menor de variables transformadas llamadas factores subyacentes o factores comunes; adicionalmente se identifica una estructura subyacente o patrón de un conjunto de datos multivariados (Yang y Trewn, 2004). En algunas áreas, el FA se utiliza para comprobar si las medidas de los factores son consistentes con el entendimiento y la naturaleza de las cargas y estructuras factoriales, lo que implica en el desarrollo de los modelos restricciones *a priori* dadas por hipótesis (Costa *et al.*, 2014). El modelo está dado por la Ec. 1:

$$x = \mu + \omega f + \epsilon \quad (1)$$

donde f: vector de mx1 que contiene las variables latentes o factores comunes, ω: matriz de pmx que contiene las cargas de los factores, y ε: vector de px1 que contiene las perturbaciones no observadas con distribución Np(0, φ), donde φ: matriz diagonal que no está correlacionada con los factores. En el modelo de FA x~(μ, Σ) cada observación está dada por la Ec. 2:

$$x_{ij} = j + \omega_{j1}f_{1i} + \omega_{j2}f_{2i} + \dots + \omega_{j2m}f_{mi} + \epsilon_{ij} \quad i = 1, 2, \dots, N; \quad j = 1, 2, \dots, N$$

La aplicación de FA a la contaminación atmosférica es idónea porque las variables originales son expresadas y están determinadas por combinaciones lineales de los factores, a diferencia de los componentes principales, que están definidos como combinaciones lineales de las variables originales (Rencher y Christensen, 2012). El FA busca explicar las covarianzas o correlaciones entre las variables presuponiendo un modelo estadístico formal de generación de las muestras dadas, donde los factores generan las variables observadas (Kim y Mueller, 1978; Basilevsky, 2009). Estas variables observadas y cuantificadas pueden ser interpretadas con base en los razonamientos ambientales de actividades propias de la zona. El número de variables se reduce a un número mínimo de variables independientes que describe adecuadamente los datos.

Por otro lado, en la técnica de *cluster analysis* (CA), el número de conglomerado puede ser elegido de acuerdo al conocimiento del usuario de la serie de datos; sin embargo, muchos investigadores lo determinan gráficamente a través de un dendrograma o a través de la regla #CA ≤ √n, donde n: número de datos (Romesburg, 2004; Pires *et al.*, 2009; Masoud, 2014). Desde el análisis de conglomerados se intenta determinar los vectores de observación que son similares y agruparlos en grupos, y muchas técnicas utilizan un índice

TABLA I
COORDENADAS Y ALTITUD DE LAS ESTACIONES

Estaciones	Coordenada N	Coordenada E	Elevación (msnm)
Barrancas (BAR)	1704078	1141838	149,6
Casitas 3 (CAS)	1703093	1146071	162
Patilla (PAT)	1713775	1152240	115,1
Provincial (PRO)	1711039	1145997	156
Roche (ROC)	1717243	1157236	119,3
Sol y Sombra (SOL)	1724465	1170756	117,6
Vivienda (VIV)	1725454	1160780	96,4
Nuevo Roche (NRO)	1700690	1138702	150
Est. Meteorológica (MET)	1723728	1160903	110

de similitud o proximidad entre cada par de observaciones (Rencher y Christensen, 2012). Una medida conveniente de la proximidad es la distancia entre dos observaciones. Una función de distancia común es la distancia euclidiana entre dos vectores $\mathbf{x} = (x_1, x_2, \dots, x_p)$ e $\mathbf{y} = (y_1, y_2, \dots, y_p)$, definidas por la Ec. 3:

$$d(\mathbf{w}, \mathbf{y}) = \sqrt{(\mathbf{w} - \mathbf{y})'(\mathbf{w} - \mathbf{y})} = \sqrt{\sum_{j=1}^p (x_j - y_j)^2} \quad (3)$$

Para el ajuste de diferentes varianzas y covarianzas entre las variables p , se podría utilizar la distancia estadística dada por la Ec. 4:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})' \mathbf{S}^{-1} (\mathbf{x} - \mathbf{y})} \quad (4)$$

donde \mathbf{S} es la matriz de covarianza de muestra.

Se han sugerido otras medidas de distancia para comparar nivel, variación y forma entre las variables, por ejemplo la distancia euclidiana al cuadrado, la cual es $d^2(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^p (x_j - y_j)^2$ y se expresa por la Ec. 5:

$$d^2(\mathbf{x}, \mathbf{y}) = (v_x - v_y)^2 + p(\bar{x} - \bar{y})^2 + 2v_x v_y (1 - r_{xy}) \quad (5)$$

donde $v_x^2 = \sum_{j=1}^p (x_j - \bar{x})^2$ y $\bar{x} = \sum_{j=1}^p x_j / p$, y , con similar expresión para v_y^2 y para \bar{y} . La correlación r_{xy} está dada por la Ec. 6:

$$r_{xy} = \frac{\sum_{j=1}^p (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{j=1}^p (x_j - \bar{x})^2 \sum_{j=1}^p (y_j - \bar{y})^2}} \quad (6)$$

Todo el análisis estadístico fue desarrollado con los programas STATGRAPHICS Centurión XVI.1 y R Project (i386 3.1.0). FA y CA fueron aplicados para el análisis de las concentraciones de PM_{10} registradas en los ocho sitios seleccionados. Los registros de las concentraciones fueron estandarizados para tener media= 0 y desviación estándar= 1. Se

minimizó el número de variables en cada factor, con coeficientes elevados en valor absoluto a través de rotaciones varimax. El tratamiento de valores faltantes se realizó por el método de *listwise*. De igual modo, para la aplicación CA se estandarizaron los valores de las variables. Las similitudes para la conformación de grupos fue determinada por el método de Ward y la métrica de distancias aplicada fue la euclidiana al cuadrado.

Resultados y Discusión

La Figura 3 muestra la variabilidad de las concentraciones de PM_{10} medida en cada estación en el de 2012 a 2013, incluyendo una comparación con la norma de calidad del aire local. Se muestra un gran porcentaje de variabilidad entre los valores mínimos y máximos registrados. La estación Casitas 3 presentó el mayor promedio de concentraciones entre las estaciones durante el periodo de estudio, con valor de $41,42 \mu\text{g}\cdot\text{m}^{-3}$, el segundo valor máximo de concentración diaria con un $86,34 \mu\text{g}\cdot\text{m}^{-3}$ y es la única estación en la que se presentó un valor superior de la mediana con respecto a la media, con una desviación estándar de $16,68 \mu\text{g}\cdot\text{m}^{-3}$, lo que sugiere una posibilidad rela-

tivamente alta de proyecciones superiores anuales a los $50 \mu\text{g}\cdot\text{m}^{-3}$. El promedio de las concentraciones en todas las estaciones registraron valores superiores a $29 \mu\text{g}\cdot\text{m}^{-3}$, a excepción de la estación Sol y Sombra, que presentó una media de $20,37 \mu\text{g}\cdot\text{m}^{-3}$. La estación Provincial registro el valor máximo de concentración diaria con $87,39 \mu\text{g}\cdot\text{m}^{-3}$. Los niveles de PM_{10} registrados en las ocho estaciones fueron inferiores a los obtenidos en minas de carbón ubicadas al norte de Colombia pero más al sur del Cerrejón (Angulo *et al.*, 2011; Huertas *et al.*, 2012). Estos niveles son mayores si se comparan con concentraciones reportadas en zonas urbanas de ciudades a menos de 180km de la zona minera (Amaya, 2008; Rojano *et al.*, 2011, 2013; Álvarez y Salazar, 2013; Candanoza *et al.*, 2013). Estos niveles de concentración no exceden el nivel máximo permisible diario para PM_{10} del estado colombiano, que es de $100 \mu\text{g}\cdot\text{m}^{-3}$ (MAVDT, 2010).

FA se aplicó como un método no paramétrico de clasificación para catalogar los sitios de monitoreo en clases que tienen el mismo comportamiento de contaminación y que difieren de otras clases. La Tabla II muestra los resultados al aplicar FA a las concentraciones

de PM_{10} en las estaciones estudiadas. El uso de cuatro factores para los datos de concentraciones considera una varianza acumulativa cercana al 100%. Las cargas factoriales rotadas, resaltadas en negritas, indican las variables que influyen sobre todo el FA correspondiente.

Para FA1 tiene significativas contribuciones de las estaciones Barrancas, Provincial y Nuevo Roche; mientras que FA2 fue muy cargado con los aportes de los sitios Patilla, Sol y Sombra, y Vivienda. FA3 tuvo importante contribución de la estación Casitas 3, mientras que FA4 fue asociado con la estación Roche. El FA1 geográficamente representa las estaciones ubicadas en la zona sur, a excepción de la estación Casitas 3, influenciada por 3 pits y por comunidades asentadas en el área. Aunque la estación Casita 3 se encuentra en esta área, su similitud y comportamientos de contaminación la ubicaron dentro de FA3, factor que probablemente está determinado por su cercanía al pit (llamado tajó 100) y la dirección predominante del viento.

CA también fue utilizado para agrupar los sitios de monitoreo basado en la similitud de las concentraciones de PM_{10} estandarizadas. Se utilizó la

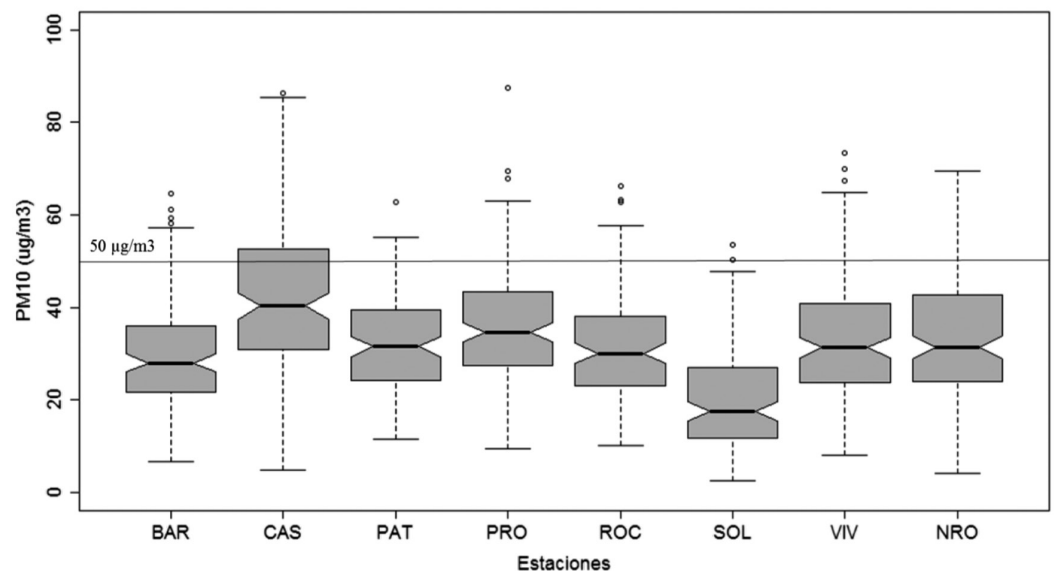


Figura 3. Boxplot de las concentraciones de PM_{10} en las estaciones de monitoreo.

TABLA II
PRINCIPALES RESULTADOS DE LA APLICACIÓN DE
FA PARA PM₁₀ EN LAS ESTACIONES DE MONITOREO

Estaciones	FA1	FA2	FA3	FA4
Barrancas	0,739	0,409	0,226	0,260
Casitas 3	0,444	0,146	0,592	0,232
Patilla	0,401	0,588	0,243	0,319
Provincial	0,805	0,337	0,183	0,149
Roche	0,363	0,428	0,367	0,629
Sol y Sombra	0,134	0,656	0,165	0,162
Vivienda	0,300	0,674	0,014	0,079
Nuevo Roche	0,789	0,237	0,360	0,190
Varianza (%)	82,08	10,43	5,69	1,49
Varianza acumulada (%)	82,08	92,50	98,19	99,68

*: Significativas contribuciones de las estaciones.

distancia euclidiana cuadrada para calcular la métrica de las distancias y el método de conglomeración fue el método de Ward para maximizar la homogeneidad dentro de los grupos. Para ello, el método plantea todas las posibles fusiones en cada etapa concreta y elige la que maximiza la homogeneidad a través de cálculo de los centroides de los grupos resultantes de las posibles fusiones; posteriormente, calcula la distancia al centroide de todas las observaciones del grupo (Basilevsky, 2009). La solución con menor suma de cuadrados total es la elegida. La Figura 4 muestra el dendograma resultante de la aplicación de CA a la concentración de masa de PM₁₀.

Los resultados obtenidos muestran tres grupos diferenciables. El grupo CA1 está compuesto por las estaciones

de monitoreo que integran al FA1; los factores predominantes están determinados por asentamientos significativamente poblados e influencia de los tajos de producción de la zona sur. El CA2 reúne como factores comunes aproximaciones cercanas a fuentes de emisiones altas en el proceso minero como vías de acarreos y remoción de estéril. El grado de impacto está relacionado con la velocidad y dirección de vientos. Igualmente, CA3 lo integran estaciones donde debido a la dirección de los vientos sus registros se ven poco influenciados por las actividades mineras; podrían considerarse estaciones de fondo. La Figura 5 muestra el comportamiento de la concentración de PM₁₀ por los grupos del CA.

Se interpretó los resultados analizando el efecto de la

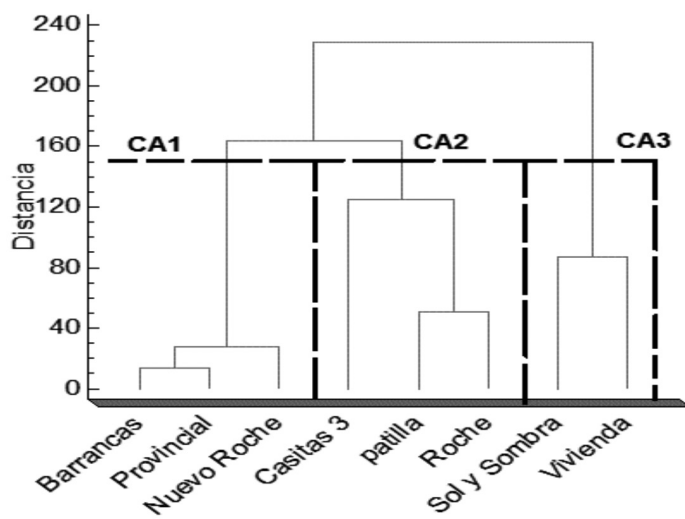


Figura 4. Dendrograma estaciones de PM₁₀.

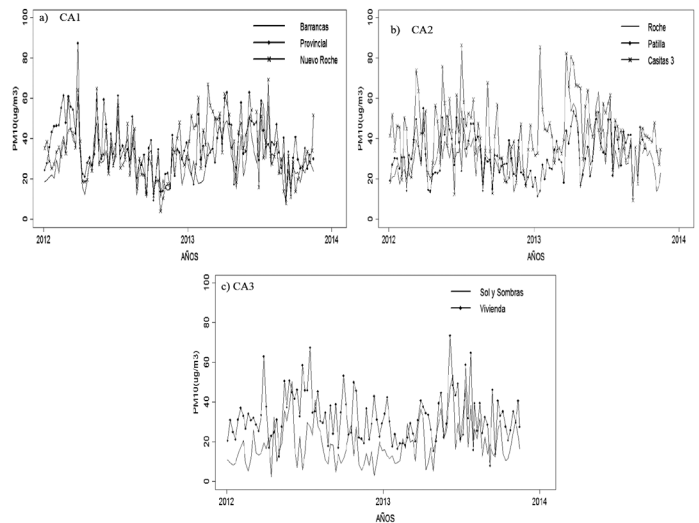


Figura 5. Identificación de grupos de concentraciones de PM₁₀ agrupados por CA. CA1: grupo 1; CA2: grupo 2 y CA3: grupo 3.

dirección del viento sobre las concentraciones de PM₁₀, con el fin de inferir posibles fuentes de emisiones que pueden impactar directamente sobre los sitios agrupados por CA y FA. La rosa de los vientos para el año 2012 (Figura 6a) muestra una dirección predominante de NE, pero baja velocidad media del viento (2,82m·s⁻¹) y bajo porcentaje de viento en calma (2,20%). El año 2013 (Figura 6b) registró mayor promedio de velocidad de viento (2,92m·s⁻¹) y mayor porcentaje de viento en calma (7,57%) que el año 2012. El año 2013 registró diferencia del 10% en la velocidad del viento en el intervalo 3,6-8,8m·s⁻¹ en relación con el 2012. Este intervalo de velocidad permite mayor dispersión de los contaminantes y

aumenta la dilución de los mismos, haciendo que se presenten menores concentraciones en los receptores.

En los dos años se presentó una frecuencia predominante de NE y un 70% de velocidad del viento en el rango 2,1-5,7m·s⁻¹. El vector de dirección resultante de 61° (NNE) se presentó en el 68% de los resultados, lo que indica un transporte de contaminante en el 68% en la dirección 241° (SOO). Estudios en zonas mineras cercanas al área de estudio mostraron resultados donde los parámetros meteorológicos de estaciones ubicadas en minería a cielo abierto al norte de Colombia (departamento del César) mostraron un alto nivel de correlación (R²>0,95) con las concentraciones de PST

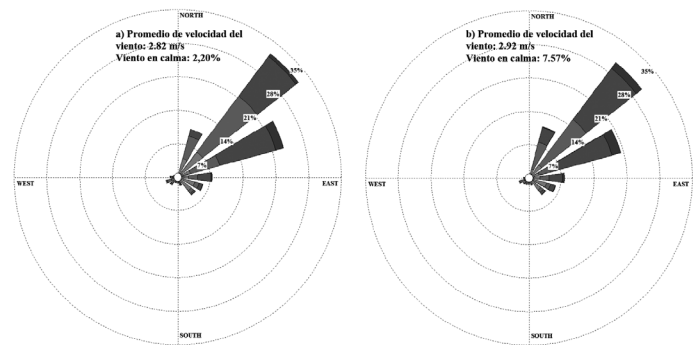


Figura 6. Rosa de viento para los años 2012 (izquierda) y 2013 (derecha) en la mina el Cerrejón.

(Huertas *et al.*, 2014), por lo que se puede inferir que todo el aporte de PM₁₀ originados en los tajos u otras actividades mineras, tiene alta probabilidad de impactar a los receptores en la dirección predominante.

En este estudio, el análisis de velocidad y dirección del viento muestra que el grupo CA1, comprendido por las estaciones de Barrancas, Provincial y Nuevo Roche, está directamente influenciado por el tajo Patilla (el tajo de mayor dimensión y actividad en la mina) y el botadero Patilla (zona de almacenaje de material estéril). La estación Provincial es la más cercana al tajo Patilla (<5km) y está ubicada directamente en la dirección predominante del viento, coincidiendo con los mayores registros y promedio en el periodo de estudio en esta estación. El análisis estadístico de CA y FA no describe un fenómeno físico de causa y efecto; por lo tanto, no se puede utilizar para identificar las fuentes de contaminación influyentes en las estaciones Barrancas, Provincial y Nuevo Roche. Sin embargo, las principales fuentes de emisión son las fuentes del área, que están asociados con las actividades de extracción de carbón, y se puede inferir una influencia del tajo y el botadero Patilla en el aporte de PM₁₀.

El grupo CA2, conformado por las estaciones Las Casitas 3, Patilla y Roche, están influenciado por los tajos: Ore-

ganal, 100, Comunero, Patilla y Caypa. La estación Las Casitas 3 está a menos de 3km de los tajos Comunero y 100, y está directamente en el vector de dirección resultante del viento de 241° (SOO). Esta estación es la que presenta las mayores concentraciones en este estudio. Entre esta estación y los tajos mencionados no hay fuentes significativas de emisión de PM₁₀, por lo que se concluye que el aporte potencial de material particulado en esta estación es producido por las actividades de extracción de los tajos Comunero y 100.

Por último, el grupo CA3 está integrado por las estaciones Sol y Sombra, y Vivienda. Estas estaciones están ubicadas viento arriba del área de extracción. Solo son afectadas cuando se presentan vientos en la dirección NNO y, según el análisis meteorológico, esta dirección registró una frecuencia de 2% en el periodo 2012 y 2013. Se puede inferir que estas dos estaciones captan aporte de PM₁₀ de actividades normales en centros urbanos, sea transporte, resuspensión de material, manipulación mecánica de suelos y otros. Los resultados de estas dos estaciones son parecidos a los reportados por Mendoza y Rojano (2010) en zonas rurales sin intervención minera a 20km al oeste de la mina el Cerrejón.

La Tabla III presenta el porcentaje de frecuencia en las concentraciones, registradas en

2012, con referencia a las direcciones del viento. La Figura 6 muestra la rosa de viento para los años 2012 y 2013. Los resultados revelan que las fuentes ubicadas NE-E ejercen mayores influencias significativas sobre las estaciones ubicadas en la zona sur. Las concentraciones pertenecientes a las estaciones que integran a CA1 se vieron influenciadas en gran manera por la dirección del viento. La dirección NE-E fue la predominante, con valores de frecuencias >34% entre los intervalos de concentraciones de 21-40µg·m⁻³, direcciones que sugieren fuentes de emisiones debido a los pits. La estación Casita 3 muestra que en mayores frecuencia del viento proveniente de NE-E se alcanzan valores >40µg·m⁻³. El CA3 infiere que alcanza concentraciones relativas provenientes de los pits de la zona norte.

La existencia de diferentes comportamientos de contaminación del aire en la red de monitoreo puede ser explicada por la ubicación geográfica de las fuentes de emisiones compuestas por los diferentes tajos, por la variabilidad de las direcciones del viento, y el tipo de actividad en la zona. El número de grupos de centros con similares comportamiento de la contaminación fue influenciado por las características geográficas de la ubicación de las fuentes y las tasas de emisiones de la zona;

esto significa que los sitios pueden ser agrupados de acuerdo a este comportamiento específico. Se muestra que FA y CA tienen un gran potencial para la gestión de aire en sistemas de monitoreo de la calidad de aire.

Conclusiones

El *factor analysis* (FA) mostró que la agrupación de las estaciones de PM₁₀ en la zona está definida por cuatro factores que contiene el 99,68% de la varianza acumulada de los datos originales, indicando una fuente principal de contaminación de PM₁₀ (operaciones mineras en los tajos). El FA3 mostró solamente el sitio de monitoreo de las Casitas 3, estación donde se presentaron los mayores niveles de PM₁₀. El *cluster analysis* (CA) indicó que las estaciones de Barrancas, Provincial y Nuevo Roche presentan comportamientos similares. Su explicación está sustentada por encontrarse en un área influenciada por las mismas actividades de minería en la zona y el asentamiento de zonas residenciales cercanas. Este método agrupó tres grupos, donde el CA1 pertenece a FA1 y donde CA3 representa los sitios de monitoreo con menor influencia de las actividades mineras por las direcciones predominantes de los vientos (dirección NE-E). En resumen, la contribución significativa de polvo mineral se debe a las fuentes locales de operación mineras.

AGRADECIMIENTOS

Los autores agradecen a Colciencias-Cerrejón, proyecto código 1115-524-30465, a la Universidad de Antioquia y a la Universidad de la Guajira por el apoyo para el desarrollo de este trabajo.

REFERENCIAS

Amaya M. (2008). *Estudio de Emisiones Atmosféricas para el Municipio Riohacha en los Contaminantes Partículas Suspensas Totales (PST) y Menores de 10 micras PM₁₀*. K2 Ingeniería-Corporación Autó-

TABLA III
FRECUENCIAS DE DIRECCIÓN DE VIENTO EN LAS CONCENTRACIONES DE PM₁₀ REGISTRADAS EN EL 2012 Y 2013, EN %

Dirección del viento	Barrancas			Casitas 3			Patilla			Provincial		
	0-20	21-40	>40	0-20	21-40	>40	0-20	21-40	>40	0-20	21-40	>40
N-NE	0,0	6,6	0,0	0,0	3,3	3,3	0,0	6,6	0,0	0,0	3,3	3,3
NE-E	6,6	39,3	9,8	6,6	16,4	32,8	8,2	39,3	8,2	0,0	34,4	21,3
E-SE	11,5	13,1	1,6	6,6	9,8	9,8	4,9	16,4	4,9	6,6	16,4	3,3
SE-S	1,6	4,9	3,3	1,6	3,3	4,9	0,0	4,9	4,9	1,6	4,9	3,3
Otros	1,6	0,0	0,0	0,0	1,6	0,0	0,0	1,6	0,0	1,6	0,0	0,0
	Roche			Sol y Sombra			Vivienda			Nuevo Roche		
	0-20	21-40	>40	0-20	21-40	>40	0-20	21-40	>40	0-20	21-40	>40
N-NE	1,8	3,6	0,0	44,3	0,0	0,0	0,0	6,6	0,0	1,6	3,3	1,6
NE-E	5,5	43,6	3,6	6,6	9,8	1,6	1,6	8,2	14,8	3,3	39,3	13,1
E-SE	9,1	16,4	3,6	16,4	8,2	1,6	4,9	11,5	9,8	11,5	9,8	1,6
SE-S	1,8	9,1	0,0	3,3	6,6	0,0	1,6	39,3	0,0	3,3	6,6	0,0
Otros	0,0	1,8	0,0	0,0	1,6	0,0	0,0	1,6	0,0	1,6	0,0	0,0

- noma regional de La Guajira. Riohacha, Colombia. 116 pp.
- Angulo LC, Huertas JI, Restrepo GM (2011) Caracterización de Partículas Suspendidas (PST) y Partículas Respirables (PM₁₀) producidas en Áreas de Explotación Carbonífera a Cielo Abierto. *Inf. Tecnol.* 22(4): 23-34.
- Álvarez ML, Salazar A (2013) Síntomas respiratorios y función pulmonar en niños de 6 a 14 años de edad y su relación con la contaminación por material particulado PM₁₀ en Santa Marta Colombia. *Rev. Esp. Salud Públ.* 87: 239-246.
- Basilevsky AT (2009) *Statistical Factor Analysis and Related Methods: Theory and Applications*. Wiley. Winnipeg, Canadá. 768 pp.
- Candanoza S, Goribar LL, García F (2013) Relación partículas respirables (PM₁₀)/partículas suspendidas totales (PST) en Santa Marta (Colombia). *Dyna* 179: 157-164.
- Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J. Hydrol.* 353: 294-313.
- Costa DR, Lachosa VH, Bazanb JL, Azevedoa CL (2014) Estimation methods for multivariate tobit confirmatory factor analysis. *Comput. Stat. Data Anal.* 79: 248-260.
- Dominick D, Juahir H, Latif MT, Zain SM, Aris AZ (2012) Spatial assessment of air quality patterns in Malaysia using multivariate analysis. *Atmos. Environ.* 60: 172-181.
- Gramsch E, Cereceda-Balic F, Oyola P, Von Baer D (2006) Examination of pollution trends in Santiago de Chile with cluster analysis of PM₁₀ and Ozone data. *Atmos. Environ.* 40: 5464-5475.
- Güler C, Thyne GD (2004) Hydrologic and geologic factors controlling surface and groundwater chemistry in Indian Wells-Owens Valley area, southeastern California, USA. *J. Hydrol.* 285: 177-198.
- Güler C, Thyne GD, McCray JE, Turner KA (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol. J.* 10: 455-474.
- Hendryx M (2008) Mortality rates in Appalachian coal mining counties: 24 years behind the nation. *Environ. Justice* 1: 5-11.
- Hendryx M, Zullig KJ (2009) Higher coronary heart disease and heart attack morbidity in Appalachian coal mining regions. *Prev. Med.* 49: 355-359.
- Hinwood A, Callan AC, Heyworth J, McCafferty P, Sly PD (2014) Children's personal exposure to PM₁₀ and associated metals in urban, rural and mining activity areas. *Chemosphere* 108: 125-133.
- Huertas J, Huertas M, Solís D (2012) Characterization of airborne particles in an open pit mining region. *Sci. Total Environ.* 423: 39-46.
- Huertas JI, Huertas ME, Cervantes G, Diaz J (2014) Assessment of the natural sources of particulate matter on the opencast mines air quality. *Sci. Total Environ.* 493: 1047-1055.
- Hynds P, Misstear BD, Gill LW, Murphy HM (2014) Groundwater source contamination mechanisms: Physicochemical profile ing, risk factor analysis and multivariate modelling. *J. Contam. Hydrol.* 159: 47-56.
- Kim JO, Mueller CW (1978) *Factor Analysis: Statistical Methods and Practical*. Sage. Thousand Oaks, CA, EEUU. 88 pp.
- Lu WZ, He HD, Dong LY (2011) Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis. *Build. Environ.* 6: 577-583.
- Makra L, Matyasovszky I, Guba Z, Karatzas K, Anttila P (2011) Monitoring the long-range transport effects on urban PM₁₀ levels using 3D clusters of backward trajectories. *Atmos. Environ.* 45: 2630-2641.
- Masiol M, Rampazzo G, Ceccato D, Squizzato S, Pavoni B (2010) Characterization of PM₁₀ sources in a coastal area near Venice (Italy): An application of factor-cluster analysis. *Chemosphere* 80: 771-778.
- Masoud AA (2014) Groundwater quality assessment of the shallow aquifers west of the Nile Delta (Egypt) using multivariate statistical and geostatistical techniques. *J. Afr. Earth Sci.* 95: 123-137.
- Mendoza YI, Rojano R (2010) *Monitoreo de la Calidad de Aire (Partículas Menores de 10 micras (PM₁₀), Polvo Sedimentales y Gases SO₂ y NO₂); en la Población de Angostura-Municipio de Hatonuevo. Departamento de La Guajira. Instituto de Estudios Ambientales. Universidad de La Guajira. Colombia. 180 pp.*
- MAVDT (2010) *Resolución 0610 de 2010. Por la cual se Establece la Norma de Calidad del Aire o Nivel de Inmisión, para todo el Territorio Nacional en Condiciones de Referencia*. Ministerio de Ambiente, Vivienda y Desarrollo Territorial *Diario Oficial* 47.672 de abril 6 de 2010, 17 pp. Bogotá, Colombia.
- Monjerezi M, Vogt RD, Aagaard P, Saka JD (2011) Hydrogeochemical processes in an area with saline groundwater in lower Shire River valley, Malawi: an integrated application of hierarchical cluster and principal component analyses. *Appl. Geochem.* 26: 1399-1413.
- Pires JCM, Sousa SIV, Pereira MC, Alvim-Ferraz MC, Martins FG (2008) Management of air quality monitoring using principal component and cluster analysis - Part I: SO₂ and PM₁₀. *Atmos. Environ.* 42: 1249-1260.
- Pires JCM, Pereira MC, Alvim-Ferraz MC, Martins FG (2009) Identification of redundant air quality measurements through the use of principal component analysis. *Atmos. Environ.* 43: 3837-3842.
- Qu Y, Tang Y, Cao D, Wu F, Liu J, Lu G, Zhang Z, Xia Z (2007) Genetic polymorphisms in alveolar macrophage response-related genes, and risk of silicosis and pulmonary tuberculosis in Chinese iron miners. *Int. J. Hygiene Environ. Health* 210: 679-689.
- Rencher AC, Christensen WF (2012) *Methods of Multivariate Analysis*. 3^a ed. Wiley. Hoboken, NJ, EEUU. 800 pp.
- Rojano R, Angulo LC, Restrepo G (2013) Niveles de Partículas Suspendidas Totales (PST), PM₁₀ y PM_{2.5} y su Relación en Lugares Públicos de la Ciudad Riohacha, Caribe Colombiano. *Inf. Tecnol.* 24: 37-46.
- Rojano R, Pérez J, Deluque J (2011) Análisis comparativo de las mediciones de material particulado PM₁₀. *Rev. Fac. Ing. Univ. Antioquia* 58: 27-35.
- Romesburg C (2004) *Cluster Analysis for Researchers (Vol. 12)*, Lulu Press. Morrisville, NC, EEUU. 334 pp.
- UPME (2006) *Plan Nacional para el Desarrollo Minero. Visión Año 2019*. Unidad de Planeación Minero Energética. Ministerio de Minas y Energías. Bogotá, Colombia. 125 pp.
- Yang K, Trewn J (2004) *Multivariate Statistical Methods in Quality Management*. McGraw Hill. Nueva York, NY, EEUU. 229 pp.
- Yaouti FE, Mandour EA, Khattach D, Benavente J, Kaufmann O (2009) Salinization processes in the unconfined aquifer of Bou-Areg (NE Morocco): A geostatistical, geochemical, and tomographic study. *Appl. Geochem.* 24: 16-31.
- Zibert J, Praznikar J (2012) Cluster analysis of particulate matter (PM₁₀) and black carbon (BC) concentrations. *Atmos. Environ.* 5: 1-12.