

# Metodología para la caracterización espacio-temporal de PM2.5 en el área urbana de la ciudad de Medellín-Colombia

✉ LIBARDO ANTONIO LONDOÑO-CIRO<sup>1</sup>

JULIO EDUARDO CAÑÓN-BARRIGA<sup>2</sup>

## Resumen

Se propone una metodología para caracterizar espacio temporalmente la concentración de PM2.5 que integra interpolación espacial, estadística espacial y econometría. Se implementa con sistemas de información geográfica y datos de concentración promedio mensual de PM2.5, temperatura y velocidad del viento; medidos en 8 sitios de monitoreo durante 2013-2014 en la ciudad de Medellín. Primero se hace una caracterización espacial multivariada de PM2.5 con regresiones geográficamente ponderadas (GWR) y luego la caracterización temporal con algoritmos econométricos. El modelo con variables explicativas temperatura, gradiente espacial del viento, dominio espacial de PM2.5 calculado con GWR y la concentración de PM2.5; con un rezago de orden 1; explica en un 87% la variabilidad de PM2.5. El desempeño de los algoritmos se calcula con validación cruzada. Finalmente se ajustan los mapas de caracterización espacio-temporal con un modelo de proximidad espacial que usa la distancia a fuentes de emisión y de mitigación de PM2.5.

**Palabras Claves:** Contaminación del Aire, Análisis Espacial, Regresión Geográficamente Ponderada, Modelos Autorregresivos de Rezago Distribuido.

## Spatio-Temporal Characterization Methodology of PM2.5 in the Urban Area of the City of Medellín-Colombia

## Abstract

In this paper, we propose a methodology to characterize the spatiotemporal concentration of PM2.5 that integrates spatial interpolation, spatial statistics and econometrics. It is implemented with geographic information systems and

<sup>1</sup> Politécnico Colombiano Jaime Isaza Cadavid; Universidad de San Buenaventura; Universidad de Antioquia, Medellín, Colombia.

<sup>2</sup> Universidad de Antioquia, Medellín, Colombia.



*Autor de correspondencia:* Londoño-Ciro, L.A. (Libardo Antonio): Cra 21 # 47-122, apto. 1501, Medellín, Colombia.  
Teléfono: (574)2692617.  
Correo electrónico: libaranto@gmail.com

### *Historia del artículo:*

Artículo recibido: 29-XI-2017 / Aprobado: 18-V-2018  
Disponible online: 3 de agosto 2018  
Discusión abierta hasta abril de 2020

monthly average concentration data of PM2.5, temperature and wind speed measured in eight monitoring sites during 2013-2014 in Medellín. First we make a multivariate spatial characterization of PM2.5 with Geographically Weighted Regressions (GWR) and then the temporal characterization with econometric algorithms. The model with explanatory variables temperature, spatial wind gradient, spatial domain of PM2.5 calculated with GWR and the concentration of PM2.5 with a lag of order one explains the variability of PM2.5 by 87%. The performance of the algorithms is calculated with cross validation. Finally, spatio-temporal characterization maps are adjusted with a spatial proximity model that uses distance to PM2.5 emission and mitigation sources.

**Keywords:** Air Pollution, Spatial Analysis, Geographically Weighted Regression, Autoregressive Distributed Lag Models.

## Metodologia para a caracterização espaço-temporária de PM2.5 na área urbana da cidade de Medellín-Colômbia

### Resumo

Propomos uma metodologia para caracterizar temporariamente a concentração de PM2.5 que integra a interpolação espacial, estatística espacial e econometria. É implementado com sistemas de informação geográfica e dados de concentração média mensais de PM2.5, temperatura e velocidade do vento medido em 8 locais de monitoramento durante 2013-2014 na cidade de Medellín. Primeiro, fazemos uma caracterização espacial multivariada de PM2.5 com regressões geograficamente ponderadas (GWR) e, em seguida, a caracterização temporal com algoritmos econométricos. O modelo com variáveis explicativas de temperatura, gradiente de vento espacial, domínio espacial de PM2.5 calculado com GWR e a concentração de PM2.5 com um atraso de ordem 1 explica a variabilidade de PM2.5 em 87%. O desempenho dos algoritmos é calculado com validação cruzada. Finalmente, os mapas de caracterização espaço-temporal são ajustados com um modelo de proximidade de espacial que usa distância para fontes de mitigação e emissão de PM2.5.

**Palavras-chave:** Poluição do Ar, Análise Espacial, Regressão Geograficamente Ponderada, Modelos Autoregressivos de Atraso Distribuído.

---

### 1. Introducción

Diversos estudios han demostrado que los contaminantes del aire, en especial el material particulado de diámetro menor a 2.5 micras (PM2.5), tienen un efecto significativo sobre la salud (mortalidad y morbilidad) de los habitantes de las grandes ciudades. El PM2.5 constituye la denominada fracción inhalable de partículas capaces de llegar a la zona bronquio-traqueal y que está compuesto por una variedad de sustancias sólidas y líquidas derivadas de fuentes

naturales y actividades humanas que forman una mezcla compleja de carbono orgánico e inorgánico, metales, nitratos, sulfatos y fosfatos (Sanhueza *et al.*, 2009). Esto motiva la configuración e implementación de redes de sitios de monitoreo diseñadas para muestrear, analizar y procesar las concentraciones de PM2.5 en forma continua durante un cierto período.

La configuración de las redes de sitios de monitoreo ha generado a nivel mundial una serie de investigaciones que tienen como factor común la problemática asociada con la insuficiencia en la

cantidad de datos para hacer una adecuada caracterización espacial y temporal de PM<sub>2.5</sub> (Li & Heap, 2014) y hacer más efectiva la toma de decisiones en cuanto a determinación de alertas tempranas para mitigar la contaminación (Paschalidou *et al.*, 2011). El problema de la insuficiencia de datos en el espacio ha sido abordado en los últimos diez años con algoritmos computacionales para estimar el valor de la concentración de un contaminante en sitios que no tienen mediciones (Alam & McNabola, 2015; Londoño *et al.*, 2017). Los algoritmos computacionales han tenido tres enfoques: determinístico, estocástico y de interpolación espacial (Shahraiyni & Sodoudi, 2016).

El enfoque estocástico se basa en la descripción de relaciones estadísticas entre la concentración del contaminante y variables explicativas. Operativamente se estima el valor de una variable, implementando regresiones lineales o no lineales con el algoritmo de mínimos cuadrados ordinarios (Goyal *et al.*, 2006) y, para el caso de datos espaciales, con el algoritmo de Regresión Geográficamente Ponderada (Geographical y Weighted Regression – GWR) (Fotheringham *et al.*, 2002). Algunas variables explicativas usadas en este enfoque han sido: la concentración de otros contaminantes (Deligiorgi & Philippopoulos, 2011), variables meteorológicas (Ramsay *et al.*, 2011), variables de uso del suelo, distancia a vías, distancia a fuentes de emisión, densidad de tráfico, entre otras (Modelos de Regresión de Usos del Suelo-Land Use Regression LUR) (Lee *et al.*, 2017; Londoño & Cañón, 2015; Dons *et al.*, 2014). Para el caso de la concentración del material particulado, los estudios han encontrado que las variables explicativas más importantes son, en su orden, la velocidad del viento y la temperatura del aire (Sayegh *et al.*, 2014). Otras variables explicativas consideradas para la variación de corto plazo en zonas urbanas son la densidad de flujo vehicular y, en caso de no tener información al respecto, algunos autores las han reemplazado por la concentración de SO<sub>2</sub> y NO<sub>x</sub> (Singh *et al.*, 2017, Chen L. *et al.*, 2010). Adicionalmente se ha encontrado que la presencia de zonas verdes y cuerpos de agua afecta la

concentración del material particulado en la medida en que impiden su resuspensión (TaheriShahraiyni *et al.*, 2015). En general los algoritmos del enfoque estocástico permiten hacer estimaciones locales, tienen buen desempeño y no requieren mucha información, pero no consideran la física detrás de los datos y además el algoritmo aplicado para un área de estudio no siempre es aplicable a otras.

El enfoque de la interpolación espacial se basa en el principio de autocorrelación espacial que exhiben las variables ambientales (en la cercanía a un sitio de monitoreo, el valor de la variable es similar y en la medida en que se aleja del sitio el valor difiere más). Operativamente, la mayoría de los algoritmos estiman el valor de la variable en un punto en donde no se tiene medición con base en una función que es el promedio ponderado de los valores medidos en otros puntos (Sangalli *et al.*, 2013). Los algoritmos de interpolación espacial se clasifican en geoestadísticos y no geoestadísticos, según el cálculo de los pesos ponderantes (Li & Heap, 2014).

Con respecto al problema de la insuficiencia de datos en el tiempo, se han usado algoritmos computacionales para estimar y pronosticar el valor de la concentración de un contaminante (Sayegh *et al.*, 2014). Estos algoritmos han tenido tres enfoques principales: Regresión Lineal Multivariada (Multivariate Linear Regression – MLR), Redes Neuronales Artificiales (*Artificial Neural Networks* – ANN) y Económico (Shahraiyni & Sodoudi, 2016).

El enfoque MLR ha sido ampliamente utilizado para el pronóstico de material particulado en centros urbanos (Goyal *et al.*, 2006). Este enfoque muestra un buen desempeño, pero tiene imprecisiones debido a la representación lineal de sistemas no lineales y muestra falencias para el pronóstico de los valores extremos de la variable (episodios). El enfoque ANN se ha utilizado para el pronóstico de la concentración de una amplia gama de contaminantes a diferentes escalas de tiempo con muy buenos resultados (Şahin *et al.*, 2011). El problema de este enfoque es que no es aplicable para otro período y lugar diferente al de los datos (Pilsung, 2013). En el

enfoque econométrico, se pronostica el valor de la variable de estudio (variable respuesta) a partir de otras variables explicativas (o regresoras) que pueden incluir a la misma variable respuesta rezagada (Zhang, 2003). Estos algoritmos tienen un buen desempeño y permiten hacer pronóstico de eventos extremos (episodios), pero tienen el problema de la representación lineal de sistemas no lineales y el requerimiento de datos históricos continuos.

En este artículo se usa el enfoque estocástico con el algoritmo GWR para la caracterización espacial multivariada de la concentración de PM2.5 y el enfoque econométrico a través de algoritmos Autorregresivos de Rezago Distribuido (AR-RD) de rezago 1 para la caracterización temporal. El desempeño de los algoritmos y la validación de los valores estimados se hace mediante pruebas de validación cruzada y el cálculo del error estándar promedio (ASE), la raíz del error cuadrático promedio (RMS), la raíz del error cuadrático promedio estandarizada (RMSS) y el error absoluto relativo promedio (MARE). Adicionalmente se introduce la variable distancia a fuentes de emisión y de mitigación, usando un Modelo de Proximidad Espacial (MPE) para hacer ajustes a los mapas obtenidos de caracterización espacio-temporal de la concentración de PM2.5.

Finalmente, los resultados de las estimaciones muestran un comportamiento similar a los datos medidos, con una estacionalidad mensual en el fenómeno de la contaminación por PM2.5, en donde marzo es el mes más contaminado con unos valores de concentración de PM2.5 alrededor de  $31 \mu\text{g}/\text{m}^3$ , valor que está por encima del máximo promedio anual que en Colombia es de  $25 \mu\text{g}/\text{m}^3$  (Resolución 610 de 2010 del Ministerio de Ambiente, Vivienda y Desarrollo Territorial) (MinAmbiente, 2017).

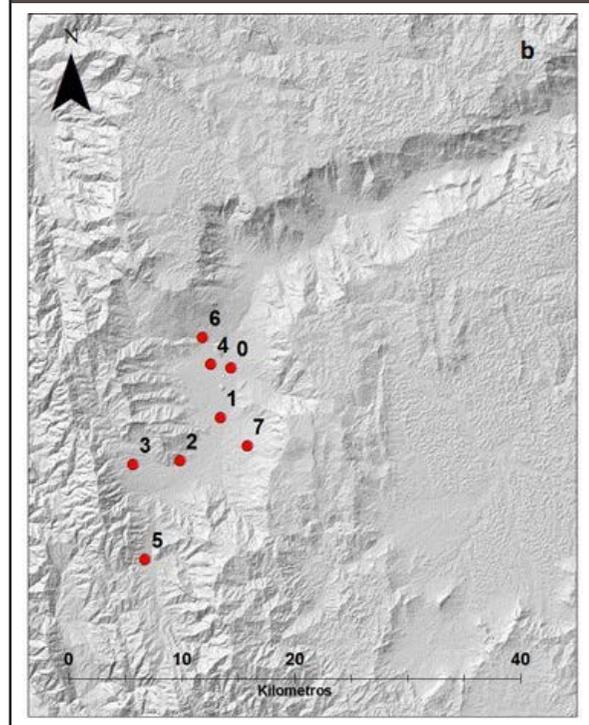
## 2. Área de estudio

La metodología se aplica en el área metropolitana de la ciudad de Medellín y se toma como variable de estudio la concentración promedio mensual de PM2.5 ( $\mu\text{g}/\text{m}^3$ ), disponible de la Red de Calidad del Aire (RedAire) en ocho sitios de monitoreo en

los años 2013 y 2014 (Sistema de Coordenadas Geográficas MAGNA Colombia Bogotá y proyección Transversade Mercator) (RedAire, 2015). Se tomaron como posibles variables explicativas a la temperatura ( $T$  en  $^{\circ}\text{C}$ ), la velocidad del viento ( $V_v$  en  $\text{m}/\text{s}$ ), el gradiente espacial de  $T$  ( $\nabla T$ ) y el gradiente espacial de  $V_v$  ( $\nabla V_v$ ). Los datos de las dos primeras del Sistema de Alertas Tempranas del Departamento de Antioquia (SIATA, 2016) y las dos últimas se calculan con técnicas de análisis espacial. Los datos fueron medidos en los sitios de monitoreo de la **Tabla 1**. La superficie de la región en donde se hicieron las mediciones es de  $535,13 \text{ km}^2$ , que corresponde a una densidad de muestreo de  $66,89 \text{ km}^2/\text{sitio de monitoreo}$ .

En la **Figura 1** se localizan los sitios de monitoreo. La metodología se implementa con las herramientas *Geostatistical Analyst*®, *GWR Analyst*® del aplicativo *ArcMap*® de ESRI para la caracterización espacial y con el software *EViews*® para la caracterización temporal.

**Figura 1.** Sitios de monitoreo de PM2.5.



**TABLA 1. DESCRIPCIÓN SITIOS DE MONITOREO DE PM2.5.**

ID del sitio de monitoreo	Xi	Yi	Descripción
0	834818,87	1183026,29	Museo de Antioquia
1	833901,20	1178597,43	Politécnico Colombiano Jaime Isaza Cadavid
2	830336,17	1174772,86	Casa de Justicia de Itagüí
3	826213,92	1174416,11	Concejo de Itagüí
4	833067,09	1183369,30	Universidad Nacional Núcleo Volador
5	827265,67	1165931,52	Corporación Universitaria Lasallista
6	832336,01	1185768,31	Universidad Nacional Facultad de Minas
7	836286,00	1176113,00	El Poblado

### 3. Metodología para la caracterización espacio-temporal de PM2.5

La metodología propuesta tiene las siguientes fases: primero se hace una caracterización espacial de PM2.5, con esta información y datos de series de tiempo se hace la caracterización temporal; finalmente se hace un ajuste al valor estimado de la concentración de PM2.5 por la proximidad a fuentes de emisión y de mitigación del contaminante.

#### 3.1. Caracterización espacial de PM2.5

En la **Figura 2** se muestra la metodología propuesta para hacer la caracterización espacial de PM2.5. En principio se valida si los datos de PM2.5 son estacionarios o no (si la función de distribución de densidad de los datos es normal) y si existen datos de otras variables que puedan afectar su concentración (variables explicativas). Dependiendo de lo anterior se usan diferentes algoritmos para realizar el proceso de caracterización.

Si no existen datos de variables explicativas, se hace una caracterización espacial univariada. Ésta tiene dos propósitos: hacer imputación espacial de datos faltantes y analizar cómo se comporta en el espacio la concentración de PM2.5 (mapa de estimación de datos). El procedimiento es el siguiente:

- Si los datos no son estacionarios, se usan algoritmos de interpolación espacial no geoestadísticos. Para el caso de PM2.5, el algoritmo del Inverso de la Distancia (Inverse DistanceWeighted – IDW) y el Spline; son los de mejor desempeño (Li & Heap, 2014).
- Si los datos son estacionarios, se usan algoritmos geoestadísticos (Kriging) (Londoño *et al.*, 2015).
- El desempeño de los algoritmos se calcula con pruebas de validación cruzada.

Si se tienen datos de variables explicativas, se hace una caracterización espacial multivariada. Se calcula el coeficiente de correlación entre la variable PM2.5 y cada una de las variables explicativas. Si hay independencia, se usan algoritmos geoestadísticos multivariados (CoKriging). Si hay dependencia, se usan algoritmos de regresión geográficamente ponderados (GWR). Para la caracterización espacial multivariada con CoKriging se usan los datos de la caracterización espacial univariada PM2.5 obtenidos mediante Kriging. El procedimiento para la caracterización espacial multivariada con GWR es el siguiente:

- Se revisan exhaustivamente las combinaciones de variables explicativas para determinar cuáles cumplen con los valores umbrales de los criterios de especificación del método de mínimos cuadrados ordinarios ( $\text{Adj}R^2 > 0.5$ , el mínimo

valor de AICc,  $(JB_{p-value}) > 0.1$ ,  $VIF < 7.5$ ,  $K(BP)_{p-value} < 0.05$ ; y  $(SA_{p-value}) > 0.1$  (Scott *et al.*, 2011).

- Con las combinaciones adecuadas de variables explicativas, se usa la herramienta *GWR*® del aplicativo *ArcMap*® de ESRI. El mejor modelo de regresión será aquel con mayor valor de  $AdjR^2$ , menor AICc y menor VIF.
- El desempeño del algoritmo se calcula con pruebas de validación cruzada.

Con base en lo anterior se obtiene el modelo con mejor desempeño que caracteriza espacialmente la concentración de PM2.5. Con este modelo se tienen datos en el dominio espacial con los cuales es posible calcular un mapa de estimación que muestra cómo se comporta la contaminación por PM2.5 en el área metropolitana de la ciudad. Estos datos se usan también como datos funcionales para la caracterización temporal de la contaminación.

### 3.2. Caracterización temporal de PM2.5.

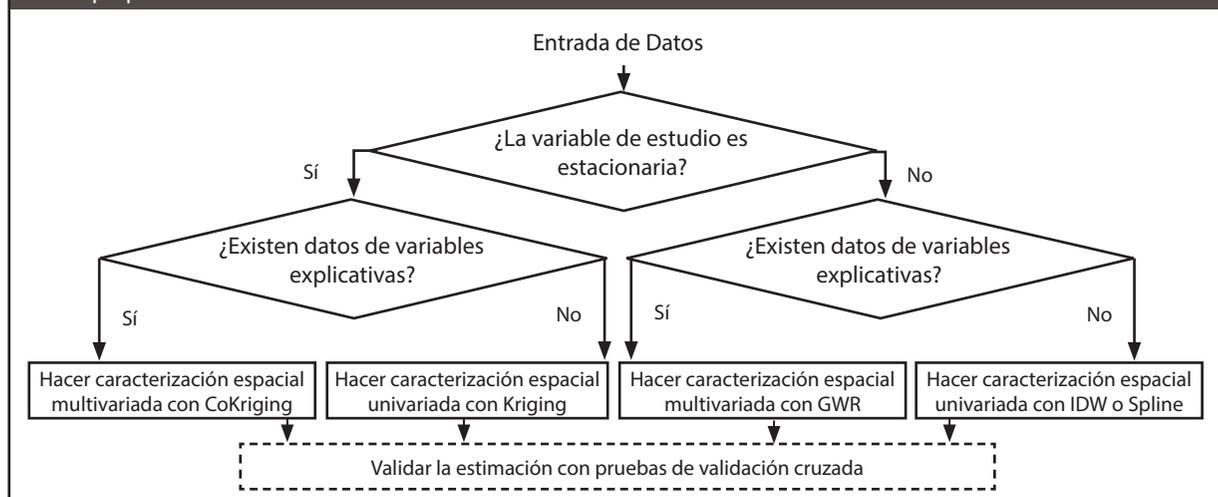
La caracterización temporal de PM2.5 se hace con algoritmos econométricos. Para ello los datos de las variables (PM2.5 y variables explicativas) se estructuran como paneles balanceados (Woolridge, 2002). A todas las series de tiempo de las

variables se le hacen pruebas de estacionariedad (prueba de raíces unitarias de Dickey-Fuller – PRU) (Dickey&Fuller, 1979) y pruebas de causalidad (prueba de causalidad de Wiener – Granger - PCG) (Granger, 1969). En la metodología propuesta se prueban algoritmos econométricos de rezagos distribuidos (RD), autorregresivos (AR) y mixtos (AR-RD). Los algoritmos AR y AR –RD han mostrado un buen desempeño para la estimación de la concentración de material particulado en zonas urbanas (Wang *et al.*, 2017), con valores de  $R^2$  que oscilan entre 0,38 y 0,98; e incorporan como variables explicativas: la velocidad del viento (Sayegh, 2014), la temperatura (Paschalidou, 2011), la densidad vehicular (para efectos de largo plazo) (Paschalidou, 2011), la concentración de CO, NO<sub>x</sub> y SO<sub>2</sub> (Sayegh, 2014), los usos del suelo, la presencia de zonas verdes (Taheri, 2015), los cuerpos de agua (Taheri, 2016), la densidad de población y variables meteorológicas (Chen, 2010). Los mejores modelos son aquellos cuya combinación de variables explicativas da un menor valor la suma de cuadrados de los residuos (SCR) y un mayor valor del coeficiente de determinación ajustado  $ADJR^2$ .

En la **Figura 3**, se muestra la metodología propuesta para la caracterización temporal.

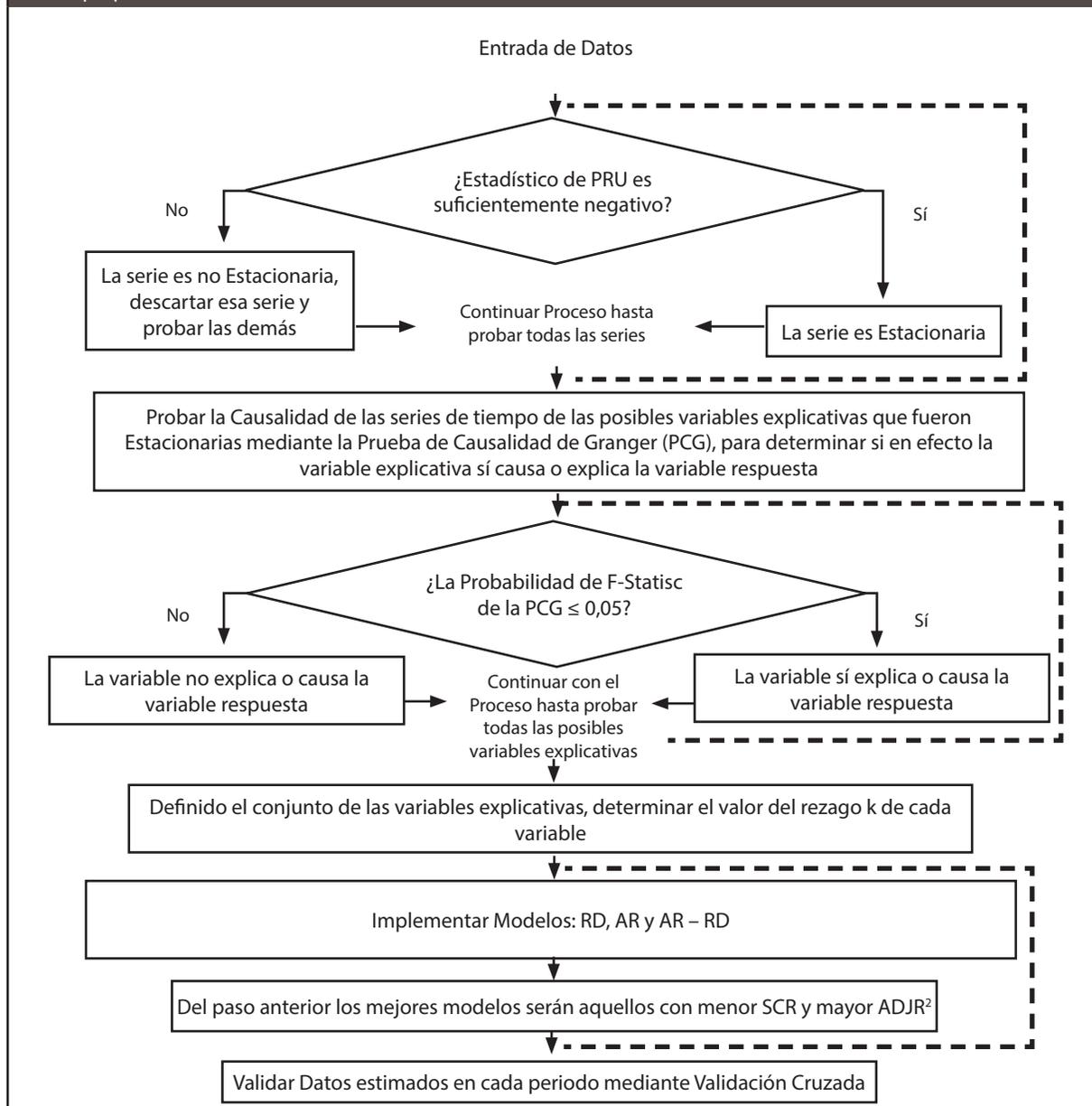
**Figura 2.** Caracterización Espacial.

Fuente propia.



**Figura 3.** Caracterización Temporal.

Fuente propia.



### 3.3. Ajuste del valor estimado de concentración de PM2.5 por proximidad a fuentes de emisión y de mitigación

Una vez hecha la caracterización espacio-temporal de la concentración de PM2.5, se implementa un Modelo de Proximidad Espacial (MPE) (Londoño *et al.*, 2017) para ajustarla incluyendo el efecto

de cercanía a las fuentes de emisión y de mitigación del contaminante. Cuando hay una aproximación a fuentes de emisión como las vías y las industrias, la concentración del contaminante aumenta y cuando hay una aproximación a las fuentes de mitigación como las zonas verdes, la concentración disminuye. El MPE calcula un factor de proximidad espacial

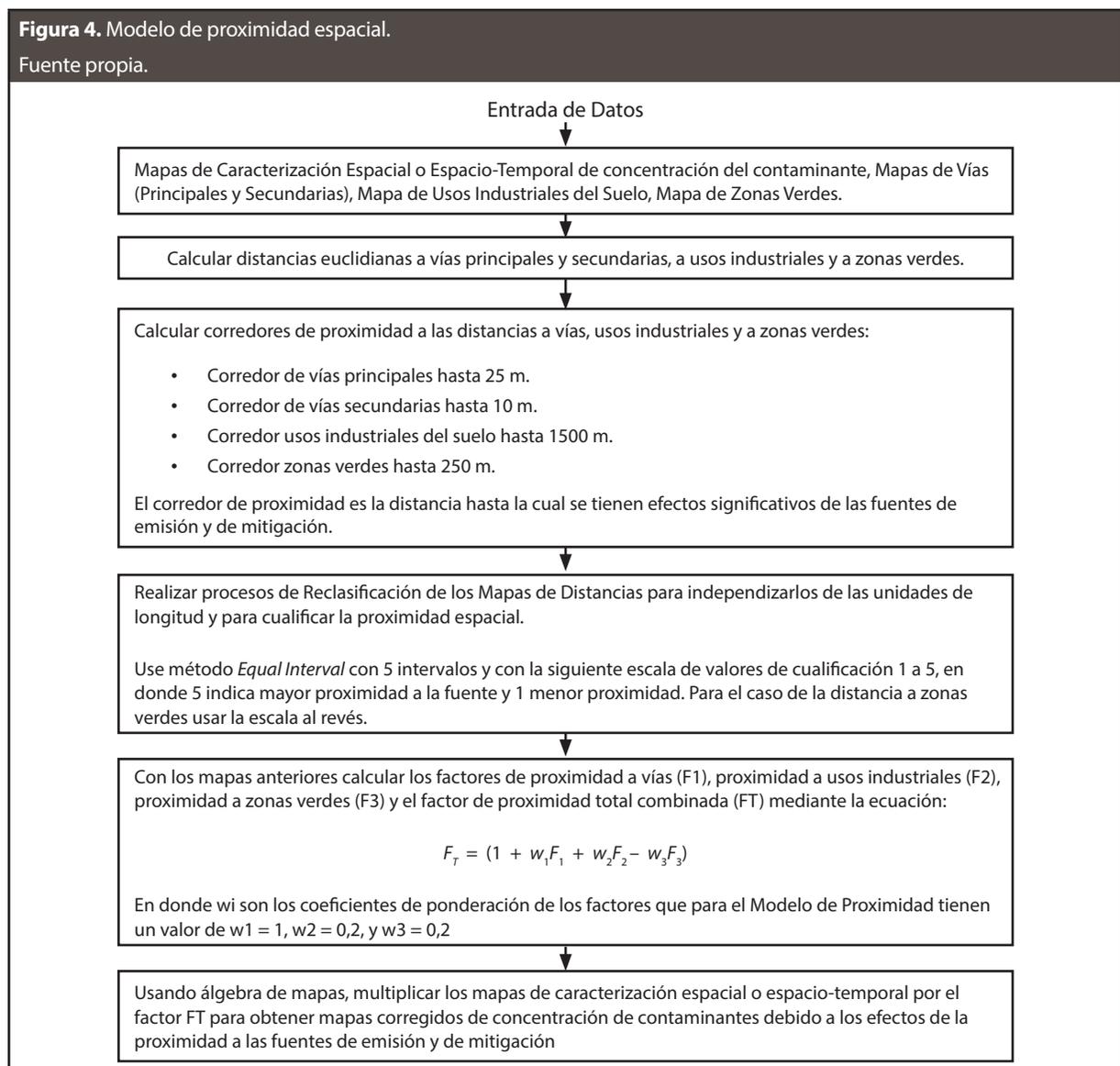
(FT) que contempla simultáneamente las cercanías a fuentes de emisión y de mitigación. Las entradas del MPE son los datos de caracterización espacio-temporal de concentración de PM2.5, las vías principales y secundarias de la ciudad (fuentes de emisión móvil de material particulado), las zonas de producción industrial (fuentes de emisión fijas) y las zonas verdes (fuentes de mitigación). El modelo se implementa mediante la herramienta *SpatialAnalyst*® del aplicativo *ArcMap*® de ESRI. En la **Figura 4**, se muestra el diagrama de flujo del MPE propuesto.

El FT, supone que el eje central de las vías principales contribuye con un 100% en el aumento de la concentración del material particulado, las zonas industriales contribuyen con un aumento del 20% y las zonas verdes contribuyen con una atenuación del 20%. En la **Figura 4**, el factor F1 está dado por la **Ecuación 1**.

$$F_1 = \frac{(\lambda_1 O_{R1} + \lambda_2 O_{R2})}{[\lambda_1 \text{Max}(O_{R1}) + \lambda_2 \text{Max}(O_{R2})]} \quad (1)$$

**Figura 4.** Modelo de proximidad espacial.

Fuente propia.



En la **Ecuación 1**,  $\lambda_1$  y  $\lambda_2$  son los pesos ponderados del mapa de reclasificación de la distancia a vías principales ( $O_{R1}$ ) y a vías secundarias ( $O_{R2}$ ), calculados mediante el método de análisis multicriterio del proceso jerárquico analítico AHP de Saaty (Gass & Fu, 2013). Siendo  $O_{R2}$  más importante que  $O_{R1}$ , los pesos calculados con el algoritmo AHP fueron  $\lambda_1 = 0,83$  y  $\lambda_2 = 0,17$ . La función  $Max(O_{Rj})$ , calcula el valor máximo de  $O_{Rj}$  tomando en cada celda una vecindad rectangular de  $3 \times 3$  celdas mediante funciones de estadística de bloque. Los factores  $F_2$  y  $F_3$  normalizan los mapas reclasificados de distancia a usos industriales y zonas verdes respectivamente y se calculan mediante la **Ecuación 2**, con  $j$  igual a 2 y 3, en donde  $O_{R3}$  es el mapa reclasificado de la distancia a usos industriales y  $O_{R4}$  es el mapa reclasificado de la distancia a zonas verdes.

$$F_j = 0,2 * O_{Rk} \quad k=3, 4. \quad (2)$$

#### 4. Resultados y análisis

El consolidado de los resultados obtenidos en la caracterización espacial y espacio-temporal de la concentración de PM2.5 se muestra en la **Tabla 2**, de donde se ve que los modelos de caracterización espacial explican poco su variabilidad. La explicación mejora significativamente con: los modelos de caracterización espacio-temporal AR-RD con rezago distribuido de orden 1, el dato funcional obtenido con GWR (que implícitamente contiene elementos de la caracterización espacial multivariada, la temperatura y el gradiente de la velocidad del viento), y como componente autorregresivo a la concentración de PM2.5 de rezago 1. Como se muestra en la **Tabla 3**, estas fueron las variables más importantes para mejorar la descripción de la variabilidad de la concentración de PM2.5.

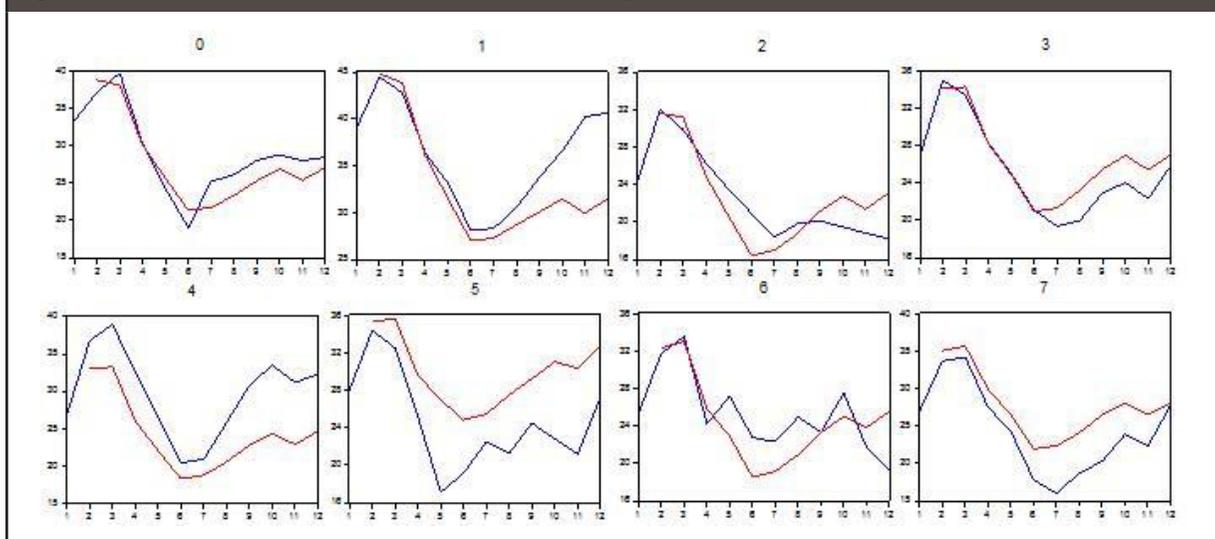
**TABLA 2. RESULTADOS DE LA CARACTERIZACIÓN ESPACIAL Y ESPACIO-TEMPORAL DE PM2.5**

Algoritmo	Variables	ADJR <sup>2</sup>	Estadísticos de Error por Mes			
<b>Caracterización Espacial Univariada</b>						
Geoestadístico Univariado (Semivariograma Teórico: Hole – Effect)			ASE	RMS	RMSS	MARE
Período 2013 - 2014	<b>PM2.5</b>	Min	0,90	1,80	0,90	0,05
Dm 66,89		Max	2,30	2,60	1,40	0,11
Área 535,13 km <sup>2</sup>		Media	1,79	2,22	1,04	0,08
8 Sitios Monitoreo (96 datos)		Desviación Estándar	0,39	0,28	0,14	0,02
<b>Caracterización Espacial Multivariada</b>						
Regresión Espacial Multivariada (GWR)	<b>T</b> (Fb, Ab, Jl). <b>T, Vv</b> (En, Mz).		ASE	RMS	RMSS	MARE
Período 2013 - 2014	$\nabla T$ (My, Ag, Nv). $\nabla Vv$ (Jn, SP, Oc). <b>Vv</b> (Dc).	0,16 Min	1,70	2,30	0,90	0,07
Dm 66,89		Max	3,30	3,70	1,10	0,14
Área 535,13 km <sup>2</sup>		Media	2,75	3,04	0,95	0,11
8 Sitios Monitoreo (96 datos)		Desviación Estándar	0,46	0,40	0,07	0,02
<b>Caracterización Espacio-Temporal</b>						
Econométrico (AR RD con componente espacial de rezago distribuido calculado mediante GWR)	<b>GWR</b> <sub>t-1</sub> <b>T</b> <sub>t-1</sub> $\nabla Vv$ <sub>t-1</sub> <b>PM2.5</b> <sub>t-1</sub>		ASE	RMS	RMSS	MARE
Período 2013 - 2014	<b>GWR</b> <sub>t-1</sub> <b>T</b> <sub>t-1</sub> $\nabla Vv$ <sub>t-1</sub> <b>PM2.5</b> <sub>t-1</sub>	0,87 Min	1,18	1,52	0,94	0,03
Dm 66,89		Max	3,90	5,05	0,94	0,16
Área 535,13 km <sup>2</sup>		Media	2,51	3,17	0,94	0,10
8 Sitios Monitoreo (96 datos)		Desviación Estándar	0,77	0,98	0,00	0,04

**TABLA 3. MEJORES MODELOS PARA LA CARACTERIZACIÓN ESPACIO-TEMPORAL DE PM2.5.**

Modelo	Variables Explicativas	ADJR <sup>2</sup>	SCR
RD	$T(-1), \nabla Vv(-1)$	0,76	497
AR	PM2.5(-1)	0,86	448
AR – RD	$T(-1), \nabla Vv(-1), PM2.5(-1)$	0,86	444
RD con GWR como variable explicativa	$GWR(-1), T(-1), \nabla Vv(-1)$	0,75	504
AR con GWR como variable explicativa	$GWR(-1), PM2.5(-1)$	0,86	431
AR-RD con GWR como variable explicativa	$GWR(-1), T(-1), \nabla Vv(-1), PM2.5(-1)$	0,87	404

**Figura 5. PM2.5 (O) en azul Vs PM2.5 (E) con AR – RD GWR en rojo.**



De la caracterización espacio - temporal de PM2.5 se puede decir que la concentración mensual y espacial de este contaminante depende de:

- La concentración en el espacio en el mes anterior. Este comportamiento se representa con la caracterización espacial multivariada hecha con el algoritmo GWR que incluye los efectos de  $T, \nabla T, Vv$  y  $\nabla Vv$  en cada mes (Tabla 2) y que define un componente espacial de rezago distribuido por medio de datos funcionales.
- El valor de  $T$  en el sitio en el mes anterior.
- El valor de  $\nabla Vv$  en el sitio en el mes anterior.
- El valor de la concentración de PM2.5 en el sitio en el mes anterior. Con base en los resultados de la Tabla 3, ésta es la variable explicativa más importante en la variabilidad de PM2.5

en el tiempo ya que, comparando los mejores modelos con el modelo AR, las demás variables explicativas solo aportan un 1% a la mejora del desempeño de los modelos.

En la Figura 5 se muestran, para cada sitio de monitoreo, las series de tiempo observadas de PM2.5 (O) vs los datos estimados de PM2.5 (E) con el modelo AR–RD de componente espacial de rezago distribuido GWR.

Los resultados de la Figura 5 se obtienen con la Ecuación 3.

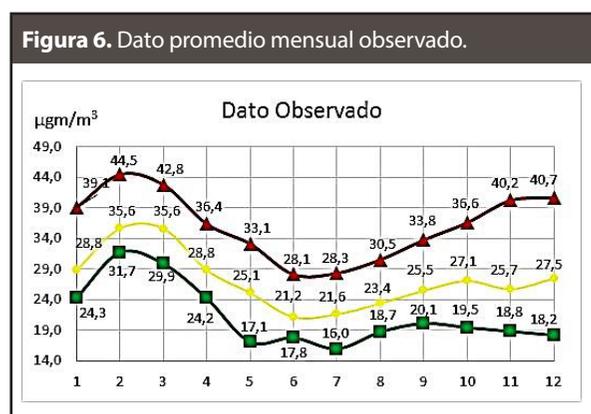
$$PM\ 2.5_t = 5.95 - 0.49GWR_{t-1} + 0.41T_{t-1} - 6.53 \nabla Vv_{t-1} + 0.85PM\ 2.5_{t-1} \quad (3)$$

En la Figura 6 se muestran los datos mensuales observados de PM2.5. En amarillo el valor pro-

medio, en verde el valor mínimo y en rojo el valor máximo. En ella se observa un comportamiento cíclico (estacional) en la concentración del contaminante (Liu *et al.*, 2017). En la curva amarilla se ve un punto de máxima entre febrero y marzo con valores de concentración entre 29,9 y 44,5  $\mu\text{g}/\text{m}^3$ , un punto de mínima entre junio y julio con valores de concentración entre 16,90 y 28,3  $\mu\text{g}/\text{m}^3$  y un punto promedio entre septiembre y octubre con valores de concentración entre 19,5 y 36,6  $\mu\text{g}/\text{m}^3$ . De allí se infiere que el período más contaminado en la ciudad se da entre febrero-marzo, el menos contaminado en junio-julio y el de contaminación cercana al valor promedio anual en septiembre-octubre.

Los valores mensuales estimados por cada uno de los algoritmos de caracterización espacial y espacio-temporal muestran un patrón similar a los datos observados en cuanto al mes de máxima y mínima concentración de PM2.5 (Figura 7).

Con relación a las estimaciones hechas en los sitios de monitoreo, en la Tabla 4 se muestran los estadísticos de error de estimación por sitio para el modelo AR-RD GWR. Con base en ellos se infiere que la estimación de mejor desempeño se da en el sitio de monitoreo 4 (Universidad Nacional Núcleo Volador) y la de menor desempeño en el 6 (Universidad Nacional Facultad de Minas).



Analizando los valores de la concentración promedio del contaminante en los sitios de monitoreo (Tabla 5), se infiere que la zona más contaminada de la ciudad está entre los sitios de monitoreo

0 (Museo de Antioquia) y 4 (Universidad Nacional Núcleo Volador), que la zona menos contaminada de la ciudad está entre los sitios de monitoreo 1 (Politécnico Colombiano Jaime Isaza Cadavid) y 7 (El Poblado), y que la zona con una contaminación tendiente al valor promedio anual está en los sitios de monitoreo 3, 5 y 6 (Corporación Universitaria La-sallista – Municipio de Caldas).

De las consideraciones anteriores se concluye que, entre febrero y marzo, en la zona entre las estaciones 0 y 4, la contaminación es crítica (esta zona corresponde a la parte central de la ciudad). En junio-julio, en la zona entre 1 y 7 la contaminación es la más baja (esta zona corresponde al sector del barrio El Poblado). En septiembre-octubre la contaminación tiende al valor promedio anual en la zona entre las estaciones 3 y 5 (que corresponde con el sur del área metropolitana), y en la estación 6 (Universidad Nacional Facultad de Minas).

En la Figura 8 se observa que las variables explicativas ( $T$ ,  $\nabla T$ ,  $Vv$  y  $\nabla Vv$ ), tienen un comportamiento mensual periódico. La  $T$  promedio mensual muestra un comportamiento simétrico, similar en el primer y el segundo semestre del año, con un máximo en el mes de julio. El  $\nabla T$  promedio mensual no es tan simétrico como  $T$ , con valores bajos entre enero y marzo y un aumento brusco hasta su valor máximo entre abril y junio para caer abruptamente entre julio y diciembre. El valor promedio mensual de  $Vv$  no es simétrico, muestra un descenso sostenido entre enero y mayo para subir rápidamente entre junio y agosto y descender nuevamente hasta casi su valor mínimo en noviembre. El  $\nabla Vv$  promedio mensual muestra un patrón similar al de  $Vv$ .

Para analizar la relación de cambio mensual de las variables explicativas con la concentración de PM2.5 se usa la siguiente convención:

- $\uparrow$ : la variable aumenta entre ambos meses.
- $\downarrow$ : la variable disminuye entre ambos meses.
- $\uparrow_{\text{Máximo}}$ : la variable tiene el máximo aumento entre ambos meses.

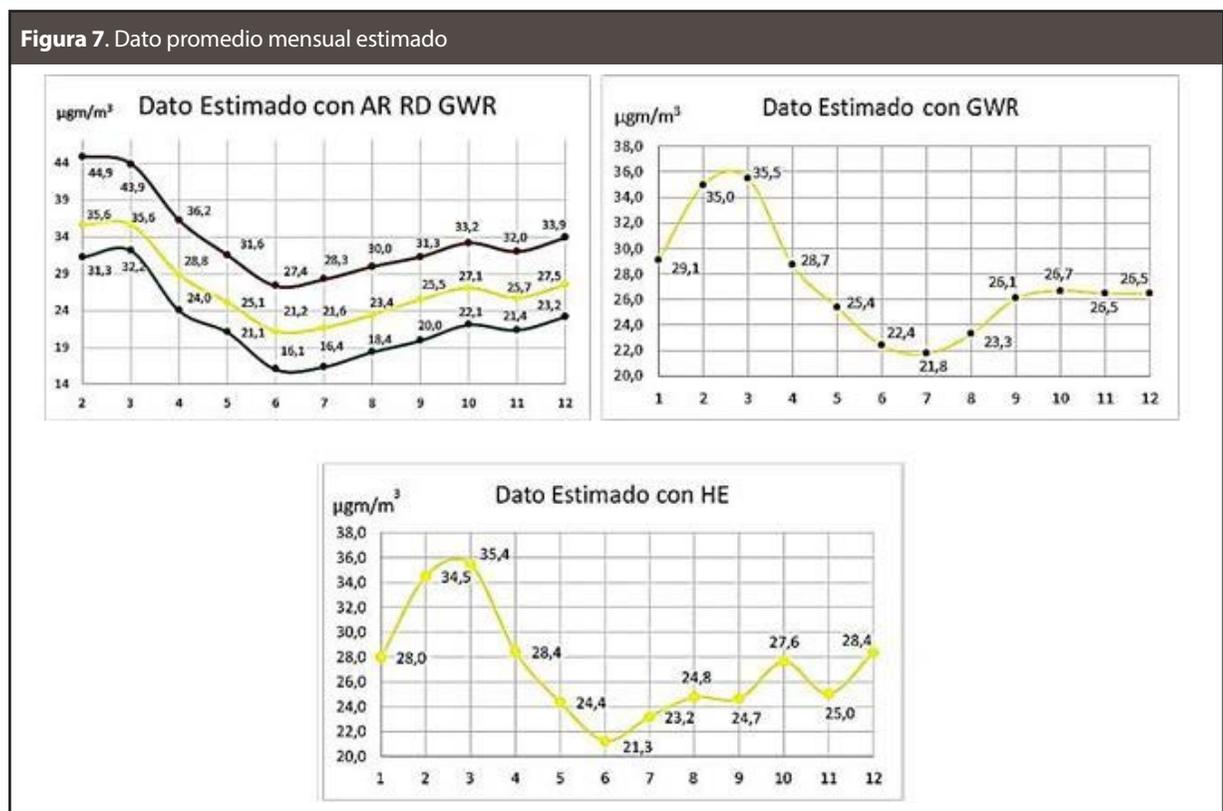
- $\uparrow_{\text{Mínimo}}$ : la variable tiene el mínimo aumento entre ambos meses.
- $\downarrow_{\text{Máximo}}$ : la variable tiene la máxima disminución entre ambos meses.
- $\downarrow_{\text{Mínimo}}$ : la variable tiene la mínima disminución entre ambos meses.
- $\leftrightarrow$ : la variable no cambia entre ambos meses.

Los resultados de las consideraciones anteriores se sintetizan en la **Tabla 6**.

Con base en los resultados de la **Tabla 6**, se encuentran los siguientes patrones:

- En el período más contaminado entre febrero y marzo, se da la mínima disminución de  $\nabla Vv$ , el mínimo aumento de  $T$  y casi el mínimo aumento de  $\nabla T$ , es decir, hay un cambio inverso de  $\nabla Vv$  con  $T$  y con  $\nabla T$ . Esto viene antecedido por la máxima disminución de  $\nabla Vv$  y ningún cambio en  $T$  entre los meses de enero y febrero. Lo anterior genera condiciones favorables para que en el mes de marzo se den los más altos valores en la concentración de PM2.5 en la ciudad. Esto no ocurre, por ahora, en los otros meses del año.

**Figura 7.** Dato promedio mensual estimado



**TABLA 4.** ESTADÍSTICOS DE ERROR DE ESTIMACIÓN DE PM2.5 POR SITIO DE MONITOREO.

Estadísticos de Error	0	1	2	3	4	5	6	7	Media	Dev.
ASE	2,49	2,44	4,21	1,52	1,52	2,71	3,14	2,38	2,55	0,87
RMS	3,58	2,81	4,62	1,77	1,72	3,57	3,98	3,19	3,16	1,02
RMSS	1,22	0,96	1,47	1,55	1,03	1,21	1,02	1,06	1,19	0,22
MARE	0,06	0,12	0,14	0,07	0,05	0,13	0,14	0,12	0,10	0,03

**TABLA 5. RESUMEN DE LOS VALORES ESTIMADOS DE PM2.5 POR SITIO**

Valores Sitio Monitoreo	0	1	2	3	4	5	6	7
Mínimo	27,4	16,1	19,6	21,5	21,2	20,7	19,8	22,6
Máximo	44,9	32,2	33,9	34,8	38,4	35,3	33,0	34,7
Promedio	<b>33,9</b>	<b>22,4</b>	<b>25,7</b>	<b>26,5</b>	<b>28,2</b>	<b>26,8</b>	<b>26,6</b>	<b>26,0</b>
Desviación estándar	5,8	5,3	4,6	4,6	5,6	4,9	4,4	4,4

**TABLA 6. COMPORTAMIENTO INTERMENSUAL DE LAS VARIABLES  $\nabla Vv$ ,  $\nabla T$  Y  $T$ .**

	En-Fb	Fb-Mz	Mz-Ab	Ab-My	My-Jn	Jn-Jl
$\nabla Vv$	↓ Máximo	↓ Mínimo	↓	↓	↑	↑ Máximo
$\nabla T$	↓	↑	↑ Máximo	↓	↑	↓ Máximo
$T$	↔	Mínimo	↑	↓	↑ Máximo	↑
	Jl-Ag	Ag-Sp	Sp-Oc	Oc-Nv	Nv-Dc	
$\nabla Vv$	↔	↓	↓	↔	↑ Mínimo	
$\nabla T$	↑	↓	↓	↓ Mínimo	↑ Mínimo	
$T$	↓ Máximo	↑	↓	↓ Mínimo	↑	

- En el período menos contaminado entre junio y julio, se da el máximo aumento de  $\nabla Vv$  y la máxima disminución de  $\nabla T$ , es decir, cambian inversamente. Esto está antecedido por el máximo aumento de  $T$  y un aumento de  $\nabla Vv$  entre los meses de mayo y junio. Lo anterior genera condiciones favorables en los meses de junio y julio para que se den los más bajos valores en la concentración de PM2.5 en la ciudad. Esto no ocurre nuevamente en los otros meses del año.
- Después del período menos contaminado, la contaminación empieza a aumentar hacia su valor promedio anual, entre septiembre y octubre. En este período  $\nabla Vv$  y  $\nabla T$  disminuyen, es decir, los gradientes cambian directamente. Esto está antecedido por la máxima disminución de  $T$  y ningún cambio de  $\nabla Vv$  entre los meses de julio y agosto. Lo anterior genera condiciones favorables en los meses de septiembre y octubre para que se den valores en la concentración de PM2.5 cercanos al valor promedio

anual. Esto no ocurre nuevamente, por ahora, en los otros meses del año, pero queda la duda de si en el futuro los niveles de contaminación subirán por encima del valor promedio anual.

Lo anterior plantea una situación que debe monitorearse a partir de la fecha. Según el valor promedio mensual de PM2.5 (curva amarilla de la **Figura 6**), la ciudad está contaminada de PM2.5 desde el inicio del año, alcanzando su valor más alto en marzo y disminuyendo en mayo a valores permisibles para continuar con la descontaminación en junio-julio y empezar nuevamente a subir hasta valores permisibles en septiembre para aumentar finalmente hasta el valor promedio anual medido de 27,1  $\mu\text{g}/\text{m}^3$  en diciembre.

Usando los resultados de la estimación espacio temporal de los datos de concentración de PM2.5 y la herramienta *SpatialAnalyst*® del aplicativo *ArcMap*® de ESRI, se calculan y se muestran en la **Figura 9** los mapas de caracterización espacio-temporal de PM2.5 en marzo (mes más contaminado), junio

(mes menos contaminado) y octubre (mes con valor tendiente al promedio anual) (se confrontan los mapas estimados y observados).

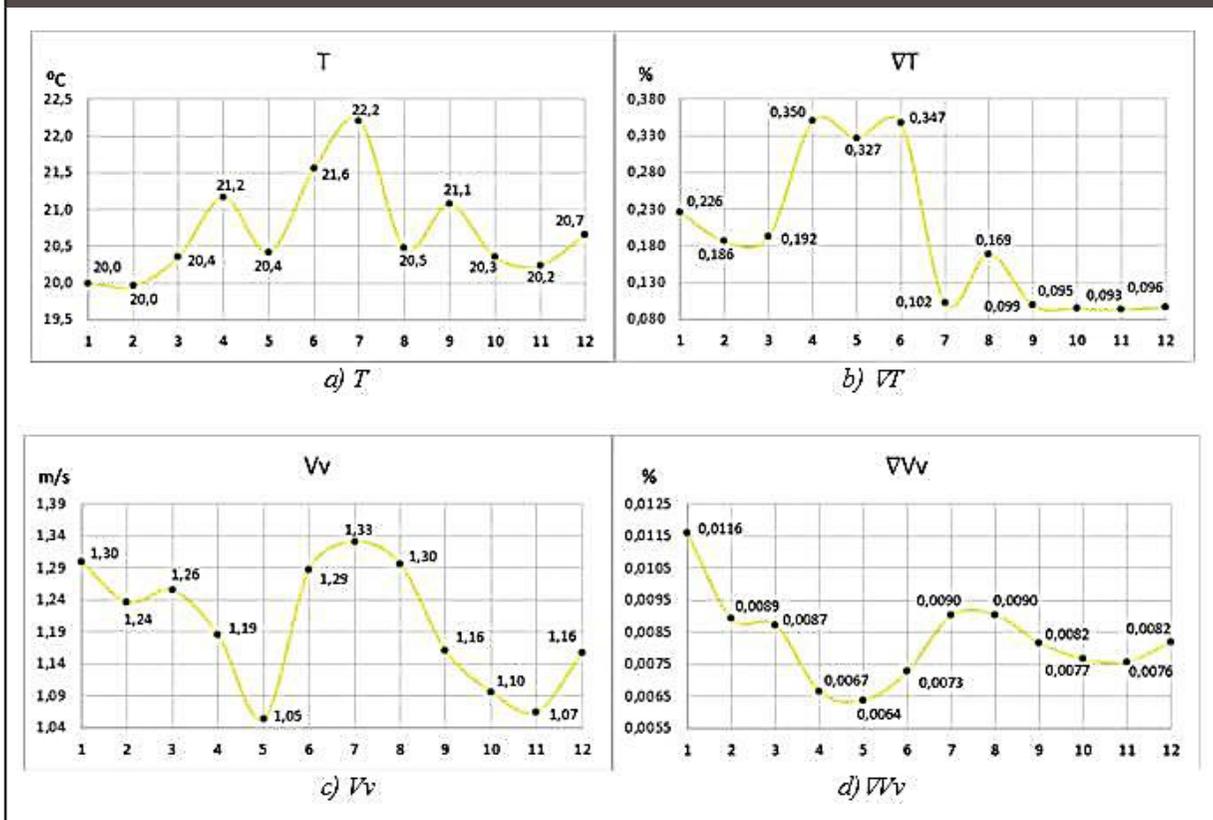
El mapa de marzo muestra un patrón similar al dato observado en el mismo mes. Las zonas más contaminadas de la ciudad están en la vecindad de los sitios de monitoreo 0 y 4. El sitio 0 corresponde a la parte céntrica de la ciudad de Medellín en donde se da una alta densidad de flujo vehicular. En general en marzo los valores de la concentración de PM2.5 están por encima de  $31 \mu\text{g}/\text{m}^3$  en toda la ciudad.

Junio y julio son los meses de menor contaminación en la ciudad. Según la **Tabla 6**, las zonas menos contaminadas de la ciudad están en la vecindad de los sitios de monitoreo 1 y 7. En el centro de la ciudad (estación 0) también se da una disminución importante de la contaminación ya que este período intersemestral coincide con la época de vacaciones

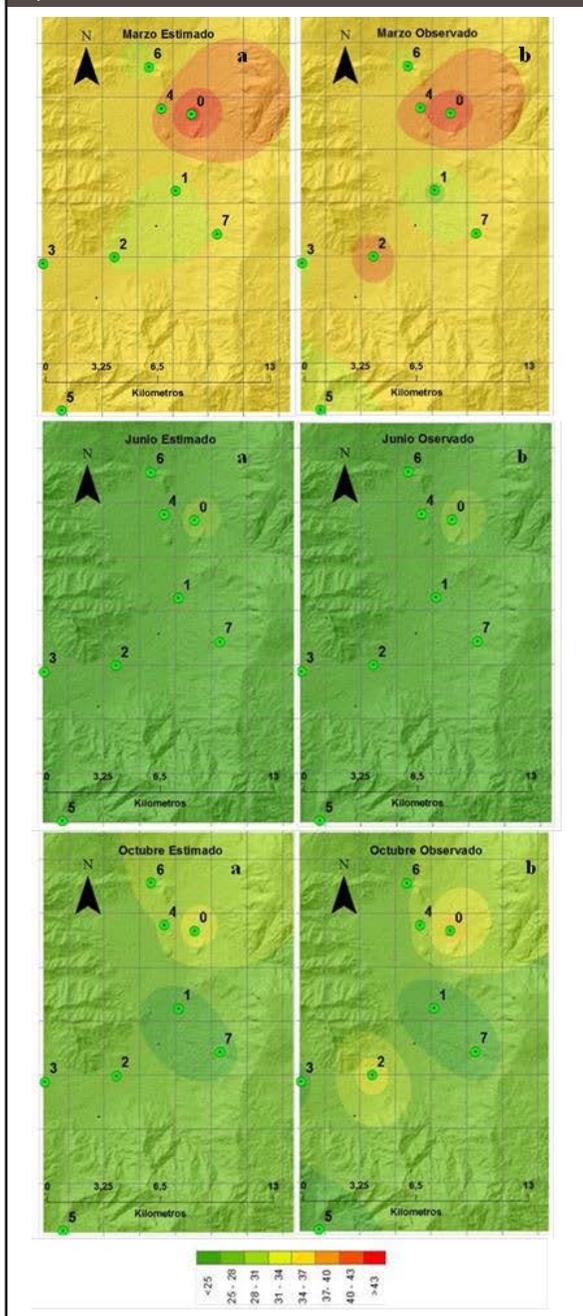
escolares y se da una disminución en la densidad de flujo vehicular. En general en junio los valores de la concentración de PM2.5 están por debajo de  $25 \mu\text{g}/\text{m}^3$  en toda la ciudad. En el mes de octubre el valor de la concentración tiende a aumentar al valor promedio anual y en general está por encima de  $28 \mu\text{g}/\text{m}^3$  en toda la ciudad.

Al aplicar el MPE a los mapas de caracterización espacio-temporal—siguiendo el procedimiento expuesto en Londoño *et al.* (2017) para PM10—y agregar las variables de proximidad a las fuentes de emisión y de mitigación, se obtienen detalles de la caracterización espacio-temporal de la concentración de PM2.5 asociados a la cercanía a las vías principales, a los usos industriales y a las zonas verdes, que se muestran en la **Figura 10** para todos los meses del año.

**Figura 8.** Valores promedio mensual de las variables T, VT, Vv y  $\nabla Vv$



**Figura 9.** Mapas de caracterización espacio-temporal de la concentración de PM<sub>2.5</sub> (µg/m<sup>3</sup>) en tres meses representativos del año.



## 5. Conclusiones

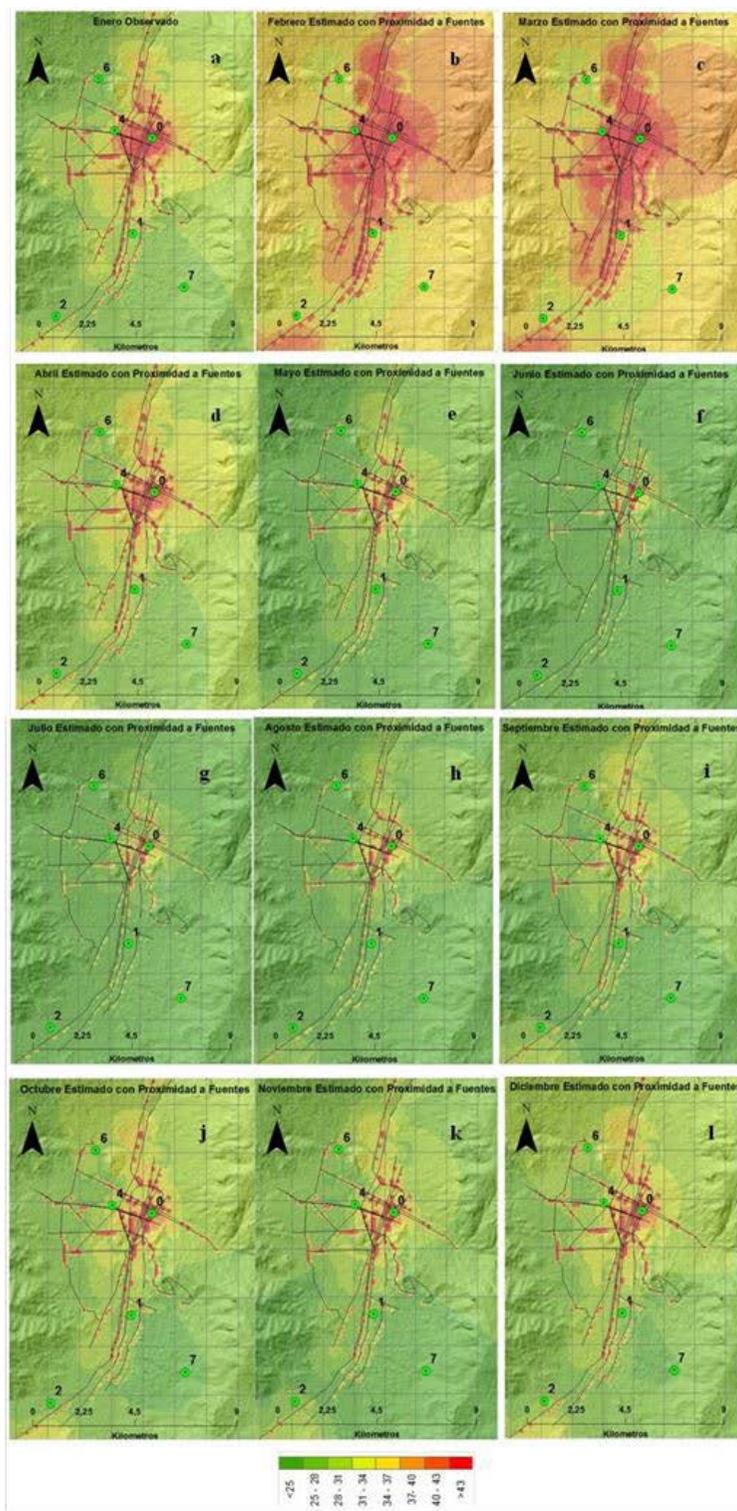
Con la caracterización espacial univariada se encuentra que el modelo *HoleEffect* es el que tiene mejor desempeño para PM<sub>2.5</sub>. En la caracterización

espacial multivariada, el algoritmo GWR mostró buen desempeño en la estimación de la concentración. Se encontró que las variables explicativas fueron la velocidad del viento y la temperatura, siendo muy importantes los gradientes espaciales de estas variables. Este hallazgo es un aporte toda vez que en la literatura no se reporta su aplicación. Para la caracterización espacio-temporal, se encontró que el modelo de mejor desempeño fue un AR-RD de rezago 1 con GWR como variable explicativa que logró explicar la variabilidad espacio-temporal de la concentración de PM<sub>2.5</sub> en un 87%. Las variables explicativas más significativas fueron: el dominio espacial calculado con GWR, el gradiente de la velocidad del viento, la temperatura y la concentración del material particulado. Al comparar los resultados de los modelos con el modelo AR, la variable más importante fue el rezago 1 de la concentración del material particulado. Las demás variables sólo aportaron un 1% de más a la variabilidad.

La concentración mensual de PM<sub>2.5</sub> permite destacar tres épocas de contaminación en el año: la de mayor impacto entre febrero y marzo, la de menor impacto entre junio y julio y la condición promedio entre septiembre y octubre. Este patrón se encontró a partir de estimaciones hechas con datos de 2013-2014, pero recientemente en los últimos tres años (2015 a 2017) se ha venido repitiendo el episodio de contaminación máxima en marzo a valores de muy alta contaminación en la ciudad con los efectos inmediatos sobre la salud de sus habitantes. Igualmente, las mediciones recientes muestran que junio es el mes menos contaminado.

Con la metodología se pudo determinar que, durante el episodio de mayor contaminación, los gradientes espaciales cambian mínimamente y de manera inversa:  $\nabla V_v$  disminuye y  $\nabla T$  aumenta. En otras palabras, en marzo no hay cambios significativos en  $\nabla V_v$  ni en  $\nabla T$  y, como estas variables afectan fuertemente la dispersión del material particulado, se favorece el aumento significativo de la contaminación.

**Figura 10.** Mapas mensuales de caracterización espacio-temporal de la concentración de PM2.5 ( $\mu\text{g}/\text{m}^3$ ) con ajuste de proximidad a fuentes.



Se encontró también que, entre enero y febrero se da la máxima disminución de  $\nabla Vv$  sin cambios significativos en la temperatura. Esto no se encontró en los cambios intermensuales de estas variables en los demás meses del año. En el período menos contaminado los gradientes nuevamente cambian de manera inversa, pero en sus valores máximos: se da el máximo aumento en  $\nabla Vv$  y la máxima disminución en  $\nabla T$ . En otras palabras, en junio se dan los máximos cambios en  $\nabla Vv$  y en  $\nabla T$  lo que favorece la dispersión del contaminante. Adicionalmente, entre abril y mayo se da el máximo aumento de temperatura y un aumento de  $\nabla Vv$ . Estas singularidades tampoco se encontraron en los cambios intermensuales de estas variables en los demás meses del año. En el segundo semestre las concentraciones de PM2.5 vuelven a incrementarse gradualmente hasta que en octubre se da otro episodio destacado en el que se alcanza un valor de concentración similar al promedio anual. En este mes, a diferencia de los casos anteriores, los gradientes cambian directamente ( $\nabla Vv$  y  $\nabla T$  disminuyen simultáneamente). En otras palabras, en octubre los gradientes empiezan nuevamente a bajar. Entre agosto y septiembre se da el máximo aumento de  $T$  y un aumento de  $\nabla Vv$ . Al igual que en los casos anteriores, tampoco se encontraron estos cambios intermensuales de las variables en los demás meses del año.

Por lo anterior se recomienda hacer un seguimiento del comportamiento de las variables  $\nabla Vv$  y  $\nabla T$ , vigilando si el patrón detectado se sigue repitiendo antes de los episodios destacados. Si así fuese, se tendrían criterios importantes de decisión para la implementación oportuna de medidas preventivas y efectivas que apunten al mejoramiento de la calidad del aire de la ciudad, teniendo en cuenta la estacionalidad de los episodios destacados detectados. Se recomienda igualmente que se midan temperatura y velocidad del viento simultáneamente a la medición de la concentración del material particulado en los sitios de monitoreo para mejorar las estimaciones (actualmente no se miden estas variables de manera simultánea en los sitios de monitoreo).

La concentración de PM2.5 en Medellín está por encima del valor máximo promedio anual permisible de  $25 \mu\text{g}/\text{m}^3$  durante la mayor parte del año. Es imperativo aplicar medidas urgentes para resolver esta situación, que ya se ha vuelto recurrente en los últimos tres años. De no ser así, existe la tendencia a futuro de que octubre tenga episodios de alta contaminación como los de marzo y que muy probablemente se afecte negativamente la baja contaminación de junio.

Otro aporte importante de la metodología es el modelo de proximidad espacial (MPE) que permite hacer ajustes a las estimaciones de la caracterización espacial-temporal de la concentración de PM2.5, al incluir la variable de proximidad a fuentes de emisión y de mitigación. Finalmente, se destaca que los resultados de esta metodología se pueden obtener de manera rápida y económica, como complemento de algoritmos determinísticos que requieren de los datos distribuidos de muchas variables para su implementación.

## Agradecimientos

Los autores agradecen al programa de Doctorado en Ingeniería de la Facultad de Ingeniería de la Universidad de Antioquia, en cuyo marco se desarrolló la presente investigación; a los estudiantes de la Maestría en Geoinformática de la Universidad de San Buenaventura (Cohorte 1), de la Especialización en Sistemas de Información Geográfica de la Universidad de San Buenaventura (Cohortes 19 y 20) y de la Especialización en Medio Ambiente y Geoinformática de la Universidad de Antioquia (Cohorte 16); por el apoyo recibido en el procesamiento previo de los datos y la aplicación de la metodología.

## Referencias

- Alam, M., & McNabola, A. (2015). Exploring the modeling of spatiotemporal variations in ambient air pollution within the land use regression framework: Estimation of PM10 concentrations on a daily basis. *J. Air Waste Manag. Assoc.*, 65, 628–640.

- Chen, L., Bai, Z., Kong, S., Han, B., You, Y., Ding, X., y otros. (2010). A land use regression for predicting NO<sub>2</sub> and PM<sub>10</sub> concentrations in different seasons in Tianjin region, China. *J. Environ. Sci*, 22, 1364–1373.
- Deligiorgi, D., & Philippopoulos, K. (2011). *Spatial Interpolation Methodologies in Urban Air Pollution Modeling: Application for the Greater Area of Metropolitan Athens, Greece, Advanced Air Pollution*. Athens: Dr. Farhad Nejadkoorki.
- Dickey, D., & Fuller, W. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, 427-431.
- Diebold, F. (1971). *Elements of Forecasting*, 2a. ed. South Western.
- Dons, E., Van Poppe, M., Panis, L., De Prins, S., Berghmans, P., Koppen, G., y otros. (2014). Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. *Science of The Total Environment*, 476–477, 378-386.
- Fotheringham, A., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: John Wiley & Sons, Ltd.
- Frees E. W. *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*, Cambridge University Press, Nueva York, 2004.
- Gass, S., & Fu, M. (2013). Saaty, T. *Analytic Hierarchy Process Encyclopedia of Operations Research and Management Science*. Springer.
- Goyal, P., Chan, A., & Jaiswal, N. (2006). Statistical models for the prediction of respirable suspended particulate matter in urban cities. *Atmos. Environ*, 40, 2068–2077.
- Granger, C. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 424-438.
- Lee, M., Brauer, M., Wong, P., Tang, R., Tsui, T., Choi, C., y otros. (2017). Land use regression modelling of air pollution in high density high rise cities: A case study in Hong Kong. *Science of the Total Environment*, 592, 306–315.
- Li, J., & Heap, D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, 53, 173 – 189.
- Liu, B., Wu, J., Zhang, J., Wang, L., Yang, J., Liang, D., y otros. (2017). Characterization and source apportionment of PM<sub>2.5</sub> based on error estimation from EPA PMF 5.0 model at a medium city in China. *Environmental Pollution*, 222, 10-22.
- Londoño, L., & Cañón, J. (2015). Metodología para la aplicación de modelos de regresión de usos del suelo en la estimación local de la concentración mensual de pm<sub>10</sub> en Medellín – Colombia. *Revista Politécnica*, 11(21), 29-40.
- Londoño, L., Cañón, J., & Giraldo, J. (2017). Modelo de proximidad espacial para definir sitios de muestreo en redes urbanas de calidad de aire. *Revista Facultad Nacional de Salud Pública*, 35(1), 111-122.
- Londoño, L., Cañón, J., Villada, R., & López, L. (2015). Caracterización espacial de PM<sub>10</sub> en la ciudad de Medellín mediante modelos geoestadísticos. *Revista Ingenierías USBMED*, 6(2), 26-35.
- Alam, M., & McNabola, A. (2015). Exploring the modeling of spatiotemporal variations in ambient air pollution within the land use regression framework: Estimation of PM<sub>10</sub> concentrations on a daily basis. *J. Air Waste Manag. Assoc*, 65, 628–640.
- Chen, L., Bai, Z., Kong, S., Han, B., You, Y., Ding, X., y otros. (2010). A land use regression for predicting NO<sub>2</sub> and PM<sub>10</sub> concentrations in different seasons in Tianjin region, China. *J. Environ. Sci*, 22, 1364–1373.
- Deligiorgi, D., & Philippopoulos, K. (2011). *Spatial Interpolation Methodologies in Urban Air Pollution Modeling: Application for the Greater Area of Metropolitan Athens, Greece, Advanced Air Pollution*. Athens: Dr. Farhad Nejadkoorki.
- Dickey, D., & Fuller, W. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, 427-431.
- Diebold, F. (1971). *Elements of Forecasting*, 2a. ed. South Western.
- Dons, E., Van Poppe, M., Panis, L., De Prins, S., Berghmans, P., Koppen, G., y otros. (2014). Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. *Science of The Total Environment*, 476–477, 378-386.
- Fotheringham, A., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: John Wiley & Sons, Ltd.

- Frees E. W. Longitudinal and Panel Data: Analysis and Applications in the Social Sciences, Cambridge University Press, Nueva York, 2004.
- Gass, S., & Fu, M. (2013). Saaty, T. Analytic Hierarchy Process Encyclopedia of Operations Research and Management Science. Springer.
- Goyal, P., Chan, A., & Jaiswal, N. (2006). Statistical models for the prediction of respirable suspended particulate matter in urban cities. *Atmos. Environ*, 40, 2068–2077.
- Granger, C. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 424-438.
- Lee, M., Brauer, M., Wong, P., Tang, R., Tsui, T., Choi, C., y otros. (2017). Land use regression modelling of air pollution in high density high rise cities: A case study in Hong Kong. *Science of the Total Environment*, 592, 306–315.
- Li, J., & Heap, D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, 53, 173 – 189.
- Liu, B., Wu, J., Zhang, J., Wang, L., Yang, J., Liang, D., y otros. (2017). Characterization and source apportionment of PM<sub>2.5</sub> based on error estimation from EPA PMF 5.0 model at a medium city in China. *Environmental Pollution*, 222, 10-22.
- Londoño, L., & Cañón, J. (2015). Metodología para la aplicación de modelos de regresión de usos del suelo en la estimación local de la concentración mensual de pm<sub>10</sub> en Medellín – Colombia. *Revista Politécnica*, 11(21), 29-40.
- Londoño, L., Cañón, J., & Giraldo, J. (2017). Modelo de proximidad espacial para definir sitios de muestreo en redes urbanas de calidad de aire. *Revista Facultad Nacional de Salud Pública*, 35(1), 111-122.
- Londoño, L., Cañón, J., Villada, R., & López, L. (2015). Caracterización espacial de PM<sub>10</sub> en la ciudad de Medellín mediante modelos geoestadísticos. *Revista Ingenierías USBMED*, 6(2), 26-35.
- MinAmbiente (2017). Contaminación Atmosférica en Colombia. Recuperado el 23 de agosto de 2017, de Ministerio de Ambiente y Desarrollo Sostenible: <http://www.minambiente.gov.co/index.php/component/content/article/1801-plantilla-#1-1-normativa>
- Paschalidou, A., Karakitsios, S., Kleanthous, S., & Kassomenos, P. (2011). Forecasting hourly PM<sub>10</sub> concentration in Cyprus through artificial neural networks and multiple regression models: Implications to local environmental management. *Environ. Sci. Pollut. Res*, 18, 316-327.
- Pilsung, K. (2013). Locally linear reconstruction based missing value imputation for supervised learning. *Neurocomputing*, 118, 65-78.
- Ramsay, J., Ramsay, T., & Sangalli, L. (2011). Spatial functional data analysis. *Recent Advances in Functional Data Analysis and related topics*. (págs. 269-275). Springer.
- RedAire. (2015). Laboratorio de Calidad del Aire (CALAIRE). Universidad Nacional de Colombia Sede Medellín, Facultad de Minas. Medellín, Colombia.
- Şahin, Ü., Bayat, C., & Uçan, O. (2011). Application of cellular neural network (CNN) to the prediction of missing air pollutant data. *Atmospheric Research*, 101(1–2), 314-326.
- Sangalli, L., Ramsay, J., & Ramsay, T. (2013). Spatial splines regressions models. *J. Roy. Statisc. Soc. Ser. B*, 75, 681-803.
- Sanhueza, P., Torreblanca, M., Diaz-Robles, L., Schiappacasse, L., Silva, M., & Astete, T. (2009). Particulate air pollution and health effects for cardiovascular and respiratory causes in Temuco, Chile: A wood-smoke-polluted urban area. *J. Air Waste Manag. Assoc.*, 59, 1481-1488.
- Sayegh, A., Munir, S., & Habeebullah, T. (2014). Comparing the performance of statistical models for predicting PM<sub>10</sub> concentrations. *Aerosol and Air Quality Research*, 14(3), 653 - 665.
- Scott, L., Rosenshein, L., & Janikas, M. (2011). Modeling Spatial Relationships using Regression Analysis. ESRI International User Conference. Technical Workshops. San Diego, CA.
- Shahraiyini, H., & Sodoudi, S. (2016). Statistical Modeling Approaches for PM<sub>10</sub> Prediction in Urban Areas; A Review of 21st-Century Studies. *Atmosphere*, 7(2) (15).
- SIATA. (2016). Sistema de Alerta Temprana de Medellín y el Valle de Aburrá. Sistema de Alerta Temprana de Medellín y el Valle de Aburrá. Medellín, Colombia.
- Singh, N., Murari, V., Kumar, M., Barman, S., & Banerjee, T. (2017). Fine particulates over South Asia: Review and meta-analysis of PM<sub>2.5</sub> source apportionment through receptor model. *Environmental Pollution*, 223, 121-136.

- Stadlober, E., Hörmann, S., & Pfeiler, B. (2008). Quality and performance of a PM10 daily forecasting model. *Atmos. Environ*, 42, 1098–1109.
- TaheriShahraiyini, H., Sodoudi, S., Cubasch, U., & Kerschbaumer, A. (2015). The influence of the plants on the decrease of air pollutants (Case study: Particulate matter in Berlin). In Presented at the Euro-American Conference for Academic Disciplines. Paris, France.
- Taheri, H., & Sodoudi, S. (2016). Statistical Modeling Approaches for PM10 Prediction in Urban Areas; A Review of 21st-Century Studies. *Atmosphere*, 7(15).
- Wang, P., Zhang, H., Qin, Z., & Zhang, G. (2017). A novel hybrid-Garch model based on ARIMA and SVM for PM2.5 concentrations forecasting. *Atmospheric Pollution Research*, In Press, Corrected Proof.
- Wiener, N. (1956). *The Theory of Prediction*. Modern Mathematics for Engineers. New York: McGraw-Hill.
- Woolridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Massachusetts: MIT Press.
- Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175.

**PARA CITAR ESTE ARTÍCULO /  
TO REFERENCE THIS ARTICLE /  
PARA CITAR ESTE ARTIGO /**

Londoño-Ciro, L.A.; Cañón-Barriga, J.E. (2018). Metodología para la caracterización espacio-temporal de PM2.5 en el área urbana de la ciudad de Medellín-Colombia. *Revista EIA*, 15(30), julio-diciembre, pp. 113-132. [Online]. Disponible en: <https://doi.org/10.24050/reia.v15i30.1217>