

Análisis de la contaminación ambiental usando técnicas de teledetección y análisis de componentes principales

Analysis of Environmental Pollution Using Remote Sensing Techniques and Principal Component Analysis

  Lizette Tello-Cifuentes¹;
 Jean P. Díaz-Paz²

¹Universidad del Valle, Cali-Colombia,
lizette.tello@correounivalle.edu.co

²Institución Universitaria Antonio José Camacho, Cali-Colombia,
jpdiaz@profesores.uniajc.edu.co

Cómo citar / How to cite

L. Tello-Cifuentes; J. P. Díaz-Paz, “Análisis de la contaminación ambiental usando técnicas de teledetección y análisis componentes principales”, *Tecnológicas*, vol. 24, nro. 50, e1710, 2021.
<https://doi.org/10.22430/22565337.1710>

ISSN-p: 0123-7799
ISSN-e: 2256-5337

Vol. 24, nro. 50, e1710

Recibido: 29 junio 2020
Aceptado: 16 septiembre 2020
Disponible: 10 noviembre 2020

©Instituto Tecnológico Metropolitano
Este trabajo está licenciado bajo
una Licencia Internacional
Creative Commons Atribución
(CC BY-NC-SA)



Resumen

En los últimos años, uno de los principales problemas en Colombia es la contaminación ambiental debido al crecimiento de la población y la industrialización. La contaminación afecta a las grandes ciudades, perjudicando el medio ambiente y la salud humana. Es por ello que se presenta una metodología para el análisis de la contaminación ambiental en Medellín, usando técnicas de percepción remota, imágenes Landsat 7 y 8 y variables de calidad del aire. La metodología está compuesta por cuatro etapas: i. Preprocesamiento de imágenes, ii. Procesamiento de imágenes, cálculo de los índices de temperatura de la superficie (*TS*), índice de vegetación normalizado (*NDVI*), índice de vegetación ajustado al suelo transformado (*TSAVI*), índice de diferencia normalizada del agua (*NDWI*) y el índice normalizado del suelo (*NSI*), iii. Interpolación de variables de calidad del aire, Material Particulado (*PM₁₀*), Material Particulado (*PM_{2.5}*), Dióxido de Nitrógeno (*NO₂*) y Ozono (*O₃*); y, iv. Análisis de componentes principales. A partir de las técnicas aplicadas combinadas con la estimación del primer componente principal, el cual contenía el 90 % de variación de la información, se obtuvo un mapa de calidad ambiental que permitió identificar que los focos de contaminación se presentan en sectores con poca cobertura de vegetación, gran cantidad de construcciones y gran flujo vehicular. Por otro lado, las zonas con calidad de aire bueno son sectores que presentan una mayor cobertura de vegetación, por lo general ubicados en los extremos de la ciudad y de estrato socioeconómico 4, 5 y 6. Este mapa puede ser insumo en la oportuna toma de decisiones en cuanto a la planificación urbana, ya que permite la pronta intervención en las zonas donde la calidad ambiental es deficiente.

Palabras clave

Percepción remota, contaminación atmosférica, calidad del aire, procesamiento de imágenes, análisis de componentes principales.

Abstract

Due to population growth and industrialization, air pollution has become one of Colombia's major issues in recent years. It affects large cities, harming the environment and human health. Therefore, in this paper, we propose a methodology to analyze air pollution in Medellín, Colombia using remote sensing techniques, Landsat-7 and Landsat-8 images, and air quality variables. The proposed methodology consists of four stages: (i) image preprocessing; (ii) image processing and calculation of the Surface Temperature (*TS*), Normalized Difference Vegetation Index (*NDVI*), Transformed Soil Adjusted Vegetation Index (*TSAVI*), Normalized Difference Water Index (*NDWI*), and Normalized Difference Soil Index (*NSI*); (iii) interpolation of air quality variables, Particulate Matter (*PM₁₀* and *PM_{2.5}*), Nitrogen Dioxide (*NO₂*), and Ozone (*O₃*); and (iv) principal component analysis. Based on the applied techniques, together with the estimation of the first major component (which contains 90% of information variation), an air quality map is obtained. According to this map, the sources of pollution are found in sectors with little vegetation cover, a great number of buildings, and high traffic flow. Conversely, areas with good air quality include sectors with greater vegetation cover, which are usually found in the limits of the city and in socioeconomic strata 4, 5, and 6. This map could be used as input for timely decision-making regarding urban planning because it allows for an early intervention in areas with poor air quality.

Keywords

Remote sensing, air pollution, air quality, image processing, principal component analysis.

1. INTRODUCCIÓN

La contaminación atmosférica se ha convertido en uno de los principales problemas en Colombia debido al deterioro continuo en la calidad del aire y el impacto sobre la salud humana y el medio ambiente [1]. Informes de diferentes instituciones públicas y privadas de vigilancia de la calidad del aire confirman el incremento en las concentraciones de contaminantes ambientales que superan lo permitido en la norma nacional de calidad del aire [2]. La concentración de contaminantes se ve afectada por factores topográficos y meteorológicos como: depresiones por altas montañas que impidan que circule el aire, precipitaciones, formación de neblina, reducción de la radiación solar y alteración de la temperatura entre otros [3]. Este contexto muestra la necesidad de usar herramientas de análisis espacial de contaminación ambiental como apoyo a las decisiones administrativas de Colombia y el mundo.

El análisis de la distribución espacial de la contaminación atmosférica se ha abordado mediante redes de vigilancia [4] y sensores remotos. Las metodologías basadas en sensores remotos incluyen las siguientes etapas: preprocesamiento, en la cual se disminuyen los efectos de ruido y se corrigen radiométrica y geoméricamente las imágenes; una etapa de extracción de características, en la que se estiman las variables o índices de interés a partir de las imágenes con el fin de discriminar áreas con características similares [5]–[7]; en la extracción de características se obtienen diferentes índices que son usados, junto a información extra de la calidad del aire [8]–[10] como variables para el análisis de la calidad ambiental urbana. Tal como presenta [11] al estimar la concentración de material particulado PM_{10} en zonas con pocas estaciones usando un modelo de regresión empírico del uso del suelo LUR e imágenes de los sensores Landsat 7 ETM+, Landsat 8 OLI y MODIS Aqua-Terra para extraer los índices $NDVI$, $NDSI$, $SAVI$, $NDWI$ y TS como variables de entrada.

Entre los métodos más usados para el análisis de la calidad ambiental se encuentra el Análisis por Componentes Principales PCA . Esta técnica estudia la relación entre variables correlacionadas permitiendo reducir su número a un conjunto de variable sin correlacionar [12]. Tal es el caso de [13], que evalúa la calidad del medio ambiente urbano por medio de un modelo lineal construido a partir de componentes principales extraídos de variables socioeconómicas, datos ecológicos e indicadores ambientales obtenidos de imágenes Landsat TM. Algo similar realiza [14] por medio del análisis de componentes principales al estimar el índice de calidad urbano usando imágenes Landsat ETM+ para el cálculo de los indicadores ambientales. Por su parte [15], mediante el análisis de componentes principales, calcula un índice de calidad ambiental EQI considerando variables agrupadas por dominios (socioeconómico, construcciones, calidad del aire, calidad del agua y uso del suelo) y aplicando PCA sobre cada uno de los grupos de variables.

Este documento presenta una metodología para el análisis de la contaminación ambiental en Medellín a partir de datos de estaciones meteorológicas e imágenes de los satélites Landsat 7 y Landsat 8. El artículo se ordena de la siguiente manera: la sección 2, muestra los materiales y métodos; sección 3, los resultados obtenidos; y, la sección 4, las conclusiones.

2. MATERIALES Y MÉTODOS

La metodología propuesta para el análisis de la calidad ambiental aplicando técnicas de teledetección y análisis de componentes principales se desarrolla en el Software ENVI 5.0 y permite el cálculo de los índices TS , $NDVI$, $TSAVI$, $NDWI$ y NSI ; la interpolación de variables de calidad del aire PM_{10} , $PM_{2.5}$, NO_2 y O_3 ; y el Análisis de Componentes Principales,

(ver Figura 1). Se utilizan imágenes entre el año 2016 al 2019 obtenidas de la *USGS* (United States Geological Survey) de los satélites Landsat 7, que contiene 3 bandas del visible (B1 azul, B2 verde y B3 rojo), 2 bandas del infrarrojo cercano (B4 y B5), 1 banda térmica (B6), 1 banda infrarrojo medio (B7) y 1 banda pancromática (B8); Landsat 8, contiene 3 bandas del visible (B2 azul, B3 verde y B4 rojo), 1 banda de aerosol (B1), 1 banda de infrarrojo cercano (B5), 2 bandas SWIR (B6 y B7), 1 banda pancromática (B8), 1 banda cirrus (B9) y 2 bandas TIRS (B10 y B11) y mediciones meteorológicas de las variables de calidad del aire NO_2 , O_3 , $PM_{2.5}$, y PM_{10} de las mismas fechas. La información relacionada con estas variables es obtenida del *SIATA* (Sistema de Alerta Temprana de Medellín y el Valle de Aburrá) (ver Tabla 1).

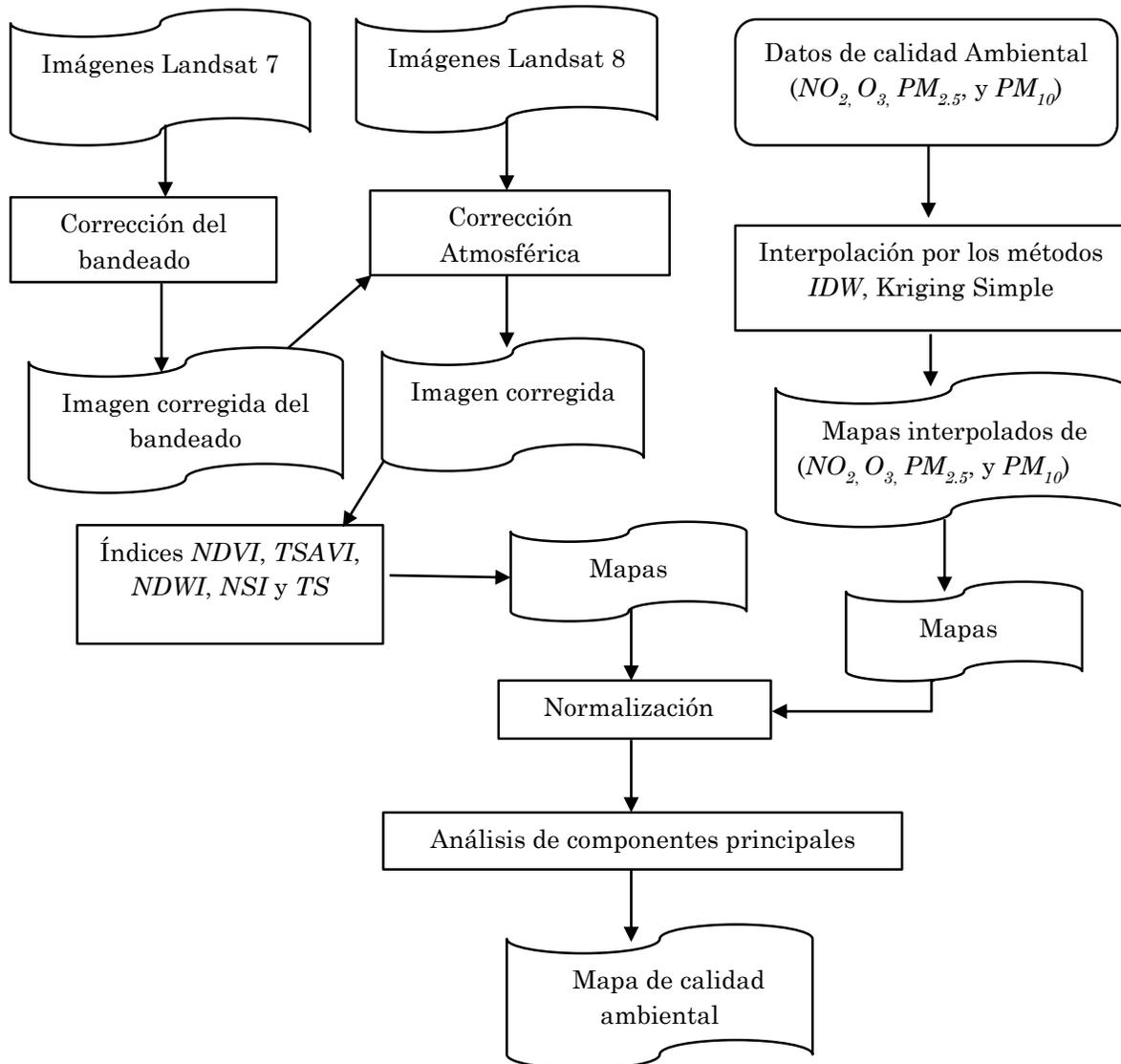


Figura 1. Metodología Propuesta. Fuente: elaboración propia.

Tabla 1. Fecha y hora de adquisición de la información. Fuente: elaboración propia.

Fecha	Hora de captura imagen	Imagen	Hora de captura variable ambiental
18/12/2017	15:18	Landsat 8	15:00
09/04/2018	15:17	Landsat 8	15:00
21/11/2016	15:21	Landsat 7	15:00
15/02/2019	15:12	Landsat 7	15:00

2.1 Preprocesamiento de las imágenes

Las imágenes requieren de un preprocesamiento para disminuir los efectos del ruido geométrico y radiométrico presente, y ser el insumo en las etapas de análisis visual o digital.

La corrección atmosférica disminuye las distorsiones producidas por la atmosfera que se introducen en los valores de radiancia en las bandas de las imágenes, esta se calcula por medio de (1) [16].

$$\rho = \frac{\pi * [L - L_a] * d^2}{\cos \theta * E_o \tau_1 \tau_2} \quad (1)$$

Donde ρ es la reflectividad, E_o es la irradiancia exoatmosférica, τ_1 es el coeficiente de transmisión atmosférica sol-tierra para cada una de las bandas $B1\tau_1 = 0.70$, $B2\tau_1 = 0.78$, $B3\tau_1 = 0.85$, $B4\tau_1 = 0.91$, $B5\tau_1 = 0.95$ y $B7\tau_1 = 0.97$, $\tau_2 = 1$ es el coeficiente de transmisión atmosférica tierra-sensor, L es la radiancia del pixel a corregir, L_a es el valor mínimo de niveles digitales por cada banda, θ es el ángulo cenital solar y $d = 1.49598 * 10^8$ Km es la distancia tierra-sol [5].

El preprocesamiento se realiza en el Software ENVI. Por medio del toolbox Radiometric Calibration se convierten los niveles digitales a valores de radiancia. Por otra parte, se realiza la corrección radiométrica del bandeo que corrige el efecto de desplazamiento del histograma de la imagen con el fin de obtener el mismo valor promedio y una misma desviación típica para todas las bandas [17]. Esto se realiza para las imágenes con bandeo del sensor Landsat 7, por medio de la herramienta Landsat_gapfill de ENVI.

2.2 Procesamiento de las imágenes

En esta etapa se usan las imágenes corregidas radiométrica y geoméricamente como insumos para estimar distintas variables que son indicadores de calidad del aire.

A continuación, se describe cómo estimar cada una de estas variables.

2.2.1 Temperatura de la superficie TS

La temperatura de la superficie se calcula por medio de (2) [6].

$$TS = \left(\frac{T_l}{1 + \left(\frac{\lambda T_l}{\rho} \right) * \ln \varepsilon} \right) \quad (2)$$

Donde T_l es la temperatura de brillo superficial obtenida en función de las constantes de calibración del sensor, $\lambda = 0.00115 \mu\text{m}$ es la longitud de onda media de la banda térmica, $\rho = 1.438 * 10^{-2} \text{ mK}$ y ε es la emisividad de la superficie, la cual se calcula teniendo en cuenta la proporción de vegetación [6].

2.2.2 Índice de vegetación normalizado NDVI

Los índices de vegetación son transformaciones matemáticas entre distintas bandas espectrales de la misma escena. Estos permiten evaluar la proporción de reflectancia en la vegetación verde para las longitudes de onda asociadas a la luz roja e infrarroja, pues la clorofila absorbe la luz roja en el proceso de fotosíntesis. El índice de vegetación normalizado *NDVI*, es uno de los índices más usados en la literatura científica, ya que facilita la interpretación de los parámetros biofísicos de la vegetación fotosintéticamente activa [18], como en (3).

$$NDVI = \frac{BIRC - BR}{BR + BIRC} \quad (3)$$

Donde *BIRC* es la banda del infrarrojo cercano y *BR* la banda rojo del visible [6].

2.2.3 Índice de vegetación ajustado al suelo transformado TSAVI

El índice *TSAVI* es un ajuste al *NDVI* en el que se busca corregir el efecto del brillo del suelo en zonas con escasa cobertura vegetal [19], (4).

$$TSAVI = \frac{BIRC - (sR + a)}{\frac{1}{s}[aBIRC + BR - (sa) + 0.08(1 + s^2)]} \quad (4)$$

En (4) el numerador aparece con la diferencia entre el valor de la banda *BIRC* y el valor que tendría esta banda si toda la luz en el rojo procediese del suelo, y el denominador se define de tal forma que se corrija los efectos debido a la visión del suelo entre la vegetación, *s* y *a* son la pendiente y el intercepto con el suelo [20].

2.2.4 Índice de diferencia normalizada del agua NDWI

El índice de diferencia normalizada del agua permite medir los cambios en el contenido de agua de la vegetación, permitiendo identificar el nivel de saturación de humedad de la vegetación [21], (5).

$$NDWI = \frac{(BG - BIRC)}{(BG + BIRC)} \quad (5)$$

Donde *BG* es la banda verde [21].

2.2.5 Índice normalizado del suelo *NSI*

El índice normalizado del suelo *NSI* analiza la cantidad de suelo edificado que abarca la superficie urbana, indicando que a mayor área construida los valores son más altos [22], como en (6):

$$NSI = \frac{(BIRS - BIRC)}{(BIRS + BIRC)} \quad (6)$$

Donde *BIRS* es la banda correspondiente al infrarrojo de onda corta.

2.3 Métodos de interpolación

La interpolación es usada para predecir valores en una superficie en lugares que no tienen información [23]. Esta es usada para la determinación de las superficies de las variables de calidad del aire *NO₂*, *O₃*, *PM_{2.5}*, y *PM₁₀*. Entre los diferentes métodos de interpolación se encuentran los métodos determinísticos y geoestadísticos. La ponderación de distancia inversa *IDW* es uno de los métodos determinísticos más comúnmente empleados. Por su parte, los interpoladores geoestadísticos cuantifican la estructura espacial de los datos por medio de variogramas. Estos están en el grupo de los kriging [24].

El kriging simple es usado si los fenómenos son estacionarios, con varianza y esperanzas conocidas y constantes [23]. Como en (7):

$$z(x_0) = m + \sum_{i=1}^n \lambda_i [z(x_i) - m] \quad (7)$$

Donde $z(x_0)$ es el valor estimado de la variable en la posición x_0 , λ_i son los pesos calculados por medio de la matriz de covarianza, m es la media del valor de la variable, $Z(x_i)$ son las mediciones de las variables de interés en los puntos $i = 1, 2, \dots, n$ [23].

El *IDW* asigna los pesos a los datos muestreados entorno a una función inversa de la distancia que los separa, por lo cual los puntos más cercanos tienen un peso mayor en el cálculo [25], como en (8):

$$\hat{Z}_j = \sum_{i=1}^n k_{ij} \cdot z_j \quad (8)$$

Donde \hat{Z}_j es el valor estimado para el punto, j . n es el número de puntos utilizados para la interpolación, z_j es el valor en el punto i -ésimo, k_{ij} es el peso asociado al dato i en el cálculo del nodo j . Los pesos k varían entre 0 y 1 según el dato y la suma total de los pesos es la unidad [25].

2.4 Análisis de componentes principales

El análisis de componentes principales es una técnica multivariante que estudia la relación que se presenta entre variables correlacionadas y reduce el número de dimensiones mediante la creación de un conjunto de variables sin correlación. Las componentes principales

son combinaciones lineales de las variables correlacionadas y se construyen en orden de importancia en cuanto a la variabilidad total de la muestra [26].

El análisis de componentes principales permite el paso de un espacio vectorial R^n a un subespacio R^m ($n > m$) sin pérdida de información relevante, maximizando su varianza. El *PCA* calcula una matriz de autocorrelación de los datos con sus vectores propios y los ordena de acuerdo con su valor propio para posteriormente normalizarlos [27]. La componente se calcula mediante un conjunto de variables originales (x_1, x_2, \dots, x_p) , siendo x un vector con p variables aleatorias [28], como en (9):

$$\alpha'_x = \alpha_{11} x_1 + \alpha_{12} x_2 + \dots + \alpha_{1p} x_p = \sum_{j=1}^p \alpha_{1j} x_j \quad (9)$$

Donde α'_x es la función lineal de los elementos de x de máxima varianza, y α_1 es un vector de p constantes y denota la transpuesta $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$. El segundo componente se calcula con $\alpha'_2 x$ incorrelacionada con $\alpha'_1 x$. Así se eligen los componentes no correlacionados entre sí, de modo que las variables aleatorias tengan cada vez menor varianza [28]. Por medio de *PCA* se estima el mapa de calidad ambiental a partir de las variables de calidad del aire NO_2 , O_3 , $PM_{2.5}$, y PM_{10} y los índices (*TS*, *NDVI*, *TSAVI*, *NDWI* y *NSI*).

3. RESULTADOS Y DISCUSIÓN

Con las imágenes Landsat 7 y 8 se realizó un preprocesamiento que consistió en la corrección de bandeo para el caso del Landsat 7, la corrección atmosférica para todas las imágenes y el recorte para ajustar a la zona de estudio la ciudad de Medellín. Seguidamente, se calculó cada uno de los índices (*TS*, *NDVI*, *TSAVI*, *NDWI* y *NSI*) para cada una de las fechas. También, con las mediciones meteorológicas se estimaron las superficies de las variables NO_2 , O_3 , $PM_{2.5}$, y PM_{10} usando los métodos de interpolación kriging simple e *IDW*.

Las superficies asociadas a todas las variables se obtienen con un tamaño de píxel de 30 metros. Después, mediante el análisis de componentes principales se generó un mapa de calidad ambiental asociado a cada fecha estudiada y se identificaron las áreas con calidad de aire más deficiente. Esta información fue cruzada con información socioeconómica descargada del Portal Geográfico del Municipio de Medellín, la cual contenía la localización de los barrios y el estrato socioeconómico.

A continuación, se describe cada superficie encontrada con el fin de identificar los sectores con la temperatura más alta, las zonas con mayor cobertura de vegetación, las áreas con mayor contenido de agua y las zonas con mayor suelo construido.

3.1 Temperatura de la Superficie *TS*

Se calculó la temperatura de la superficie para las imágenes entre 2016 y 2019. Al igual que en el estudio de [29] las zonas con mayores temperaturas están ubicadas alrededor del corredor vial del río Medellín y la zona centro de la ciudad como se observa en la Figura 2.

En estos sectores se presentan temperaturas por encima de los 30°C y son sectores con construcciones grandes y de baja altura.

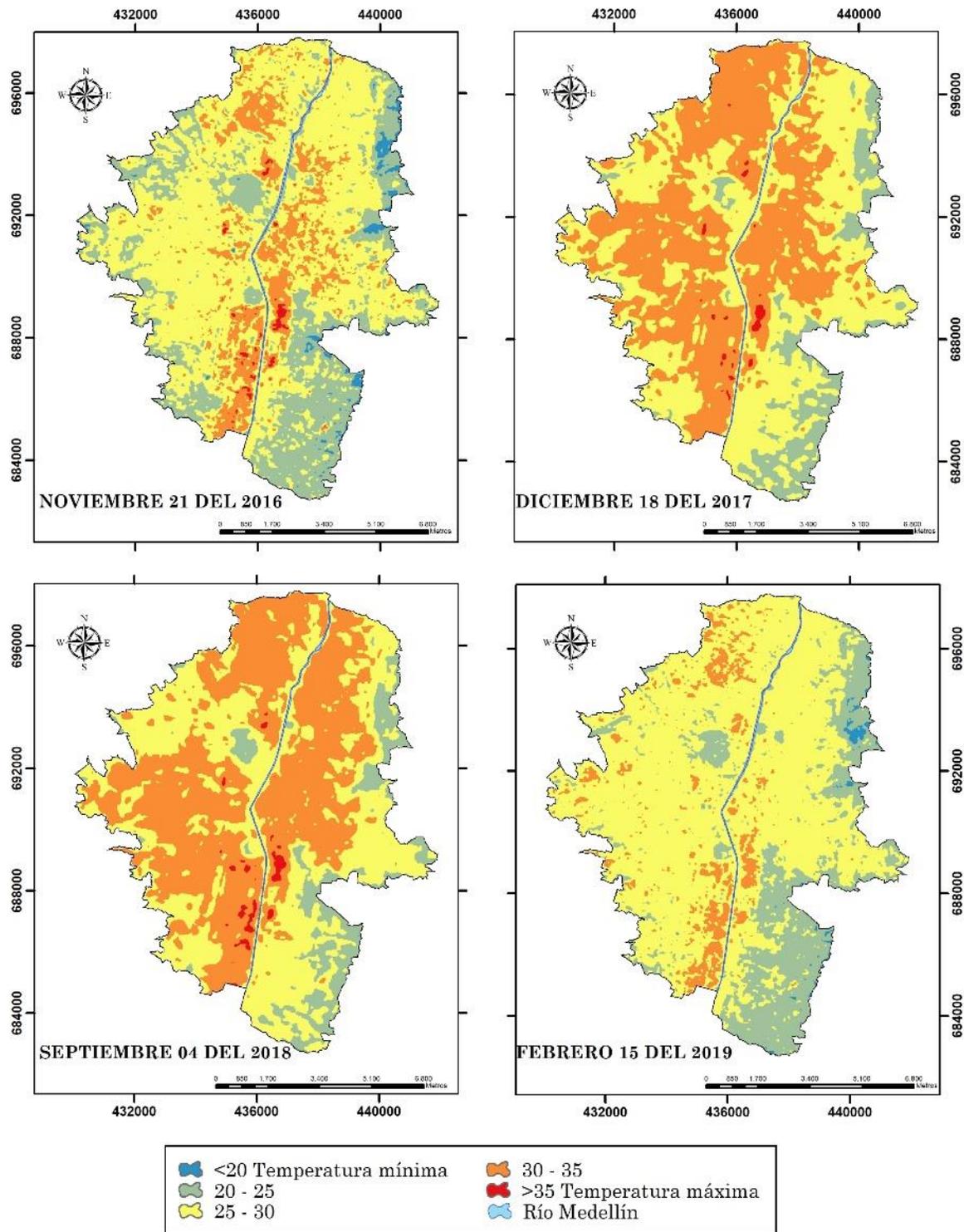


Figura 2. Temperatura de la superficie obtenida para cada una de las fechas de estudio.
Fuente: elaboración propia.

3.2 Índice de vegetación *NDVI*

El índice de vegetación *NDVI* permitió identificar las zonas con cobertura de vegetación. Se evidenció que en los extremos del casco urbano se presenta una vegetación dispersa. El sector de la ciudad con mayor vegetación dispersa se encuentra en el sureste, siendo un sector residencial de estrato socioeconómico 5 y 6, con una gran cobertura de árboles. Por otro lado, el centro de Medellín es la zona donde se presenta menos cobertura de vegetación al igual que en los barrios aledaños al corredor vial del río Medellín. La Figura 3 muestra el índice *NDVI* obtenido en las fechas estudiadas.

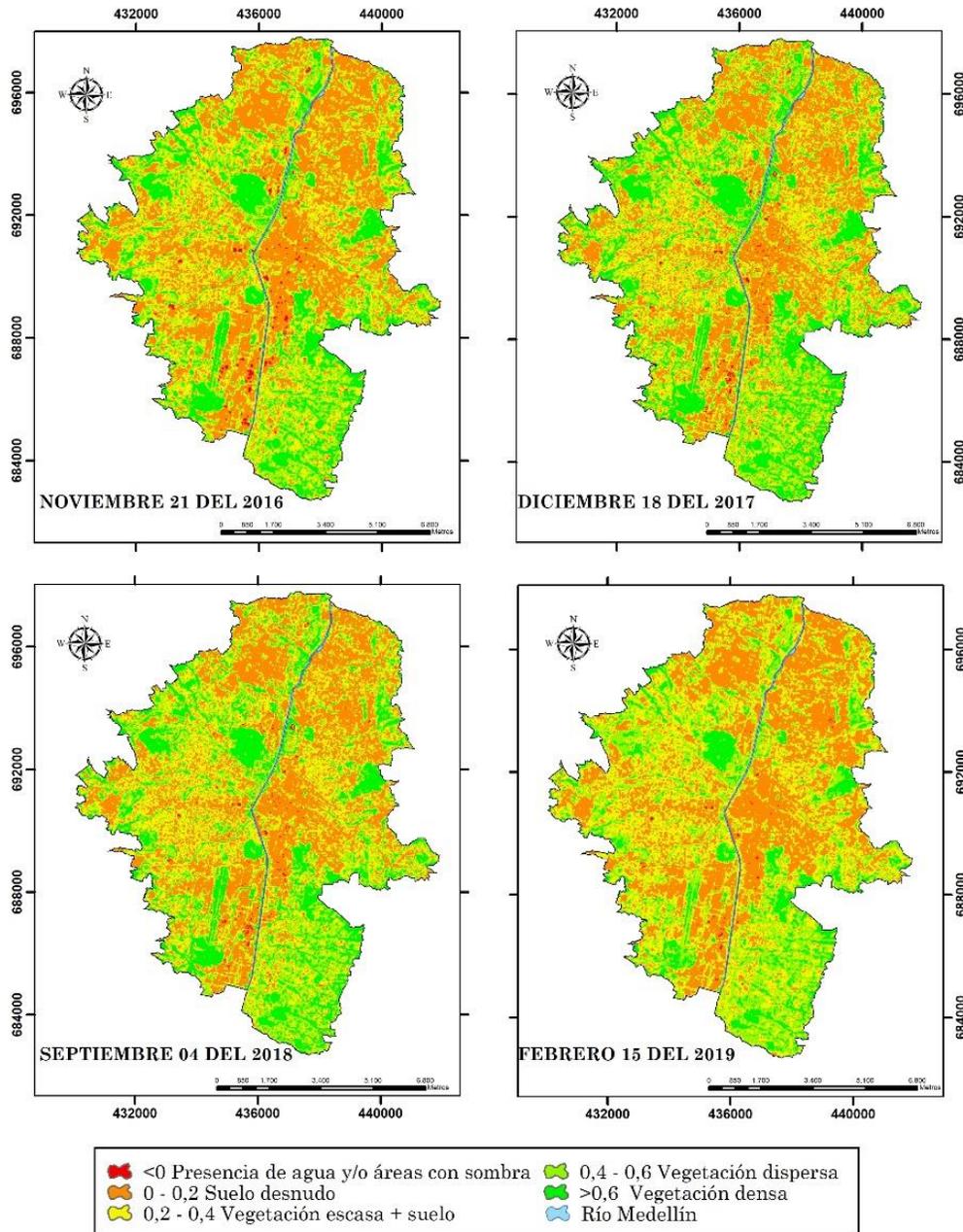


Figura 3. Índice de vegetación *NDVI* obtenido para cada una de las fechas. Fuente: elaboración propia.

3.3 Índice de vegetación *TSAVI*

El índice de vegetación *TSAVI* puede verse afectado por las variaciones en el brillo del suelo, por lo cual los valores de la cubierta vegetal son más independientes al reflejo del suelo [30]. Este índice permitió determinar las coberturas de vegetación en el casco urbano de Medellín. A diferencia del *NDVI*, *TSAVI* identificó menos áreas con vegetación en la zona céntrica de la ciudad, lo cual se asemeja a lo identificado por [29], en su estudio de islas de calor urbana en el Valle De Aburrá. La Figura 4 ilustra el índice calculado para cada una de las fechas de estudio.

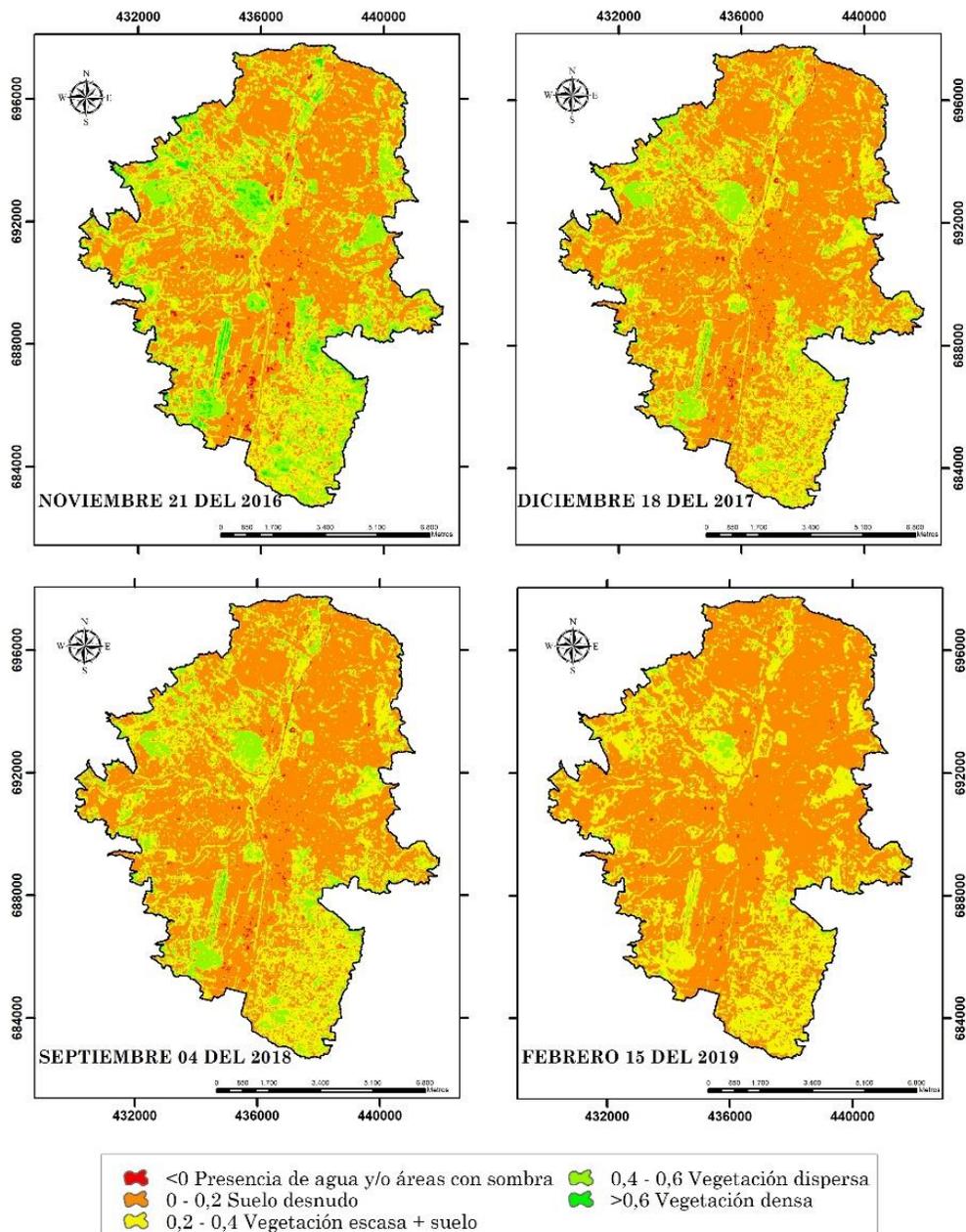


Figura 4. Índice de vegetación *TSAVI* obtenido para cada una de las fechas. Fuente: elaboración propia.

3.4 Índice de vegetación *NDWI*

Este índice determinó las zonas con mayor concentración de humedad, donde al igual que el *NDVI* y el *TSAVI* en la zona suroriental y noroccidental, es donde se presenta mayor vegetación y por ende mayor contenido de humedad. En la Figura 5 se presenta el índice obtenido para cada una de las fechas.

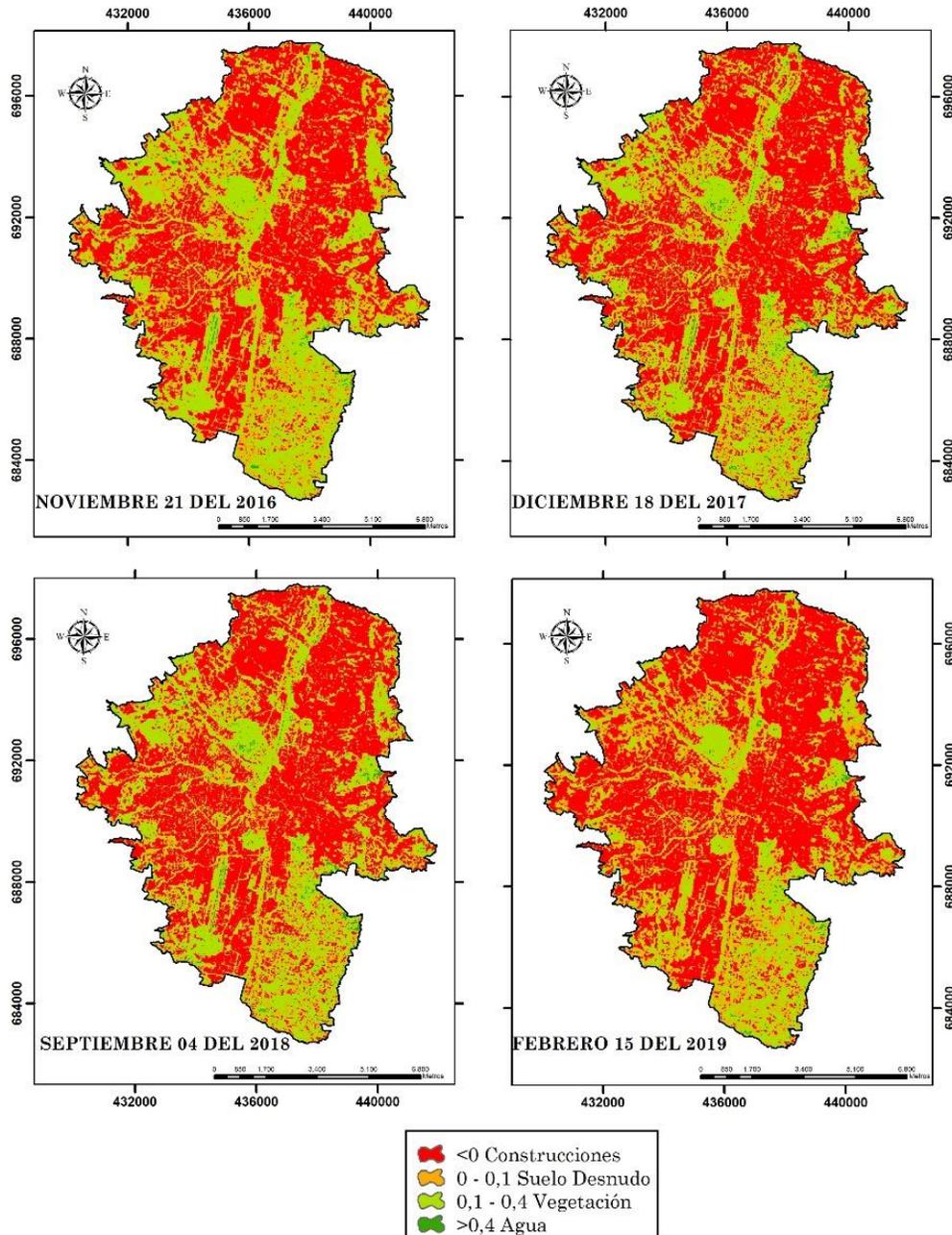


Figura 5. Índice de vegetación *NDWI* obtenido para cada una de las fechas. Fuente: elaboración propia.

3.5 Índice de vegetación *NSI*

Este índice determinó las áreas construidas con el fin de conocer cómo ha cambiado el uso del suelo y su expansión a lo largo del tiempo y el espacio. Para el año 2019 se presentaron cambios en el uso del suelo (ver Tabla 2), siendo ese año el que mostró la mayor variación. Por otro lado, entre los años 2017 y 2018 no se presentó variación en cuanto al área construida y por lo general las fechas de estudio no presentaron una variación significativa. La Figura 6 ilustra el índice de vegetación obtenido.

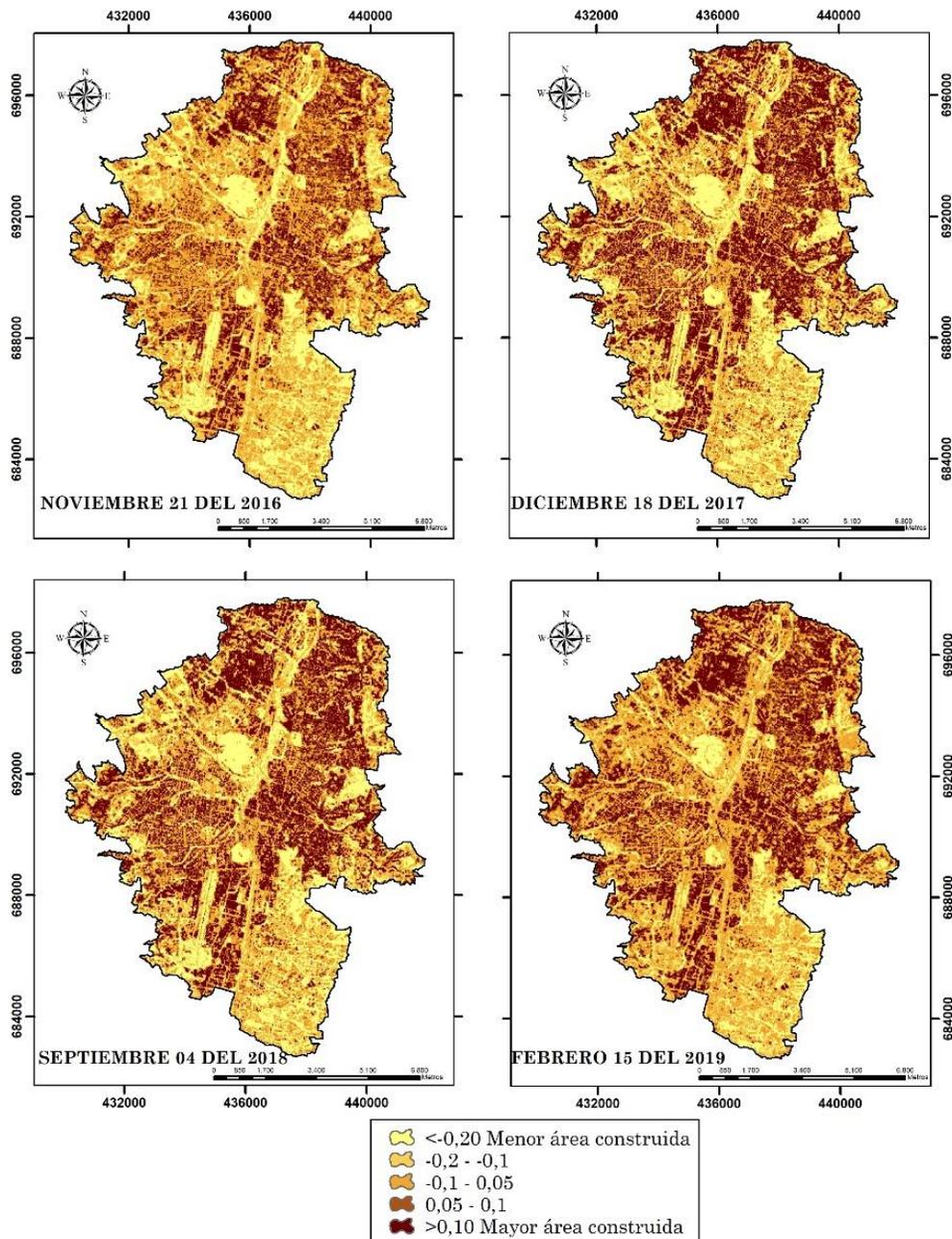


Figura 6. Índice de vegetación *NSI* obtenido para cada una de las fechas. Fuente: elaboración propia.

Tabla 2. Matriz de correlación del *NSI*. Fuente: elaboración propia.

<i>NSI</i>	2016	2017	2018	2019
2016	1.000	0.919	0.919	0.890
2017	0.919	1.000	1.000	0.876
2018	0.919	1.000	1.000	0.876
2019	0.890	0.876	0.876	1.000

3.6 Análisis de relaciones entre los índices

Por medio de matrices de correlación se analizaron las relaciones entre los diferentes índices (*NDVI*, *TSAVI*, *NDWI*, *NSI* y *TS*) para cada una de las fechas estudiadas (ver Tabla 3). Se observó que se presenta una correlación alta entre el *NDWI*, *TSAVI* y *NDVI*, debido a que el primero analiza la cantidad de humedad en el suelo mientras el *TSAVI* y *NDVI* identifican las zonas con cobertura de vegetación.

Tabla 3. Correlación de los índices para cada uno de los años de estudio. Fuente: elaboración propia.

Año 2016					
Variables	<i>TS</i>	<i>NDVI</i>	<i>TSAVI</i>	<i>NDWI</i>	<i>NSI</i>
<i>TS</i>	1.000	-0.697	-0.680	-0.695	0.695
<i>NDVI</i>	-0.697	1.000	0.993	0.905	-0.905
<i>TSAVI</i>	-0.680	0.993	1.000	0.901	-0.901
<i>NDWI</i>	-0.695	0.905	0.901	1.000	-1.000
<i>NSI</i>	0.695	-0.905	-0.901	-1.000	1.000
Año 2017					
Variables	<i>TS</i>	<i>NDVI</i>	<i>TSAVI</i>	<i>NDWI</i>	<i>NSI</i>
<i>TS</i>	1.000	-0.583	-0.526	-0.590	0.590
<i>NDVI</i>	-0.583	1.000	0.976	0.913	-0.913
<i>TSAVI</i>	-0.526	0.976	1.000	0.892	-0.892
<i>NDWI</i>	-0.590	0.913	0.892	1.000	-1.000
<i>NSI</i>	0.590	-0.913	-0.892	-1.000	1.000
Año 2018					
Variables	<i>TS</i>	<i>NDVI</i>	<i>TSAVI</i>	<i>NDWI</i>	<i>NSI</i>
<i>TS</i>	1.000	-0.641	-0.612	-0.640	0.640
<i>NDVI</i>	-0.641	1.000	0.987	0.891	-0.891
<i>TSAVI</i>	-0.612	0.987	1.000	0.876	-0.876
<i>NDWI</i>	-0.640	0.891	0.876	1.000	-1.000
<i>NSI</i>	0.640	-0.891	-0.876	-1.000	1.000
Año 2019					
Variables	<i>TS</i>	<i>NDVI</i>	<i>TSAVI</i>	<i>NDWI</i>	<i>NSI</i>
<i>TS</i>	1.000	-0.582	-0.528	-0.639	0.639
<i>NDVI</i>	-0.582	1.000	0.984	0.870	-0.870
<i>TSAVI</i>	-0.528	0.984	1.000	0.831	-0.831
<i>NDWI</i>	-0.639	0.870	0.831	1.000	-1.000
<i>NSI</i>	0.639	-0.870	-0.831	-1.000	1.000

Por otro lado, se presentó una correlación negativa o inversa entre el *NDVI* y el *NSI*, dado que el *NSI* busca identificar las zonas con mayores construcciones y el *NDVI* las coberturas de vegetación.

La temperatura *TS* presentó una correlación positiva con el *NSI*, dado que a mayor área construida mayor aumento de temperatura, es por ello por lo que en la zona centro de la ciudad donde se presenta una alta densidad de construcciones, la temperatura es mayor.

Lo contrario ocurre con los índices *NDVI*, *TSAVI* y *NDWI*, ya que la correlación es negativa debido que a mayor temperatura, menor cobertura de vegetación y menor humedad del suelo lo cual concuerda con el estudio de [14], por lo tanto en la zonas con mayor cobertura de vegetación la temperatura es menor, lo cual se ve reflejado en la zona sureste de la ciudad.

El *TSAVI* y el *NDVI* presentaron una alta correlación debido a que ambos índices identifican cobertura de vegetación.

3.7 Interpolación de variables de calidad del aire

Las variables de calidad del aire descargadas del *SIATA* fueron adquiridas de la misma fecha de la toma de las imágenes. La red de monitoreo del *SIATA* cuenta con 43 estaciones a lo largo del Valle de Aburrá, de las que se tomó la información de 30 (ver Figura 7).

Las variables *PM₁₀*, *PM_{2.5}*, *NO₂* y *O₃* fueron interpoladas por el método kriging simple para cada una de las fechas de estudio. Este método generó una superficie homogénea que representaba la tendencia de los datos y la distribución de los mismos.

Para las variables *PM_{2.5}* y *PM₁₀* del año 2016, se utilizó el método de interpolación *IDW* debido a que los datos poseían un distanciamiento grande y el Kriging muestra problemas en la generación de los variogramas, ocasionando que los datos no interpolen o no representen la realidad, por lo cual el *IDW* se perfiló mejor, para este tipo de distribución de datos.

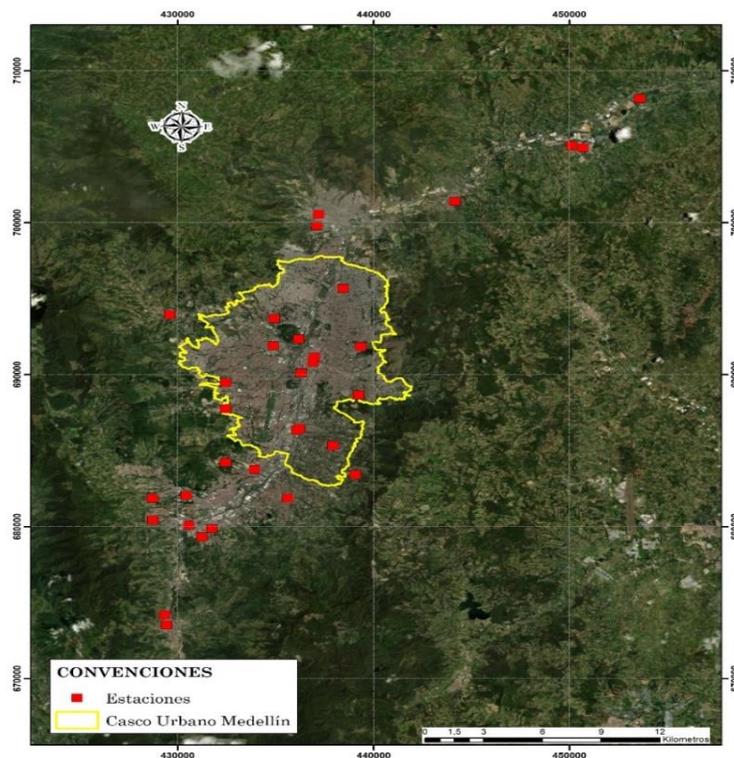


Figura 7. Distribución de estaciones a lo largo del Valle de Aburrá. Fuente: elaboración propia.

3.8 Mapa de calidad ambiental

Al igual que en [14], se determinó que las variables ambientales (*NDVI*, *TSAVI*, *NDWI*, *NSI* y *TS*) presentaban una correlación alta, positiva o negativa, y por ello eso se decidió usar el *PCA*. Se calcularon los componentes principales con variables normalizadas entre 0 y 1, dado que la información debe estar en un mismo rango. El primer componente principal contenía el 90 % de la variabilidad en la información y se tomó este como el mapa de calidad de ambiental. Este mapa fue categorizado mediante la desviación típica de cada uno de los datos para cada fecha de estudio. La Figura 8 ilustra los mapas de calidad ambiental para cada una de las fechas.

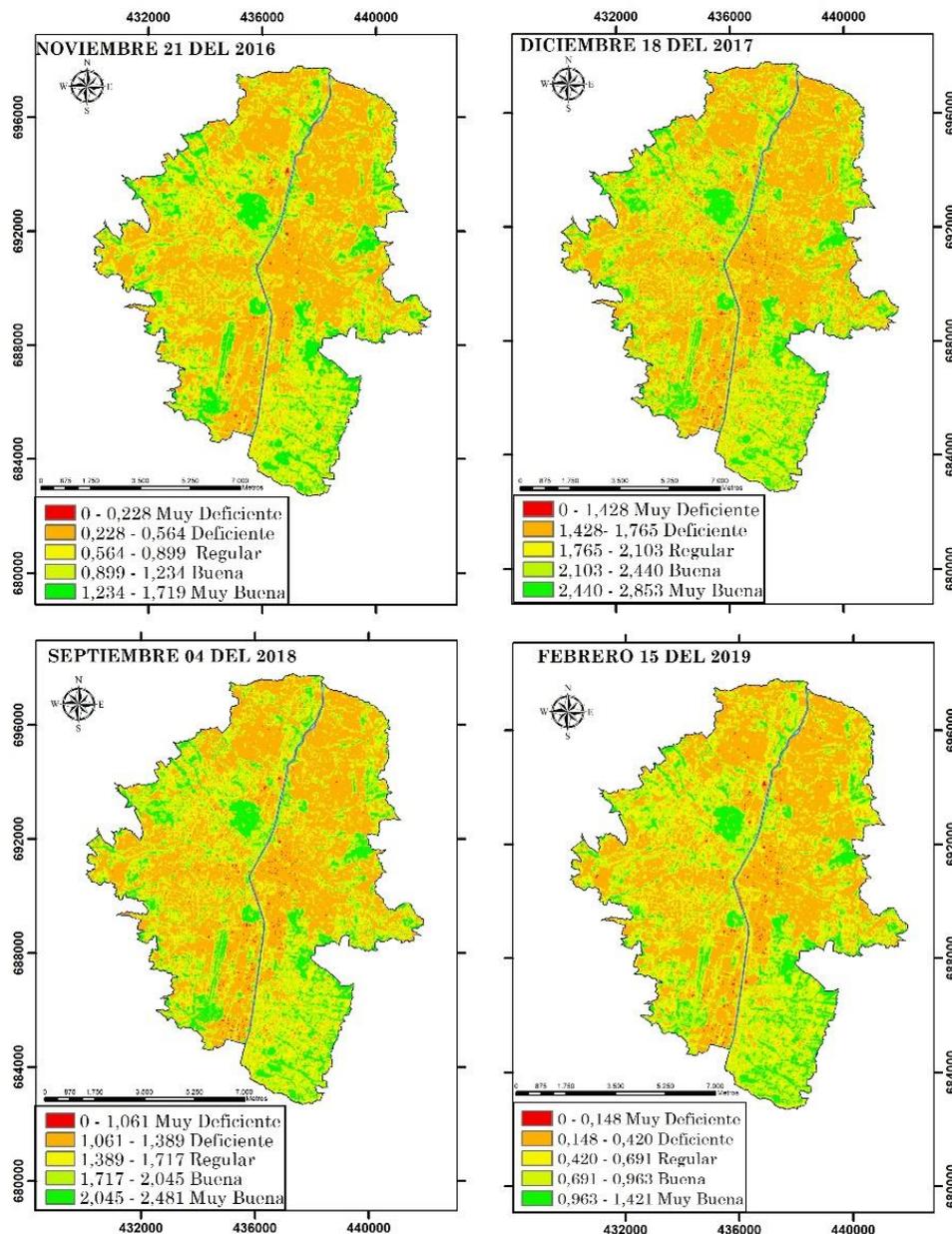


Figura 8. Mapas de Calidad Ambiental para cada una de las fechas de estudio. Fuente: elaboración propia.

En el año 2016, 16.47 hectáreas presentaban calidad ambiental muy deficiente y para el año 2017 ese valor aumentó a 40.32 hectáreas, ya que zonas con calidad ambiental deficiente pasaron a la categoría muy deficiente. En el año 2018 el área disminuyó a 36.36 Ha y esta tendencia se mantuvo hasta el año 2019 con 31.32 Ha.

La calidad ambiental buena tendió a aumentar, ya que para el año 2016 era de 1884.24 Ha y para el año 2019 paso a ser 1984.05 Ha, debido a que zonas que presentaban una calidad ambiental moderada pasaron a tener una buena calidad ambiental. Por el contrario, zonas categorizadas con muy buena calidad pasaron a tener calidad buena, como en la zona sureste de la ciudad.

Se evidenció que las zonas con calidad muy deficiente son sectores con una gran cantidad de construcciones y poca cobertura de vegetación, como la zona centro de la ciudad, donde se integran rutas de buses, el metro y se presenta una gran concurrencia de personas.

Las zonas con una calidad del aire muy buena y buena presentan una mayor cobertura de vegetación. Por lo general están ubicadas en los extremos de la ciudad, zonas de estrato socioeconómico 4, 5, y 6, lo cual evidencia que la calidad ambiental está muy asociada a la cobertura de vegetación que es un factor clave de balance y al estrato socioeconómico, debido al área construida y el diseño urbanístico. Los focos de contaminación se ubican en el centro de la ciudad y en algunos sectores aledaños al Aeropuerto Olaya Herrera, debido a su diseño urbanístico con poca cobertura de vegetación, una gran cantidad de construcciones y un gran flujo vehicular.

En este estudio se analizó la calidad ambiental por medio de *PCA* usando variables ambientales y variables de calidad de aire, a diferencia [12] que usó en su estudio en los Estados Unidos variables de calidad de aire, calidad de agua, índices ambientales, variables de construcción y sociodemográficas. Por otro lado [13] realizó su estudio usando variables socioeconómicas, ecológicas e índices ambientales.

La metodología propuesta para el análisis de la contaminación ambiental se usó para analizar un periodo de 4 años con datos anuales, pero puede adaptarse para utilizar datos con periodicidad trimestral para analizar efectos estacionales. Asimismo, incluyendo variables de calidad del agua, socioeconómicas y sociodemográficas que permitan la creación de índice de calidad ambiental que analice la contaminación ambiental desde varios componentes.

4. CONCLUSIONES

Se propuso una metodología para analizar la distribución espacial de la contaminación ambiental en el casco urbano de Medellín utilizando técnicas de percepción remota y análisis estadísticos. Con la metodología propuesta se obtiene un mapa de calidad ambiental que se enfoca en la contaminación atmosférica dado el uso de variables de calidad del aire como PM_{10} , $PM_{2.5}$, NO_2 y O_3 . Este mapa se obtuvo a partir del primer componente, el cual contenía el 90 % de variación de la información y permitió ubicar los sectores con la calidad ambiental más deficiente, siendo insumo para la planificación urbana y para dar atención prioritaria a las áreas más críticas.

De acuerdo con la metodología propuesta, los principales focos de contaminación del casco urbano de Medellín se presentaron en la zona centro de la ciudad y en sectores aledaños al Aeropuerto Olaya Herrera. Estos focos corresponden a zonas con poca cobertura de vegetación, un área construida extensa y elevado flujo vehicular.

El mapa de calidad ambiental puede ser usado como insumo en la oportuna toma de decisiones en cuanto a la planificación urbana por parte de entidades como la Alcaldía y las secretarías de Medio Ambiente y de Planeación, ya que posibilita la pronta intervención en

las zonas donde la calidad ambiental es muy deficiente. Proporciona también que las entidades de salud puedan identificar los sectores más vulnerables por la Muy Deficiente calidad ambiental y hacer seguimiento a los residentes de estos sectores, dado los problemas de salud que esto conlleva.

La metodología propuesta puede ser aplicada en diferentes zonas y se recomienda incluir variables socioeconómicas como uso del suelo urbano, densidad de población y variables de calidad de agua, entre otras, debido a que la incorporación de otros indicadores robustece el análisis y permite una mejor comprensión de los factores que influyen en la contaminación ambiental.

5. AGRADECIMIENTOS

Los autores agradecen al SIATA por suministrar la información de las variables de calidad del aire Material Particulado (PM_{10} y $PM_{2.5}$), Dióxido de Nitrógeno (NO_2) y Ozono (O_3) y a la Universidad de Manizales por el apoyo para el desarrollo de la investigación. El artículo no es financiado por ninguna entidad.

CONFLICTOS DE INTERÉS DE LOS AUTORES

Los autores declaran que no se presentan conflictos de interés.

CONTRIBUCIÓN DE LOS AUTORES

Lizette Tello-Cifuentes: conceptualización, diseño, desarrollo y análisis de la investigación para tesis de maestría y escritora del artículo.

Jean P. Díaz Paz: contribuyó como director de tesis, supervisó y revisó durante el planteamiento, desarrollo y análisis de la investigación y apoyó en la escritura del artículo.

6. REFERENCIAS

- [1] Ministerio de Salud y Protección Social Sostenible, Resolución 2154. “Por la cual se establece el reglamento técnico sobre los requisitos sanitarios que deben cumplir los aceites y grasas de origen vegetal o animal que se procesen, envasen, almacenen, transporten, exporten, importen y/o comercialicen en el país, destinados para el consumo humano y se dictan otras disposiciones.”, Agosto. 2012. [URL](#)
- [2] A. M. Muñoz D.; J. J. Paz V.; C. M. Quiroz P., “Efectos de la contaminación atmosférica sobre la salud de adultos que laboran en diferentes niveles de exposición,” *Rev. Fac. Nac. Salud Pública*, vol. 25, no. 2, pp. 85–94, 2007. [URL](#)
- [3] M. Romero Placeres; F. D. Olite; M. Álvarez Toste, “La contaminación del aire: su repercusión como problema de salud,” *Rev. Cuba. Hig. Epidemiol.*, vol. 44, no. 2, pp. 1–14, 2006. [URL](#)
- [4] E. Aránguez *et al.*, “Contaminantes atmosféricos y su vigilancia,” *Rev. Esp. Salud Pública*, vol. 73, no. 2, pp. 123–132, Mar. 1999. [URL](#)
- [5] S. Hantson *et al.*, “Cadena de preprocesamiento estándar para las imágenes Landsat del Plan Nacional de Teledetección,” *Rev. Teledetección*, no. 36, pp. 51–61, 2011. [URL](#)
- [6] M. Preciado; A. Aldana, “Análisis de presencia de islas de calor en Santiago de Cali empleando técnicas de teledetección,” *Ventana Inform.*, no. 24, pp. 95–114, Jul. 2011. <https://doi.org/10.30554/ventanainform.24.162.2011>

- [7] B. P. Page; L. G. Olmanson; D. R. Mishrac, “A harmonized image processing workflow using Sentinel-2/MSI and Landsat-8/OLI for mapping water clarity in optically variable lake systems,” *Remote Sens. Environ.*, vol. 231, p. 111284, Sep. 2019. <https://doi.org/10.1016/j.rse.2019.111284>
- [8] W. Mueller *et al.*, “Urban greenspace and the indoor environment: Pathways to health via indoor particulate matter, noise, and road noise annoyance,” *Environ. Res.*, vol. 180, p. 108850, Jan. 2020. <https://doi.org/10.1016/j.envres.2019.108850>
- [9] I. Moreira; C. Linares; F. Follos; G. Sánchez-Martínez; J. M. Vellón; J. Díaz, “Short-term effects of Saharan dust intrusions and biomass combustion on birth outcomes in Spain,” *Sci. Total Environ.*, vol. 701, p. 134755, Jan. 2020. <https://doi.org/10.1016/j.scitotenv.2019.134755>
- [10] Y. Zhang; Z. Li, “Remote sensing of atmospheric fine particulate matter (PM_{2.5}) mass concentration near the ground from satellite observation,” *Remote Sens. Environ.*, vol. 160, pp. 252–262, Apr. 2015. <https://doi.org/10.1016/j.rse.2015.02.005>
- [11] C. I. Alvarez-Mendoza; A. C. Teodoro; N. Torres; V. Vivanco, “Assessment of Re-mote Sensing Data to Model PM₁₀ Estimation in Cities with a Low Number of Air Quality Stations: A Case of Study in Quito, Ecuador,” *Environments*, vol. 6, no. 7, p. 85, Jul. 2019. <https://doi.org/10.3390/environments6070085>
- [12] L. C. Messer; J. S. Jagai; K. M. Rappazzo; D. T. Lobdell, “Construction of an environmental quality index for public health research,” *Environ. Health*, vol. 13, no. 1, p. 39, May 2014. <https://doi.org/10.1186/1476-069X-13-39>
- [13] M. A. Musse; D. A. Barona; L. M. Santana Rodríguez, “Urban environmental quality assessment using remote sensing and census data,” *Int. J. Appl. Earth Obs. Geoinformation*, vol. 71, pp. 95–108, Sep. 2018. <https://doi.org/10.1016/j.jag.2018.05.010>
- [14] L. M. Santana Rodríguez; L. A. Escobar Jaramillo; P. A. Capote, “Estimación de un índice de calidad ambiental urbano, a partir de imágenes de satélite,” *Rev. Geogr. Norte Gd.*, no. 45, pp. 77–95, May. 2010. <http://dx.doi.org/10.4067/S0718-34022010000100006>
- [15] S. C. Grabich *et al.*, “Additive Interaction between Heterogeneous Environmental Quality Domains (Air, Water, Land, Sociodemo-graphic, and Built Environment) on Preterm Birth,” *Front. Public Health*, vol. 4, pp. 232, Oct. 2016. <https://doi.org/10.3389/fpubh.2016.00232>
- [16] H. Aguilar-Arias; R. Mora-Zamora; C. Vargas-Bolaños, “Metodología para la corrección atmosférica de imágenes Aster, Rapideye, Spot 2 y Landsat 8 con el módulo Flaash del software ENVI,” *Rev. Geográfica América Cent.*, vol. 2, no. 53, pp. 39–59, May. 2014. <https://doi.org/10.15359/rgac.2-53.2>
- [17] M. G. Martínez Mena, “Detección de cambio en reservorios acuíferos basados en el índice espectral de sequía,” (Tesis de Maestría), Universidad Politécnico de Madrid, Madrid, 2017. [URL](#)
- [18] C. Gonzaga Aguilar, “Aplicación de índices de vegetación derivados de imágenes satelitales Landsat 7 ETM+ y Aster para la caracterización de la cobertura vegetal en la zona centro de la provincia de Loja, Ecuador,” (Tesis de Maestría), Universidad Nacional de La Plata, Argentina, 2014. [URL](#)
- [19] L. M. Boix, “Integración de sensores y uso de índices de vegetación para evaluar la producción en plantaciones de cítricos en Málaga,” (Tesis Doctoral), Universidad de Málaga, España. 2016. [URL](#)
- [20] D. Rodríguez Pérez; N. Sánchez Carnero; J. A. Domínguez Gómez; C. Santa Marta Pastrana, *Cuestiones de Teledetección*. Madrid: Editorial UNED, 2015.
- [21] H. Xu, “Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery,” *Int. J. Remote Sens.*, vol. 27, no. 14, pp. 3025–3033, Jul. 2006. <https://doi.org/10.1080/01431160600589179>
- [22] L. Jiménez, “Estimación de la concentración del contaminante gaseoso NO_x en el Distrito Metropolitano de Quito, a partir de imágenes de satélite Landsat 7 ETM / Landsat 8 OLI de los años 2010, 2013 y 2015, considerando modelos estadísticos a partir de correlación entre estaciones de monitoreo atmosférico y los índices ambientales,” (Tesis de Maestría), Universidad de las fuerzas armadas, Espe, Sangolquí. 2018. [URL](#)
- [23] L. Nikroo; M. Kompani-Zare; A. R. Sepaskhah; S. R. F. Shamsi, “Groundwater depth and elevation interpolation by kriging methods in Mohr Basin of Fars province in Iran,” *Environ. Monit. Assess.*, vol. 166, no. 1, pp. 387–407, Jul. 2010. <https://doi.org/10.1007/s10661-009-1010-x>
- [24] M. Villatoro; C. Henríquez; F. Sancho, “Comparación de los interpoladores IDW Y Kriging en la variación espacial de pH, Ca, CICE y P del suelo,” *Agron. Costarric.*, vol. 32, no. 1, pp. 95–105, 2008. [URL](#)
- [25] D. S. Álvarez Osorio; A. Cárdenas Contreras; J. C. Matiz León, “Modelos digitales batimétricos generados por métodos de interpolación IDW, KRIGING, SHEPARD y B-SPLINE en el Archipiélago de Islas del Rosario,” *UD Geomática*, no. 5, Dec. 2011. <https://doi.org/10.14483/23448407.3641>
- [26] Á. León González; H. Llinás Solano; J. Tilano, “Análisis multivariado aplicando componentes principales al caso de los desplazados,” *Ing. Desarro.*, no. 23, pp. 119–142, Jan. 2008. [URL](#)
- [27] H. Ramírez Murillo; C. A. Torres-Pinzón; E. F. Forero-García, “Estimación del potencial fotovoltaico mediante minería de datos en cuatro ciudades de Colombia,” *Tecnológicas*, vol. 22, no. 46, Sep. 2019. <https://doi.org/10.22430/22565337.1345>

- [28] C. Chávez, S. Sánchez, and J. DelaCerde, “Análisis De Componentes Principales Funcionales En Series De Tiempo Económicas (Analysis of Principal Functional Components in Economic Time Series),” *Int. Gest. Conoc. Technol.*, vol. 3, no. 2, pp. 13–25, Dec. 2015. [URL](#)
- [29] A. Martínez Osorio, “Islas de calor en el área urbana del Valle de Aburrá,” (Tesis de Maestría), Universidad Nacional de Colombia-Sede Medellín, Medellín, 2017. [URL](#)
- [30] V. M. Rodríguez-Moreno; S. H. Bullock, “Comparación espacial y temporal de índices de la vegetación para verdor y humedad y aplicación para estimar LAI en el Desierto Sonorense,” *Rev. Mex. Cienc. Agríc.*, vol. 4, no. 4, pp. 611–623, Jun. 2013. [URL](#)