



REVISTA AIDIS

de Ingeniería y Ciencias Ambientales:
Investigación, desarrollo y práctica.

REGRESIÓN PONDERADA GEOGRÁFICAMENTE PARA EL ESTUDIO DE LA TEMPERATURA SUPERFICIAL EN MEDELLÍN, COLOMBIA

*Engelberth Soto Estrada¹

*GEOGRAPHICALLY WEIGHTED REGRESSION TO ANALYZE
THE LAND SURFACE TEMPERATURE OVER MEDELLIN,
COLOMBIA*

Recibido el 9 de agosto de 2013; Aceptado el 29 de noviembre de 2013

Abstract

Rural areas have different thermal properties than urban materials such as concrete and asphalt, which show higher evidence of the urban heat island (UHI) effect. The intensity of an UHI is related to the land cover configuration and physical factors affecting the urban terrain. In this research a multiple regression analysis was carried out to examine the spatial patterns of the UHI phenomena in Medellin, Colombia. Landsat 7 ETM+ from 2010 and WorldView-2 from 2011 and 2012 satellite images were used. The regression analysis was based on the Ordinary Least Squares and the Geographically Weighted methods using average information from the 344 of the 537 barrios situated in the Aburra valley, the place where Medellin is located. Results showed that the valley's rugged relief influences UHI spatial patterns and that this phenomena depends not only on the presence of hard surfaces or vegetation, but also on physical factors such as solar exposure and average height of the terrain.

Key Words: geographically weighted regression, satellite imagery, urban heat island.

¹ Departamento de Movilidad, Planeación Territorial y SIG, Escuela de Ingeniería de Antioquia. Colombia

*Autor correspondiente: Departamento de Movilidad, Planeación Territorial y SIG, Escuela de Ingeniería de Antioquia. Km 2 + 200 Vía al Aeropuerto José María Córdova, Envigado. Código Postal 055428. Colombia. Email: pfesoto@eia.edu.co

Resumo

Las superficies urbanas se caracterizan por su alta capacidad para absorber calor y liberarlo lentamente originando el efecto conocido como “isla de calor urbana”. La intensidad de este fenómeno depende de la composición de coberturas en el entorno urbano y de otros factores físicos que inciden en la superficie. En esta investigación se llevó a cabo un análisis de regresión con múltiples variables para explicar los patrones espaciales de la isla de calor que tiene lugar en la ciudad de Medellín, Colombia. Para ello se utilizaron imágenes satelitales Landsat 7 ETM+ de 2010 y WorldView-2 de 2011 y 2012. El análisis de regresión se realizó mediante dos métodos: mínimos cuadrados ordinarios y ponderación geográfica empleando datos de 344 de los 537 barrios asentados en el valle de Aburrá, sitio donde se localiza Medellín. Los resultados sugieren que la morfología del valle tiene influencia en la distribución de temperatura superficial urbana y que dicha distribución depende no solamente de la presencia de superficies artificiales o vegetación, sino también de condiciones físicas como la exposición solar de las superficies y la altura de éstas sobre el nivel medio del mar.

Palabras clave: imágenes satelitales, isla de calor urbana, regresión ponderada geográficamente.

Introducción

Las islas de calor urbanas son incrementos en la temperatura atmosférica y superficial de áreas pobladas, en comparación con la temperatura de las zonas suburbanas o rurales que las rodean. Este fenómeno se origina cuando las superficies naturales son reemplazadas o aisladas por materiales artificiales como concreto y asfalto, los cuales tienen una baja reflectividad solar y una alta capacidad para absorber calor y liberarlo lentamente (USEPA, 2013).

Las islas de calor generan diversos impactos ambientales, entre ellos: el incremento del consumo energético por el uso intensivo de sistemas de refrigeración y la disminución en la eficiencia de los motores de combustión interna (Akbari, 2005), la degradación de la calidad del aire por el consumo de energía y la formación de ozono a nivel de superficie, la polución térmica del agua de lluvia a su paso sobre superficies calientes, la aparición de enfermedades relacionadas con las altas temperaturas y, en general, la reducción de la calidad de vida de la población.

Debido a ello, el estudio de este fenómeno y la implementación de estrategias para combatirlo son una prioridad en diversas ciudades del mundo como Nueva York (NYSERDA, 2006), Montreal (CCAP, 2007), Las Vegas (CofLV, 2010) y Tokio (Yamamoto, 2006). En las ciudades de países en vías de desarrollo, como Medellín, la aparición de islas de calor puede tener un impacto mayor debido a que cuentan con una mayor cantidad de personas expuestas a las variaciones de temperatura. En el grupo de habitantes vulnerables se incluyen las personas que habitan viviendas construidas con materiales que no poseen propiedades aislantes, aquellos que no cuentan con recursos económicos para proveerse de equipos de refrigeración y los que tienen acceso limitado a los servicios de salud.

El objetivo de esta investigación fue estimar la magnitud de la isla de calor superficial que tiene lugar en el área metropolitana asentada en el valle de Aburrá, en el departamento colombiano de Antioquia, y posteriormente identificar los factores que influyen en sus patrones de distribución espacial mediante análisis de regresión. Con los análisis de regresión fue posible examinar la relación que existe entre la variable de interés, en este caso la temperatura superficial urbana, y posibles variables explicativas. Para ello se utilizaron los métodos de mínimos cuadrados ordinarios (OLS, por sus siglas en inglés) y ponderación geográfica (GWR, por sus siglas en inglés).

OLS es un modelo global que se basa en el uso de una sola ecuación para explorar la relación entre las variables. En este modelo se asume que la relación es consistente en toda el área de estudio. GWR es un modelo local que crea una ecuación para cada elemento del conjunto de datos de la variable dependiente. En este modelo se asume que la relación entre las variables varía dependiendo de su ubicación en el área de estudio. Adicionalmente, en el modelo GWR se considera que los valores cercanos al punto de análisis tienen un mayor peso para la estimación de los parámetros de regresión. La ponderación geográfica se determina mediante un esquema de ponderación conocido como *kernel*, el cual puede basarse en una distancia de análisis o en un número de datos vecinos (basado en Charlton & Fotheringham, 2009).

A través este estudio se espera construir un modelo confiable mediante el cual se puedan tomar decisiones sobre las acciones de mitigación de temperatura superficial en zonas conflictivas del área urbana mencionada. En este artículo se presentan los resultados del análisis de regresión utilizando las variables explicativas vegetación, cobertura de superficies artificiales, altura y exposición solar del terreno.

En la siguiente sección se desglosa la metodología de la investigación. En el tercer apartado se presentan y discuten los resultados, mientras que en la cuarta se incluyen las conclusiones y reflexiones finales.

Metodología

En esta investigación se utilizó una imagen satelital Landsat 7 ETM+ tomada el 21 de enero de 2010, dos imágenes WorldView-2 de 2011 y 2012, un modelo digital de elevación (MDE) de 30 m de resolución y otras capas con información vectorial del área de estudio. Para el procesamiento de la información se emplearon los programas informáticos ArcGIS 10.1, ERDAS 9.1 y Geomatica 2013.

La isla de calor urbana se estimó utilizando la imagen Landsat. Luego de ajustar sus características geométricas, radiométricas y eliminar las áreas nubosas, se transformaron los números digitales de la capa térmica a magnitudes de irradiación y posteriormente a grados

Kelvin, para finalmente convertirlos a grados centígrados. Las Ecuaciones (1) y (2) y la Tabla 1 muestran las variables y valores considerados en el cálculo de temperatura superficial (NASA, 2007).

$$T = \frac{K2}{\ln\left(\frac{K1}{L\lambda} + 1\right)} \quad \text{Ecuación (1)}$$

Donde

T: Temperatura en grados Kelvin

K1: Constante de calibración 1

K2: Constante de calibración 2

Lλ: Irradiación espectral en la apertura del sensor en watts/(m² * ster * μm)

El valor de las constantes de calibración K1 y K2 considerado fue 666.09 y 1282.71, respectivamente (NASA, 2011).

Por otra parte,

$$L\lambda = G_{rescale} * Q_{CAL} + B_{rescale} \quad \text{Ecuación (2)}$$

Donde

G_{rescale}: ganancia ajustada en (W/m² * sr * μm)/DN

B_{rescale}: sesgo ajustado en W/(m² * sr * μm)

Q_{CAL}: valor discreto calibrado del pixel en DN

El valor de las constantes de ajuste para la radiancia espectral se incluye en la Tabla 1.

Tabla 1. Constantes de ajuste para la radiancia espectral de la banda térmica de imágenes LANDSAT

	G _{rescale} (W/m ² * sr * μm)/DN	B _{rescale} W/(m ² * sr * μm)
Landsat 7 – Low gain	0.067087	-0.07
Landsat 7 – High gain	0.037205	3.16

FUENTE: (NASA, 2011)

Para el análisis de regresión se calculó, como una primera aproximación al estudio del fenómeno, el índice de vegetación de diferencia normalizada (NDVI, por sus siglas en inglés) utilizando la imagen Landsat. Este indicador da cuenta de la densidad de vegetación en el terreno, factor que ha sido relacionado con la intensidad de la isla de calor urbana (USEPA, 2013). Posteriormente, mediante las imágenes satelitales de alta resolución se estimó la cobertura de vegetación (discriminándola en vegetación alta y baja) y la cobertura de superficies artificiales. Estas variables brindaron datos más precisos de la composición de coberturas en el área urbana con los que es posible planear intervenciones futuras para el desarrollo de la ciudad.

Las imágenes WorldView-2 se clasificaron mediante un procedimiento supervisado basado en la ecuación de distancia de Mahalanobis (ERDAS, 1997), el cual fue verificado utilizando 250 puntos de control localizados aleatoriamente. La Tabla 2 muestra el porcentaje de precisión global y el valor del coeficiente Kappa encontrados al aplicar tres reglas paramétricas en el proceso de clasificación.

Tabla 2. Precisión de las clasificaciones supervisadas

Imagen	Fecha	Regla paramétrica	Precisión global (%)	Coficiente Kappa
Landsat 7 ETM	21/1/2010	Maximum Likelihood	87.20	0.7530
		Mahalanobis Distance	88.80	0.7771
		Minimum Distance	77.20	0.5807
WorldView-2	20/2/2012	Maximum Likelihood	68.00	0.5168
	2/9/2011	Mahalanobis Distance	78.00	0.6169
		Minimum Distance	70.90	0.5564

Posteriormente, se estimó la altura promedio (en metros sobre el nivel medio del mar) y la exposición solar del terreno (considerada como el inverso de la presencia de sombras). La altura promedio se determinó utilizando el MDE de 30 X 30 m (JLP, 2004). La exposición solar se calculó utilizando el MDE y la posición del sol el día y la hora en que la imagen Landsat fue tomada (alrededor de las 10 de la mañana). Dicha posición se estableció mediante la calculadora solar del Servicio Nacional de Meteorología de Estados Unidos de América (NOAA, 2012).

Como se mencionó anteriormente, el análisis de regresión se realizó mediante los modelos OLS y GWR. Para ello se consideraron valores promedio de 344 de los 537 barrios incluidos en la base de datos del área metropolitana del valle de Aburrá, incluyendo todos los barrios de la ciudad de Medellín. Para la ponderación geográfica se consideró una función de *kernel* adaptativa, la cual se recomienda para valores que se distribuyen de manera irregular (Charlton

& Fotheringham, 2009), que es nuestro caso pues los barrios varían en tamaño y forma. Para establecer el umbral de análisis se seleccionó el Criterio de Información de Akaike (AICc). En la figura 1 se puede apreciar la extensión (aproximadamente 20 X 30 km) y las características topográficas del área de estudio.

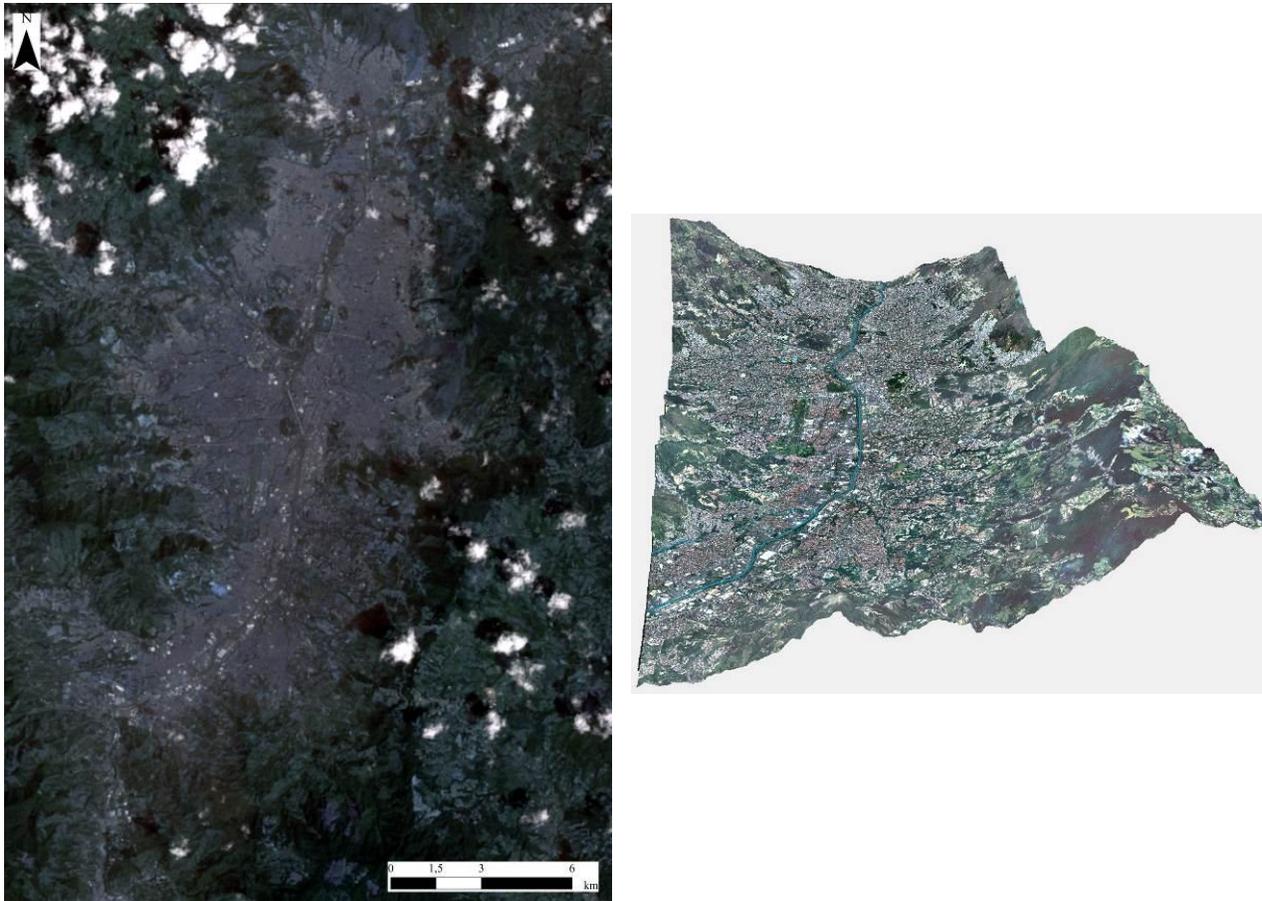


Figura 1. Extensión del área de estudio y perspectiva de la zona urbana en 3D

Discusión de resultados

Los mapas de temperatura superficial (izquierda) y densidad de vegetación (derecha) se presentan en la Figura 2. En ellos se puede apreciar a simple vista el polígono que conforma el área urbana localizada en el valle de Aburrá, así como el río que da nombre al valle y fluye de sur a norte atravesando la parte central de la región habitada. Entre estas dos variables se estimó un coeficiente de correlación de -0.3917 , lo que indica una relación inversa que es congruente con lo que se aprecia en los mapas: a menor presencia de vegetación (áreas en color café en el mapa de la derecha), mayor temperatura superficial en el mapa de la izquierda.

No obstante, la regresión OLS con estas variables arrojó un rendimiento del modelo (R^2) de 0.0642, lo que indica que la temperatura superficial no puede ser explicada únicamente por la presencia de vegetación. En un valle estrecho como el de Aburrá (ver Figura 1) otros factores, como la variación de altura del área habitada y la exposición a la radiación solar, podrían tener una influencia significativa en la temperatura de las superficies urbanas.

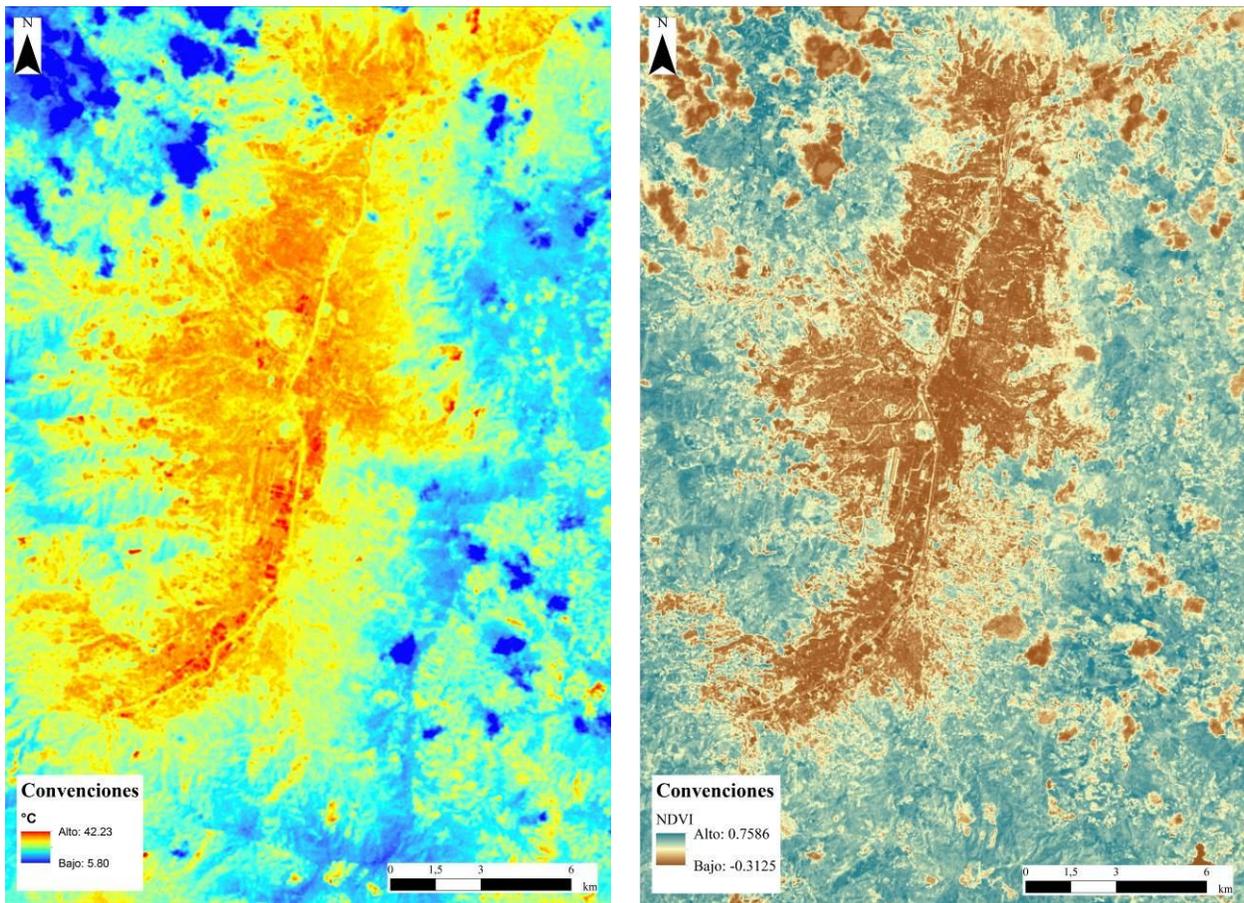


Figura 2. Temperatura superficial y densidad de vegetación el 21 de enero de 2010

La Tabla 3 muestra el rendimiento de las regresiones OLS al analizar los valores de temperatura superficial contra las variables explicativas. Los valores de R^2 representan el porcentaje en que cada una de las variables puede explicar la variable dependiente. En principio, se puede afirmar que la temperatura superficial en el área de estudio no puede ser explicada por una sola variable.

Tabla 3. Rendimiento de las regresiones OLS

Variable explicativa	Rendimiento del modelo (R^2)
NDVI	0.0642
Altura	0.3742
Exposición solar	0.5605
Vegetación baja	0.2966
Vegetación alta	0.2451
Superficies artificiales	0.4024

Las regresiones OLS indicaron que la exposición solar y la cobertura de superficies artificiales son las variables que mejor explican, de manera individual, el fenómeno de temperatura superficial (56 y 40% de los valores, respectivamente). Nótese que la vegetación, estimada mediante el indicador NDVI, sólo puede explicar el 6% de los valores de temperatura superficial en el área urbana. Si ésta se considera como cobertura (ya sea vegetación baja o alta) este porcentaje sube hasta casi 30%. En la Figura 3 se puede observar gráficamente la relación existente entre la temperatura superficial y las variables explicativas.

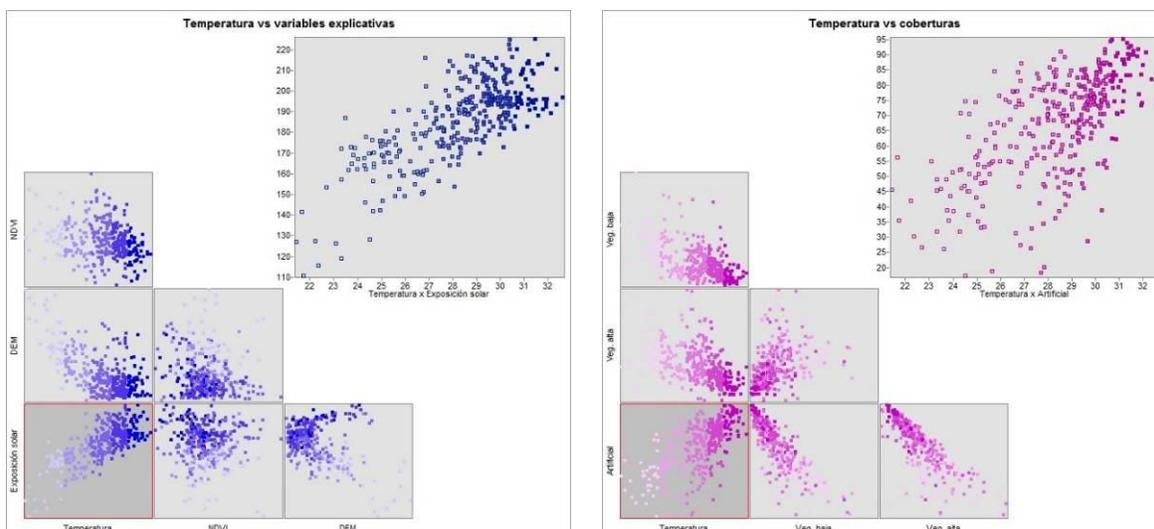


Figura 3. Matrices de distribución de temperatura superficial vs. variables explicativas

Luego de examinar las variables de manera individual se analizaron diferentes combinaciones de factores para determinar el modelo que mejor explicara los patrones de distribución de temperatura superficial. El rendimiento de cada combinación se analizó utilizando el indicador R^2 ajustado (el cual genera valores de 0.0 – 1.0), el criterio de información de Akaike corregido

(entre menor sea el valor del indicador AICc mejor será el modelo) y el factor de inflación de la varianza VIF (mayor a 7.5 indica que la variable es redundante y debe retirarse del modelo).

El modelo explicativo compuesto por la exposición solar, la intensidad de vegetación (NDVI) y la cobertura de superficies artificiales fue el que arrojó un mejor rendimiento. Esta combinación pudo explicar el 74% de los valores de temperatura superficial. Este modelo se ajustó localmente, los resultados se muestran en la Tabla 4.

Tabla 4. Indicadores de rendimiento de la correlación GWR para el modelo explicativo seleccionado

Indicador	Valor
Neighbors	56
ResidualSquares	147.25
EffectiveNumber	56.56
Sigma	0.7157
AICc	785.94
R2	0.9175
R2Adjusted	0.9016

Como se aprecia en la Tabla 4, el modelo explicativo, ajustado localmente, pudo explicar el 90% de los valores de la variable dependiente, lo que demuestra que la temperatura superficial es función de estas variables. A pesar de estos resultados, el modelo no pudo ser utilizado para predecir valores de temperatura superficial en Medellín. Para ello es necesario que los valores residuales de la regresión, es decir, la diferencia entre los valores ajustados y los observados, se distribuyan aleatoriamente de manera normal, lo que no sucedió con el modelo explicativo.

La distribución de valores residuales se examinó utilizando la prueba Jarque Bera (JB) y el índice I de Moran. Cuando el valor p (probabilidad) de la prueba JB es menor a 0.05 para un nivel de confianza del 95% los residuales no se distribuyen normalmente. El modelo seleccionado arrojó un valor de probabilidad de 0.000144, lo que indica que la distribución de residuales no sigue una distribución normal. Por su parte, el índice I de Moran determina si el patrón de distribución es agrupado, disperso o aleatorio. Cuando los valores de los residuales tienen a agruparse espacialmente (es decir, valores altos con valores altos y bajos con bajos) el índice es positivo. En este caso, el modelo generó un índice de 0.5413, lo que significa que la distribución de residuales es agrupada.

La distribución agrupada de valores residuales puede ser el resultado de errores en el diseño del modelo explicativo (¿se eligieron las variables correctas?) o por la falta de información (¿las variables utilizadas tienen al menos 30 datos?). Se considera que en esta primera etapa de la

investigación se eligieron variables relevantes para el fenómeno examinado, con un buen número de observaciones, aunque es probable que existan otros factores influyendo en la distribución de temperatura superficial urbana como el porcentaje de humedad o el viento al que está expuesta la superficie.

Todas las combinaciones de variables que se plantearon arrojaron distribuciones de residuales agrupadas. Esto puede indicar que el comportamiento de la temperatura superficial está gobernado, al menos a la hora del día en que se tomó la imagen satelital, por los aspectos físicos propios de cada sector del área urbana (principalmente la exposición solar que no se distribuye aleatoriamente en el territorio) y no por la composición de coberturas. Lo anterior puede apreciarse en la distribución de coeficientes del modelo explicativo (ver Figura 4), la cual muestra una estructura que coincide con las laderas del valle.

Esta distribución también se observa al representar en un mapa la temperatura promedio de los barrios examinados (ver Figura 4). Resulta evidente que los barrios localizados en la ladera este del valle de Aburrá (lado derecho en la figura del lado derecho) presentan menores temperaturas que los de la ladera oeste. Esto puede deberse a que el sol a esta hora del día ha calentado más un extremo del valle que el otro, lo que refuerza la hipótesis de que los aspectos físicos analizados en esta investigación no cuentan con una distribución normal en el territorio y por lo tanto no pueden producir un modelo confiable para realizar predicciones.

Conclusiones

Con base en los resultados se concluye que los patrones de distribución de la isla de calor urbana que tiene lugar en el valle de Aburrá son función de varios factores como la presencia de vegetación, la cobertura de superficies artificiales y la exposición solar de las superficies. No obstante, la distribución de valores de la variable dependiente no permite la construcción de un modelo de predicción consistente y preciso que sirva para analizar escenarios de desarrollo urbano. A pesar de ello, mediante esta investigación se pudo identificar zonas urbanas críticas considerando su temperatura superficial promedio.

Para construir un modelo de predicción es necesario contar con datos de temperatura superficial y exposición solar representativos, esto es, cuanto las superficies reciben la misma cantidad de energía solar. Para obtener esta información se utilizará un vehículo aéreo no tripulado (UAV, por sus siglas en inglés). Además, en la siguiente fase de la investigación se considerarán nuevos datos, como la estimación de la cantidad acumulada de energía solar o la humedad, así como unidades homogéneas de análisis espacial que no dependan del tamaño del barrio.

Por último, se deberá avanzar en el análisis estadístico espacial considerando diferentes condiciones de la función *kernel*, analizando la relación entre las variables independientes y logrando un mayor entendimiento de la distribución de coeficientes.

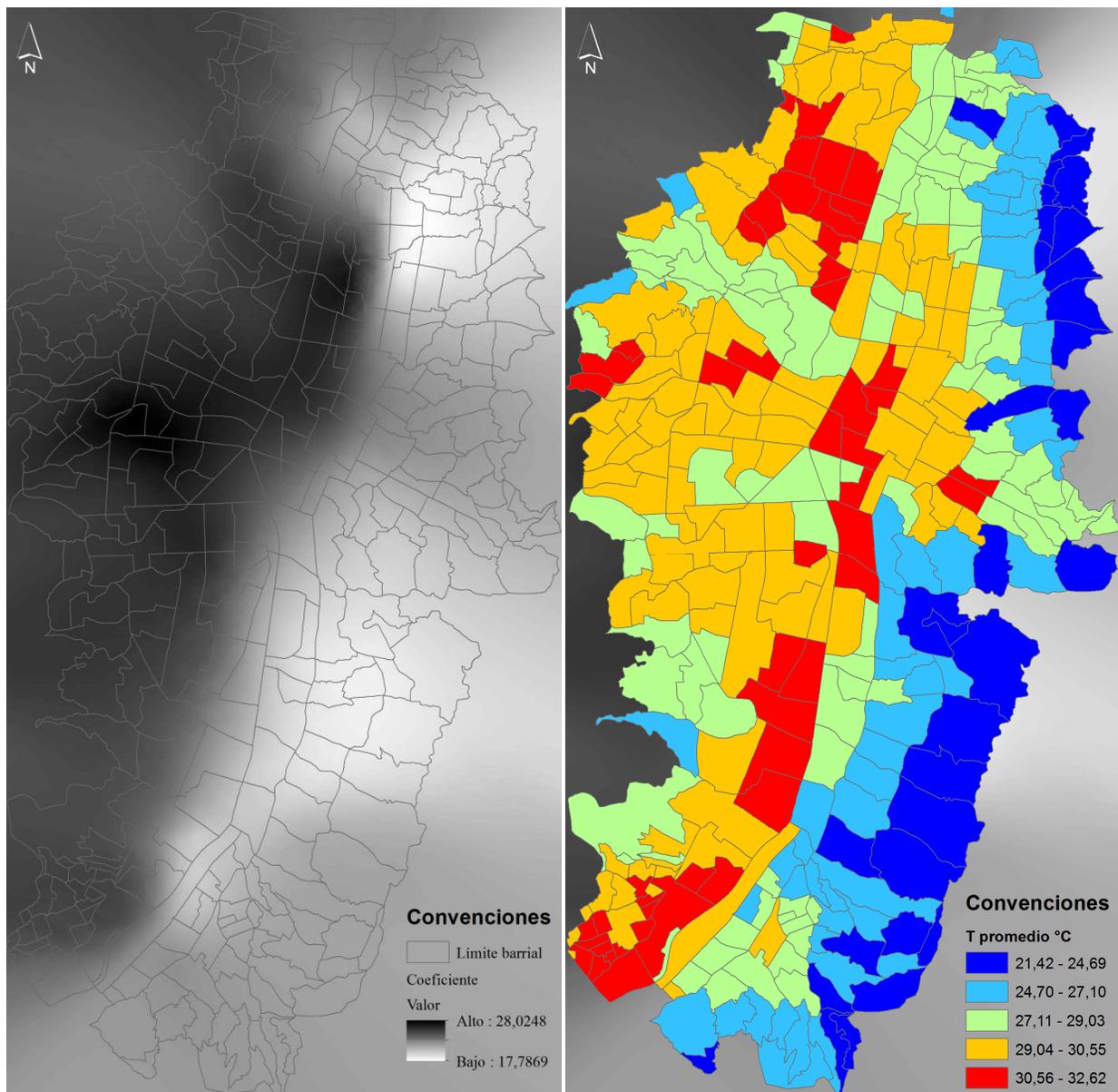


Figura 4. Distribución de coeficientes del modelo explicativo (izquierda) y temperatura superficial promedio por barrio alrededor de las 10 de la mañana (derecha)

Referencias bibliográficas

- Akbari, H., (2005) Energy Saving Potentials and Air Quality Benefits of Urban Heat Island Mitigation. Lawrence Berkeley National Laboratory, p. 19.
- CCAP, (2007) Urban Heat Islands: A Climate Change Adaptation Strategy for Montreal, Montreal: Services des infrastructures, transport et Environnement of the City of Montreal.
- Charlton, M. & Fotheringham, S., (2009) Geographically Weighted Regression, White Paper. National Centre for Geocomputation. Maynooth, Co Kildare, Ireland, pp. 1-13.
- CofLV, 2010. Summary Report Urban Heat Island Effect, City of Las Vegas: Sustaining Las Vegas.
- ERDAS, (1997) ERDAS Field Guide. Cuarta edición ed. Atlanta: ERDAS Inc.
- JLP, (2004) ASTER Global Digital Elevation Map. [En línea] Disponible en: <http://asterweb.jpl.nasa.gov/gdem.asp> [Último acceso: 25 9 2012].
- NASA, (2007) Landsat 7 Science Data Users Handbook. US National Aeronautics and Space Administration.
- NASA, (2011). Landsat 7 Science Data Users Handbook. [En línea] Disponible en: <http://landsathandbook.gsfc.nasa.gov/>
- NOAA, (2012) NOAA Solar Calculator. [En línea] Disponible en: <http://www.esrl.noaa.gov/mgd/grad/solcalc/> [Último acceso: 3 10 2012].
- NYSERDA, (2006) Mitigating New York City's Heat Island with Urban Forestry, Living Roofs, and Light Surfaces, New York: New York State.
- USEPA, (2013) Heat Island Effect - Basic Information. [En línea] Disponible en: www.epa.gov [Último acceso: 10 7 2013].
- Yamamoto, Y., (2006) Measures to Mitigate Urban Heat Islands. Quarterly Review, Issue 18, pp. 6 -83.