

CLASIFICACIÓN ESPACIAL DEL SUELO URBANO POR EL VALOR ESPECULATIVO DEL SUELO

E IMÁGENES MSI SATELITALES USANDO K-MEANS,
HUANCAYO, PERÚ

SPATIAL CLASSIFICATION OF URBAN LAND BY SPECULATIVE LAND VALUE
AND MSI SATELLITE IMAGERY USING K-MEANS, IN HUANCAYO, PERU

70

GONZALO RODOLFO PEÑA ZAMALLOA 1

1 Magíster en Ciencias Sociales
Universidad Continental, Huancayo, Perú
Profesor investigador
<https://orcid.org/0000-0002-6141-6849>
gzamalloa@outlook.com



La ciudad de Huancayo, como otras ciudades intermedias en Latinoamérica, enfrenta problemas de cambios de uso de suelo poco planificados y una acelerada dinámica del mercado del suelo urbano. La escases y desactualización de información sobre el territorio urbano impiden la adecuada clasificación de áreas urbanas, limitando la forma de su intervención. Esta investigación tuvo como objetivo la incorporación de métodos no asistidos y mixtos para la clasificación espacial de zonas urbanas considerando el valor especulativo del suelo, la proporción del suelo urbanizado y otras variables geoespaciales. Entre los medios de recolección de datos, se usó imágenes Multi-Espectrales (MSI) del satélite Sentinel-2, el sistema vial primario y una muestra de puntos de observación directa. Los datos procesados fueron incorporados en mapas georreferenciados, a los cuales se añadió además los límites urbanos y pendientes oficiales. Durante el procesamiento de los datos se empleó el algoritmo K-Means, junto a otros métodos de *machine learning* y juicio asistido. Como resultado, se obtuvo una caracterización objetiva de zonas urbanas que difiere de la planificación existente.

Palabras clave: planificación urbana, mercado inmobiliario, periferia urbana, inteligencia artificial

The city of Huancayo, like other intermediate cities in Latin America, faces problems of poorly planned land-use changes and a rapid dynamic of the urban land market. The scarce and outdated information on the urban territory impedes the adequate classification of urban areas, limiting the form of its intervention. The purpose of this research was the adoption of unassisted and mixed methods for the spatial classification of urban areas, considering the speculative land value, the proportion of urbanized land, and other geospatial variables. Among the data collection media, Multi-Spectral Imagery (MSI) from the Sentinel-2 satellite, the primary road system, and a sample of direct observation points, were used. The processed data were incorporated into georeferenced maps, to which urban limits and official slopes were added. During data processing, the K-Means algorithm was used, together with other machine learning and assisted judgment methods. As a result, an objective classification of urban areas was obtained, which differs from the existing planning.

Keywords: urban planning, real estate market, urban periphery, artificial intelligence.

I. INTRODUCCIÓN

Las ciudades intermedias enfrentan procesos de especulación de valor del suelo y venta de terrenos, que definen la forma urbana con mayor rapidez que la intervención de los gobiernos locales. El valor del suelo urbano prioriza la demanda de agentes privados (Gasic, 2018), sin embargo, el Estado determina los límites de esta actuación (Sabatini y Arenas, 2000). Así, la segregación socio-espacial se relaciona con la disposición del mercado y las políticas poco adecuadas, moldeando el exceso de demanda u oferta del mercado inmobiliario y generando una dispersión en el estándar de vida urbano (Saleh, Hwa y Majid, 2016; López Navarrete y Peña Medina, 2017; Li, Sun, y Boersma, 2019). Ante ello, la regulación del suelo y su aplicación pueden promover o detener el desarrollo de zonas urbanas emergentes (Yu, Zhou y Yang, 2019). Estas condiciones son comunes en ciudades latinoamericanas de limitada acción pública (Sabatini y Arenas, 2000). Entre las causas de ello, se reconoce prácticas clientelistas, producto de fallas en las prácticas de libre mercado, intereses particulares, condiciones de ilegalidad, reglamentaciones ambiguas y una generalizada aceptación popular (Pimentel Sánchez, 2020; Espinoza y Fort, 2017).

El valor de suelo, el indicador más importante de la dinámica del mercado inmobiliario, no es fácil de estimar o predecir, aunque es común que las áreas consolidadas se encarezcan, volviendo la periferia más atractiva por su bajo precio (Glaeser y Ward, 2009; García y Peralta, 2016; Gasparenienea, Venclauskienea y Remeikienea, 2014). En el largo plazo, las conductas del mercado del suelo pueden aproximarse con series de tiempo (Gaete, 2021). No obstante, una aproximación con datos heterogéneos, o bien, escenarios de alta incertidumbre pueden usar la inteligencia artificial para clasificarlos (Durduran, 2015; Belhadia *et al.*, 2020; Forestier y Wemmer, 2016). El algoritmo K-Means ha resultado útil y altamente adaptable para la clasificación de imágenes, el estudio del crecimiento urbano y el análisis espacial (Liu *et al.*, 2021; Belhadia *et al.*, 2020).

El Perú atraviesa su bicentenario y enfrenta grandes retos económicos y sociales. La Política Nacional de Vivienda y Urbanismo considera al bajo impacto de la planificación urbana-territorial y la limitada práctica de cumplimiento normativo un problema mayor (Ministerio de Vivienda Construcción y Saneamiento [MVCS], 2021). Los procesos del mercado inmobiliario, de naturaleza formal e ilegal o bajo figuras mixtas, son también frecuentes en ciudades peruanas (Espinoza y Fort, 2017; Pimentel Sánchez, 2020). Aun cuando la cartera de fondos de apoyo para la vivienda, como el Fondo Mivivienda (FMV), se multiplican, su implementación se limita por las adversas condiciones urbanas y el valor del suelo (Calderón, 2015). Durante el auge inmobiliario de 2018 a 2019, al menos el 70% de municipios distritales no

contaba con planes de desarrollo urbano (FMV, 2018a, FMV, 2018b). En Huancayo, la principal ciudad del centro del Perú, la Municipalidad Provincial de Huancayo (2016) propone un desarrollo fundado en principios sostenibles e inclusivos, pero que requiere del conocimiento de la realidad urbana local y su caracterización objetiva.

Este artículo se plantea cómo clasificar espacialmente las zonas urbanas en la ciudad de Huancayo a partir de datos heterogéneos. La investigación propone la clasificación diferenciada de zonas urbanas incorporando métodos no asistidos y mixtos, y considerando el valor especulativo del suelo en el mercado inmobiliario, la proporción de suelo urbanizado, la distancia al viario principal y la pendiente del terreno. El trabajo se desarrolló en cuatro etapas articuladas: (1) construcción de mapas base; (2) procesamiento de imágenes satelitales para el análisis de la ocupación del suelo actual; (3) aplicación de métodos de machine learning para clasificación; y (4) caracterización poligonal de las zonas urbanas en la ciudad de Huancayo.

II. MARCO TEÓRICO

La planificación urbana y el mercado inmobiliario

Debido al crecimiento poblacional, una mejor planificación de las ciudades representa un problema continuo en todo el mundo (Mouratidis, 2021). Esta temática ha capturado la atención nacional para promover su desarrollo desde un enfoque sostenible (Aceid y Fundación ACS, 2018; Naciones Unidas, 2018; Castillo-García, 2021), aunque durante la pandemia su reducida presencia fue resaltante (Moreno, Allam, Chabaud, Gall y Pratlong, 2021). Se necesita, en este contexto, una revisión de la idea de la proximidad en la economía urbana vinculada a la generación de valor del suelo (Tricarico y De Vidovich, 2021).

La planificación urbana requiere un equilibrio entre el uso del suelo y la expansión urbana, no siempre alineada a los modos de vida reales y la conducta del mercado inmobiliario (López Navarrete y Peña Medina, 2017). Existe una brecha entre la generación sostenible del espacio urbano y las prácticas reales en las zonas periurbanas que colindan con áreas rurales y espacios naturales, rápidamente devastados por procesos de urbanización formales e informales (Carvajal, Moreira, Salazar, Leguia y Jorquera, 2019).

La segregación socio-espacial se relaciona con la disposición del mercado inmobiliario y políticas poco adecuadas, y afecta directamente la planificación urbana (López Navarrete y Peña Medina, 2017; Glaeser y Ward, 2009; Migueltorena y Lan, 2013). El exceso de demanda u oferta del mercado inmobiliario y la dispersión generan variaciones en el estándar de vida (Saleh et

al., 2016). La fluctuación del valor del suelo, el crecimiento urbano y la densidad inicial condicionan dichas variaciones (Glaeser y Ward, 2009; Li *et al.*, 2019).

En una etapa de crecimiento del sector, muchas de las reglas que dirigen las acciones del mercado no son fáciles de adaptar a los instrumentos de gestión, ampliando las brechas en la planificación urbana (Glaeser y Ward, 2009). Entre estas, la reglamentación del acceso a servicios urbanos formales es la que más se ha agravado (Baer y Kauw, 2016). Estas disparidades pueden llegar a ser insalvables, repercutiendo también en la generación de nuevas políticas y la recaudación tributaria (Hindi, Moreira y Rossi, 2020; Foldvary y Minola, 2017). Además, el valor del suelo tiene mayor variabilidad que las edificaciones (Kok, Monkkonen y Quigley, 2014). Debido a ello, este valor no puede asignarse idóneamente para su uso en las reglamentaciones, hipotecas y préstamos, pues su fluctuación real está caracterizada por la especulación (Hwang, Park y Lee, 2013; Gaspareniena *et al.*, 2014; Foldvary y Minola, 2017).

La necesidad de suelo urbanizable y el valor del suelo

Contar con una vivienda es uno de los aspectos más importantes en la vida de las personas (Saleh *et al.*, 2016). Aquellas financiadas con fondos sociales promueven la inversión inmobiliaria y tienen gran interés por la disponibilidad de terrenos sin urbanizar (Scotiabank, 2015; FMV, 2018c). Pero el valor del suelo urbano vacante está sujeto a especulación, y presenta menores precios en áreas periurbanas (Gedal y Ellen, 2018; Parias, 2008), lo cual promueve exclusiones al incrementarse la distancia física y de precio (Gaete, 2021; Klafus, Van Lindert, Van Noorloos y Steel, 2017). De esta forma, se ingresa en un ciclo incremental del valor especulado que impide el alcance de condiciones más homogéneas (Amézquita, Rodríguez y Murillo, 2015; Gaete, 2021; Gasic, 2018; Araque Solano y Caballero Quintero, 2009; Glaeser y Gyorko, 2003).

El nivel de consolidación y la cercanía a vías constituyen indudables atractivos del suelo urbano (Peña-Zamalloa, 2018; Gedal y Ellen, 2018). Los espacios agrícolas con cierto acceso viario son, por lo tanto, objetivos de cambio de uso de suelo (Salazar, 2014; Cardó, 2017; Migueltorena y Lan, 2013). Al tratarse de lotes informales, se prioriza sobre ellos la autoconstrucción, con el objetivo de reducir los costos de vivienda sin medir los efectos a largo plazo (GRADE, 2020; Salazar, 2014). La búsqueda continua de mayores ganancias por venta del suelo socava la intención popular positiva de hacer ciudad (Delgadillo, 2016; Araque Solano y Caballero Quintero, 2009). La rápida variación de precios genera, en el territorio, una ocupación desorganizada y de baja densidad, aun cuando en ellos se promueve la vivienda social (Calderón, 2015). Ello afecta al entorno rural, natural

y dificulta el acceso a servicios urbanos de áreas dispersas, así como el cumplimiento regulatorio (Carvajal *et al.*, 2019; Li *et al.*, 2019). Esta realidad se opone a la idea de ciudad compacta (Vorontsova, Vorontsova y Salimgareev, 2016).

La expansión urbana de baja densidad involucra altos costos en infraestructura urbana (Nabil y Eldayem, 2015). Frente a ese panorama, un modelo urbano óptimo prioriza la accesibilidad y se traduce en distancias cortas a múltiples centros urbanos y en una reducción del tiempo de movilidad (Yu *et al.*, 2019; Gedal y Ellen, 2018; Graells-Garrido, Serra, Rowe, Cucchiatti y Reyes *et al.*, 2021). La idea de un urbanismo cronológico es, en efecto, un intento por mejorar la calidad de vida de los habitantes en diversas escalas geográficas (Moreno *et al.*, 2021; Graells-Garrido *et al.*, 2021).

La caracterización del suelo urbano y el método K-Means

Los procesos expansivos urbanos poco comprendidos, sin un análisis sistémico articulado, se dan de forma dispersa y en oposición a la capacidad de generar ciudades compactas (Vorontsova *et al.*, 2016; Alfasi y Migdalovich, 2020). Adicionalmente, las métricas suelen ser unidimensionales (Tellier, 2020), cuando la complejidad urbana requiere el uso de métricas de análisis multidimensional para su clasificación (Steurer y Bayr, 2020; Tellier, 2020). El *Machine learning* ofrece una alternativa para la clusterización a partir de datos heterogéneos (Joshi, 2020). Esta clasificación puede ser asistida, no asistida o mixta (Liu *et al.*, 2018; Steurer y Bayr, 2020).

En concreto, K-Means es uno de los algoritmos de clasificación no supervisada más usado en imágenes, datos aleatorios y no etiquetados (Liu *et al.*, 2018; Zhou *et al.*, 2017). Este algoritmo permite la generación de clústeres, agrupando los datos bajo características similares (Campesato, 2020) y diferenciando elementos como vegetación, espacios urbanos vacantes e, incluso, usos rurales (Feng, Peng y Wu, 2020). Aunque el análisis jerárquico de clúster, la media móvil y maximización de la expectativa podrían considerarse alternativas apropiadas, el uso de una distancia euclidiana permite que la clasificación hecha con K-Means pueda superponerse a coordenadas bidimensionales y sea adecuada para entornos geográficos (Campesato, 2020; Joshi, 2020; Liu *et al.*, 2018).

III. ESTUDIO DE CASO

La ciudad de Huancayo se encuentra en la parte central del país. Su geografía es moldeada por el río Mantaro y constituye uno de los más amplios valles en los Andes

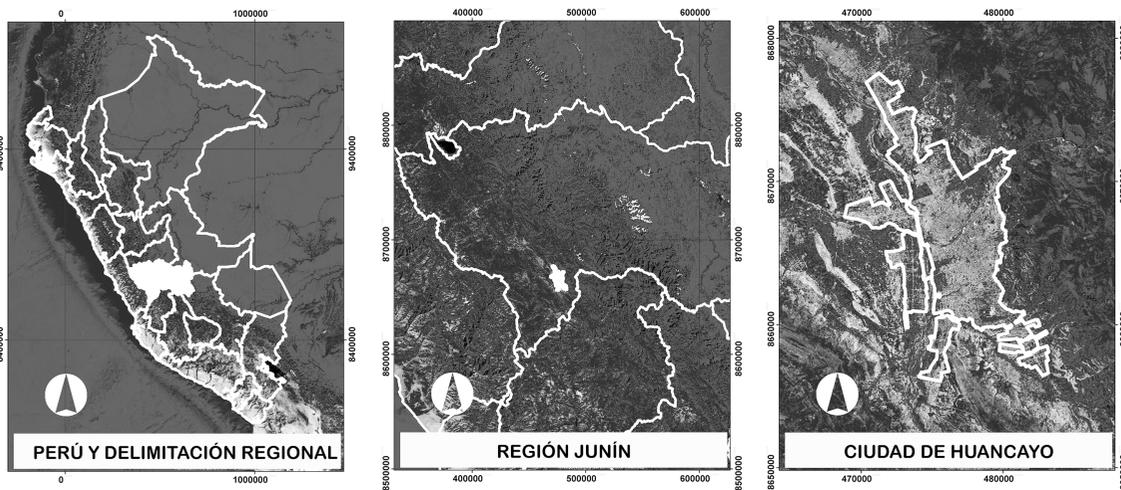


Figura 1. Ubicación del área de estudio. Fuente: Elaboración del autor

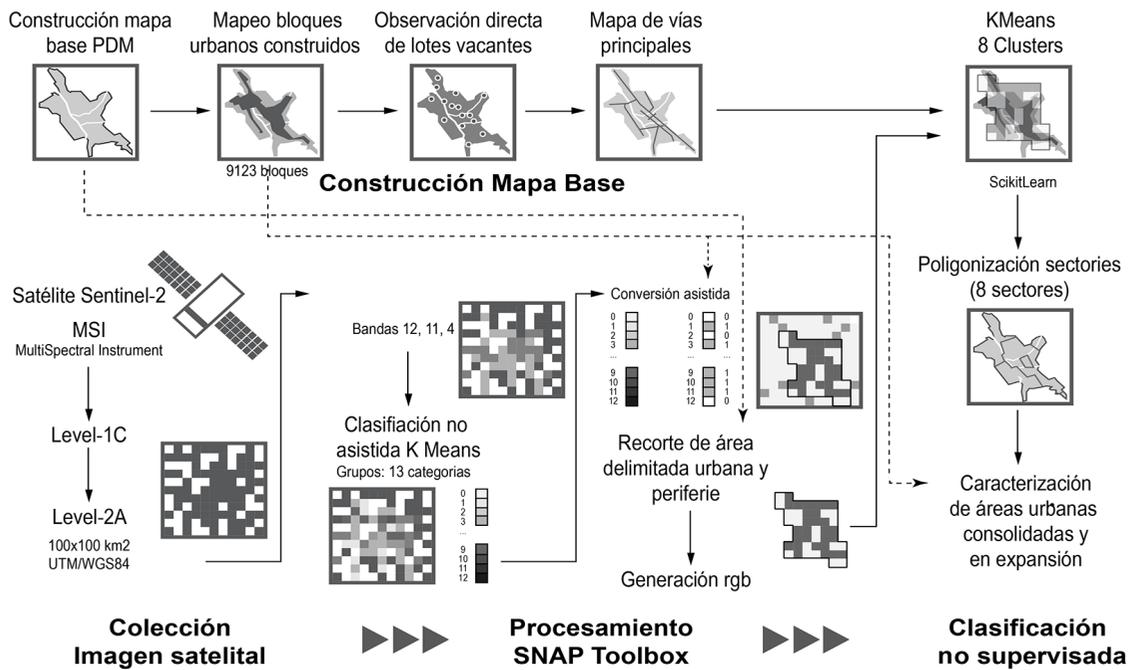


Figura 2. Metodología de recolección y procesamiento de datos. Fuente: Elaboración del autor.

Dimensión	Variable	Valores
Geografía	Pendiente del terreno	Porcentaje de inclinación de pendiente.
	Coordenada X	UTM Este
	Coordenada Y	UTM Norte
Viarío urbano	Distancia a la vía	Distancia en metros (m) a la vía más cercana del sistema de ruta.
Ocupación urbana	Proporción de área ocupada	Área ocupada por edificaciones en un radio de 400m / Área de manzanas proyectadas en un radio de 400m
Límite urbano	Distancia del límite urbano	Distancia en metros (m) al punto más cercano del límite urbano proyectado.
Precio suelo	Precio ofertado del suelo por m ²	Precio del suelo por metro cuadrado (m ²)

Tabla 1. Variables consideradas para la clasificación K-Means. Fuente: Elaboración del autor.

peruanos, disponiendo de una alta capacidad de suelo urbanizable que compite con el suelo rural. El alcance geográfico del estudio aquí expuesto, consideró los distritos Huamancaca y 3 de Diciembre, de la provincia anexa de Chupaca, además de los distritos de Pilcomayo, Chilca, Sapallanga, Huancán, El Tambo y Huancayo de la provincia de Huancayo, dada su ubicación geográfica en el margen derecho del río y la conexión directa que tienen con la ciudad. La ubicación se muestra en la Figura 1.

IV. METODOLOGÍA

La recolección de datos tuvo diversas fuentes: observación directa, imágenes satelitales y mapeos. Estos métodos fueron informatizados y procesados utilizando sistemas de información geográfica QGIS 3.12, SNAP Toolbox v8.0, la librería scikit-learn 0.24 y otras del lenguaje python. La secuencia del proceso puede visualizarse en la Figura 2, desde la construcción del mapa base hasta la generación final de los polígonos de sectores urbanos.

De acuerdo a lo anterior, se utilizó la imagen producida por el instrumento MultiSpectral Instrument (MSI), del satélite Sentinel-2, producto Level-2A que provee una imagen de reflectancia del fondo atmosférico derivada de la asociación del Level-1C, en un área compuesta de 100 x 100 km² bajo una proyección cartográfica UTM/WGS84. Se requirió realizar un redimensionamiento para su adecuada superposición y re-proyección. Así, se procesó las imágenes con el software SNAP v8.0, realizando un redimensionado de la imagen para las bandas 12, 11 y 4, con las que se generó una imagen rgb en falso color. Una vez aisladas las bandas, se procedió a generar una clasificación usando el algoritmo de clasificación no supervisada K-Means. La cantidad de categorías se estableció después de examinar los resultados de entre 3 a 15 categorías, siendo 13

categorías las que expresaban mejor la diversidad de usos de suelo.

El mapeo de los bloques urbanos consolidados fue una tarea semi manual de identificación de polígonos vacantes dentro del límite urbano propuesto, desarrollada sobre las manzanas proyectadas del Plan de Desarrollo Metropolitano e imágenes satelitales de color real. El mapeo de los terrenos periféricos consideró un tamaño del lote mínimo de aproximadamente 100m², similares a los 107m² propuestos por el FMV (2018c). Se identificaron 9123 bloques con un total de 34.22 km², que representan el 33.12% del total del territorio urbano considerado, que fue de 103.32 km². Esta definición de bloques permitió la comparación caracterizada de las áreas que la imagen satelital. Las vías principales fueron identificadas en base a planes existentes presentes en el repositorio del Ministerio de Transporte e indicadas en el Plan de Desarrollo Metropolitano. Asimismo, se usaron las pendientes procesadas a partir de las curvas definidas en la carta nacional, las cuales se expresaron en forma porcentual. Luego, se recolectó la información de 228 lotes válidos de un total de 273 calculados para una muestra simple NC=90%, E=5%, p=50%. Los puntos de observación se distribuyeron aleatoriamente sobre el plano en cantidad proporcional a la densidad poblacional mostrada en el plan vigente. Las características de los puntos de observación que se consideraron se detallan en la Tabla 1 y han sido utilizadas como características para la determinación de la clasificación de las zonas urbanas a través de un algoritmo K-Means, implementado con la librería scikit-learn.

V. RESULTADOS

En la Figura 3, se visualiza la imagen producto Level-2A, capturada y procesada en falso color rgb, usando las bandas 12,11 y 4, respectivamente. Con ello se logra diferenciar en un

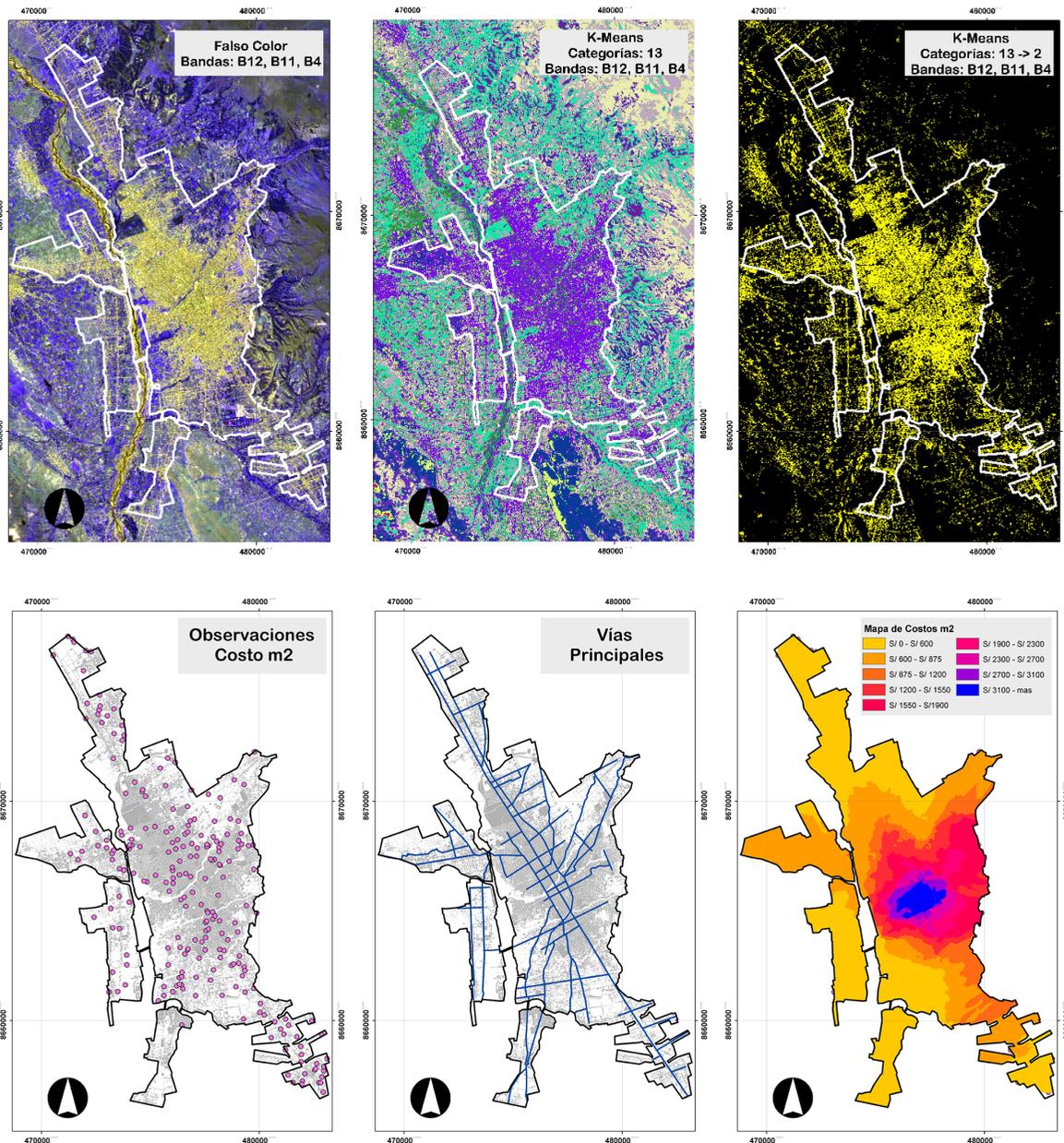


Figura 3. Identificación y clasificación de las imágenes satelitales usando K-Means. Fuente: Elaboración del autor.
 Figura 4. Puntos de muestreo, vías y mapa de valor de suelo. Fuente: Elaboración del autor.

color entre amarillo y violeta las posibles áreas edificadas y otros suelos. Para generar una escala que pueda ser manualmente discriminada, se clusterizó el falso color en 13 categorías usando el algoritmo K-Means. Después, se etiquetó cada cluster como edificado o no edificado, reduciendo los resultados a 2 categorías, las que se distinguen en color amarillo y negro. Debido a que otros tipos de suelo suelen confundirse, se

consideró solo el área limitada de expansión urbana, mejorando la precisión del resultado. A partir de la tercera imagen se calculó el porcentaje de suelo urbano ocupado de los buffers definidos posteriormente.

En la Figura 4, se distingue la distribución de los puntos muestreados observados, distribuidos de forma aleatoria en

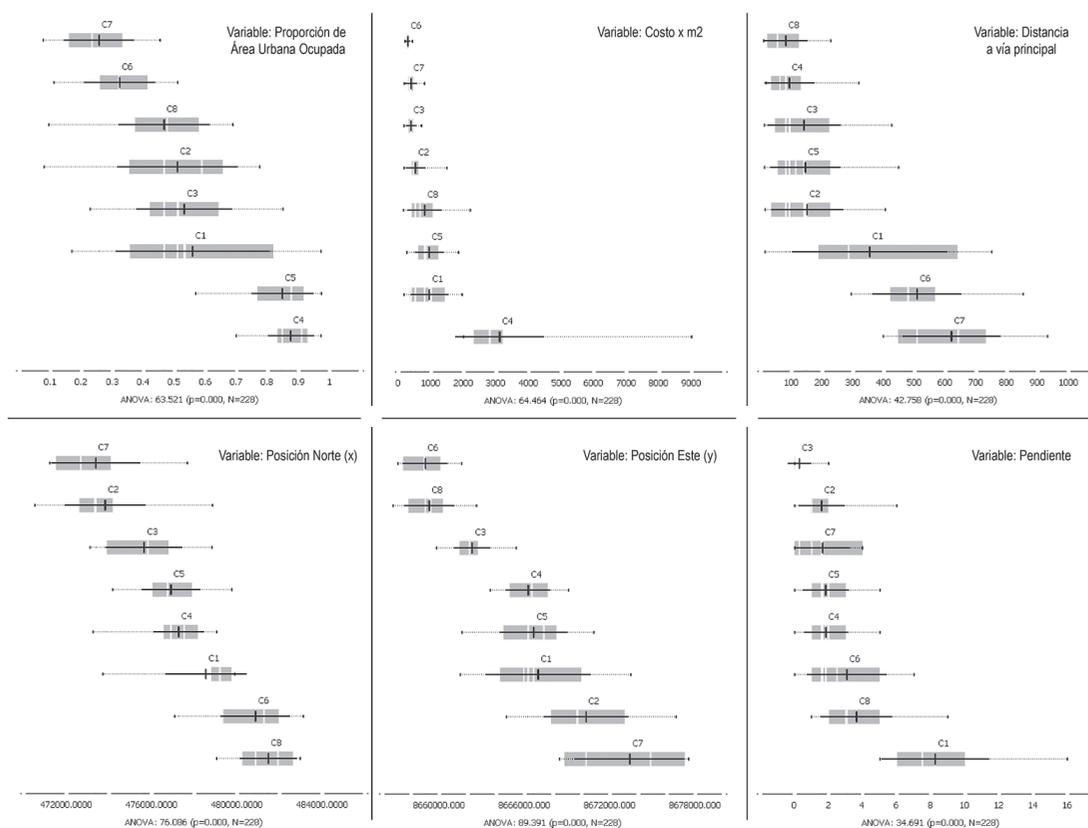
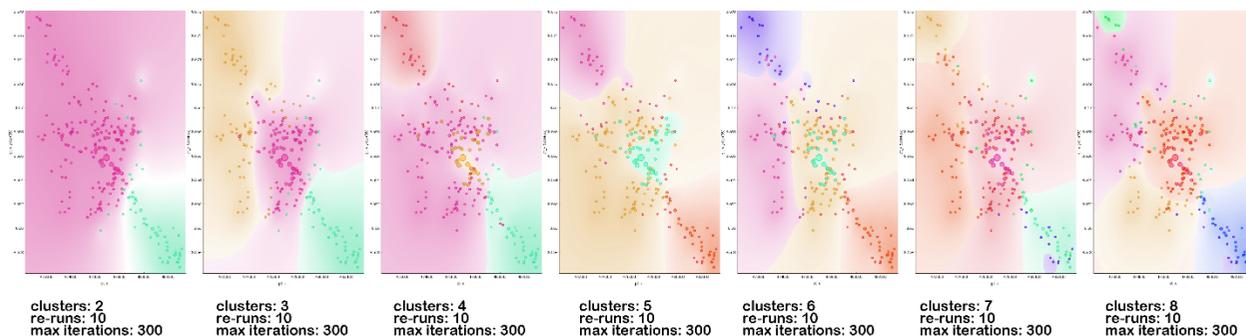


Figura 5. Clasificación usando el algoritmo no supervisado K-Means para la generación de centroides. Fuente: Elaboración del autor.
Figura 6. Vista comparativa entre clústeres por cada variable del modelo. Fuente: Elaboración del autor.

la zona ocupada: la red vial principal indicada en los planos existentes, ambos superpuestos en el mapa de manzanas urbanas consolidadas y en proceso de urbanización. Con los datos de precio x m² ofertados de suelo, se proyectó una imagen DEM generada usando una clasificación de división de optimización de Jenks de 9 categorías. Estas tareas se realizaron con la herramienta de software QGIS 3.12.

Empleando un buffer de 400 m de diámetro, que usó como centro a los puntos de observación, se promedió la distancia a la vía principal más cercana, el costo x m² ofertado, la pendiente, el porcentaje de área urbana ocupada y su referencia de ubicación geográfica. A partir de estos datos, se alimentó al algoritmo K-Means, implementado con la librería de scikit-learn. Se utilizaron parámetros fijos de 10 re-runs y 300 iteraciones para

Variable	Cluster	Estadístico	Cluster						
			C2	C3	C4	C5	C6	C7	C8
Porcentaje de área ocupada	C1	D. media			-0.3166	-0.2898		0.3008	
		P			0.0202	0.0358		0.0355	
	C2	D. media			-0.3648	-0.338	0.1873	0.2526	
		P			<.0001	<.0001	0.0024	0.0002	
	C3	D. media			-0.3432	-0.3164	0.2088	0.2742	
		P			<.0001	<.0001	0.0007	<.0001	
	C4	D. media					0.552	0.6174	0.4076
		P					<.0001	<.0001	<.0001
	C5	D. media					0.5253	0.5906	0.3808
		P					<.0001	<.0001	<.0001
	C6	D. media							-0.1444
		P							0.0457
	C7	D. media							-0.2098
		P							0.0033
Pendiente	C1	D. media	6.6591	7.9643	6.4113	6.4246	5.1786	6.6136	4.5833
		P	0.0003	<.0001	0.0004	0.0005	0.0042	0.0002	0.0084
	C2	D. media		1.3052					-2.0758
		P		<.0001					0.0033
	C3	D. media			-1.553	-1.5397	-2.7857		-3.381
		P			<.0001	<.0001	0.0134		<.0001
	C4	D. media							-1.828
		P							0.0163
	C5	D. media							-1.8413
		P							0.0105
Distancia a vía principal	C2	D. media					-357.389	-469.3501	
		P					<.0001	<.0001	
	C3	D. media					-367.0552	-479.0163	
		P					<.0001	<.0001	
	C4	D. media					-414.7302	-526.6912	
		P					<.0001	<.0001	
	C5	D. media					-363.2758	-475.2369	
		P					<.0001	<.0001	
	C6	D. media							427.3883
		P							<.0001
C7	D. media							539.3494	
	P							<.0001	

Costo m ²	C1	D. media			-2159.2823		659.3929		
		P			<.0001		0.0364		
	C2	D. media			-2587.4413	-426.6631	231.2338		
		P			<.0001	<.0001	0.0003		
	C3	D. media			-2718.6751	-557.8968			-410.8929
		P			<.0001	<.0001			0.0223
	C4	D. media				2160.7783	2818.6751	2725.8504	2307.7823
		P				<.0001	<.0001	<.0001	<.0001
	C5	D. media					657.8968	565.0722	
		P					<.0001	<.0001	
	C6	D. media						-92.8247	-510.8929
		P						0.6962	0.0024
	C7	D. media							-418.0682
		P							0.0303

Tabla 2. Prueba de Games Howell. Fuente: Elaboración autor.

un rango de 2 a 8 clusters. Estos se presentan en la Figura 5, usando los ejes de coordenadas UTM este y norte como x e y, respectivamente.

Para validar las diferencias entre los clústeres resultantes se efectuaron pruebas de ANOVA, todas las cuales resultaron significativas con un valor $p < .0001$. Las diferencias de distribución de los valores se muestran en la Figura 6, a través de gráficos de cajas, con una referencia del estadístico F de cada prueba. Mientras, las diferencias significativas entre grupos, realizadas con la prueba post-hoc de Games-Howell, se ilustran en la Tabla 2.

En la Figura 6, se identifica diferencias entre clústeres (C) por variable. La ubicación geográfica es significativamente distinta para todos los clústeres. La proporción de área urbana ocupada, para C7 y C6 es menor a 0.5 y para C4 y C5 es mayor a 0.7. El costo por m², para el C4 es altamente variable y mayor al de otros clústeres; le siguen el C1, C2 y C8. La distancia a la vía principal más cercana se da un rango menor a 200 m para C8 y C4; menor a 300 m para C3, C5 y C2; entre 400 m a 800 m, para el C6 y C7; y de 100 m a 800 m para C1. Una pendiente mayor a 5% se aprecia en el C1 y menor a 5%, en los otros clústeres.

La Tabla 2 permite identificar diferencias significativas entre clústeres emparejados. La proporción de área urbana ocupada es significativamente diferente entre C1 y C4, C5 y C7; entre C2 y el intervalo que va desde C4 a C7; entre C3 y el intervalo que va desde C4 a C7; entre C4 y C6, C7 y

C8; entre C5 y C6, C7 y C8; entre C6 y C8; y entre C7 y C8. La pendiente, por su parte, es significativamente diferente entre C1 y el intervalo que va desde C2 a C8; entre C2 y C3 y C8; entre C3 y C4, C5, C6 y C8; entre C4 y C8; y entre C5 y C8. La distancia a la vía principal es significativamente diferente entre C2 y C6 y C7; entre C4 y C6 y C7; entre C5 y C6 y C7; entre C6 y C8; y entre C7 y C8. Por último, el costo por m² es significativamente diferente entre C1 y C4 y C6; entre C2 y el intervalo desde C4 a C6; entre C3 y C4, C5 y C8; entre C4 y el intervalo desde C5 a C8; entre C5 y C6 y C7; entre C1 y C7 y C8; y entre C7 y C8. En suma, se identificaron diferencias significativas en todas las variables.

Una vez validadas las diferencias significativas entre clústeres emparejados, se delimitó los polígonos urbanos superpuestos con los centroides y las áreas presentados en la Figura 5. El mapa resultante se expone en la Figura 7 a nivel de manzanas, distinguiendo las consolidadas y las que están en proceso de urbanización.

VI. DISCUSIONES

La investigación utilizó un modelo que prioriza el porcentaje de área ocupada por edificaciones, en contraste con Liu *et al.* (2018) y Steurer y Bayr (2020), quienes utilizan el crecimiento poblacional en base a la densidad cercana. En todos los casos, el algoritmo K-Means hizo

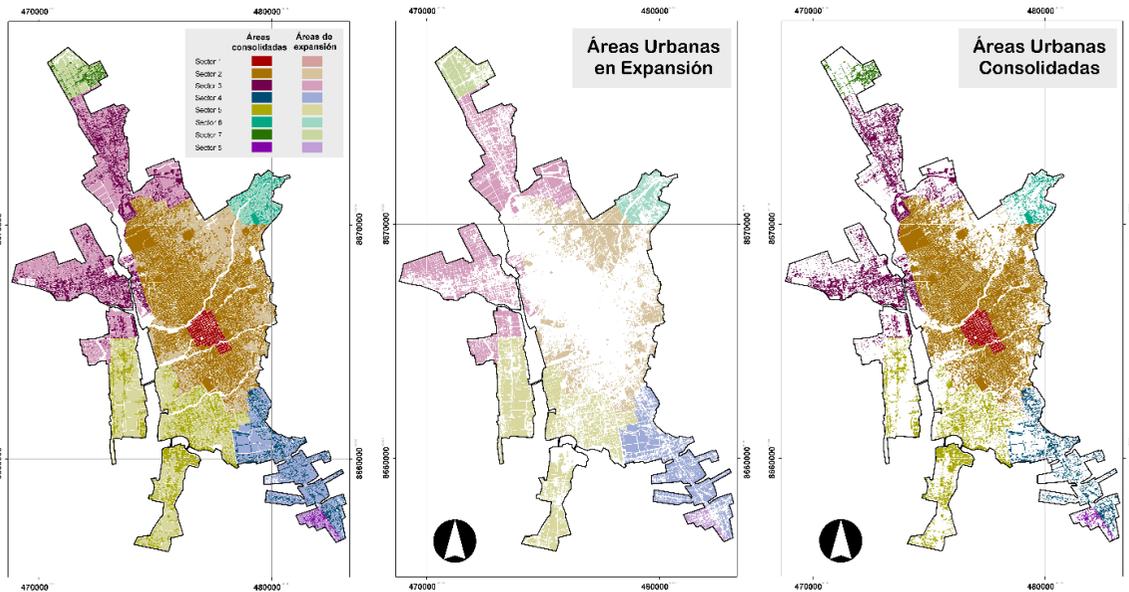


Figura 7. Polígonos de caracterización de zonas urbanas resultantes. Fuente: Elaboración del autor.

posible la clasificación multidimensional. En este sentido, Steurer y Bayr (2020) proponen medios que pueden complementarse con los resultados para investigaciones futuras. En cuanto al ajuste de las fuentes de datos basados en imágenes, se encontró ruido en la clasificación del suelo urbano; sin embargo, su reducción no siguió los parámetros de Zhou *et al.* (2017), sino que se redujeron las capas clasificadas hasta obtener una imagen con datos de 2 valores, que representan el suelo ocupado.

Considerando que la conducta espacial de los fenómenos urbanos es compleja e incierta (Pickard y Meentemeyer, 2019), se requieren cambios para su adecuado estudio. Por tal razón los límites urbanos que habían sido definidos por la MPH (2016) debieron ser ajustados manualmente para poder abarcar las zonas de expansión periurbana y anexar los distritos colindantes en la margen izquierda del río Mantaro. Tras generar la clasificación de clúster, los polígonos de delimitación de la propuesta de MPH (2016) mantienen un contraste de variación que podría contemplarse en futuros planes urbanos. En este aspecto, se debe recordar que la compleja realidad demanda flexibilidad al momento de establecer los límites urbanos y no solo su consideración política y administrativa (Steurer y Bayr, 2020).

El desarrollo de la expansión urbana en Huancayo es disperso y de baja densidad: las diferencias entre clústeres identificadas con la prueba de Games-Howell

revelan que las áreas colindantes al centro, C1 y C2, poseen una proporción mayor distinta a todos los otros 6 clústeres y que esta diferencia se extiende al precio, el cual es en extremo elevado en el C1 y C2, y menor en C3, pero más similar en los restantes. Esta conducta se asemeja al patrón de ocupación en el que se busca un constante precio más bajo y la valorización desproporcionada del mismo en las áreas cercanas a las zonas urbanas consolidadas (Baer y Kauw, 2016; Gasic, 2018). La forma de crecimiento en la ciudad de Huancayo parece contraponerse a la visión ideal de un urbanismo cronológico propuesto por Graells-Garrido *et al.* (2021) y Moreno *et al.* (2021). La creciente distancia de los centralizados servicios urbanos y un deficiente transporte deben ser prioritarios (Vorontsova *et al.*, 2016).

Tal como Araque Solano y Caballero Quintero (2009) señalan, los precios en mercados informales ingresan a su formalización en condiciones similares que a sectores consolidados. En la ciudad de Huancayo, se distingue esta variación en los colindantes de los C2 y C3 que cierran un proceso de consolidación a un precio mayor. El fraccionamiento identificado podría estar vinculado a la escasa participación del sector público sobre el control del mercado (López Navarrete y Peña Medina, 2017). Este veloz incremento del precio periférico que persigue un incremento del capital es una situación común en otros escenarios como los analizados por Amézquita, Rodríguez y Murillo (2015), Gaete (2021) o Gasic (2018). A

lo dicho hay que añadir los efectos reglamentarios e intervenciones económicas particulares o individuales como política de Estado (Li *et al.*, 2019; Garza Puentes y Tovar Vanegas, 2009).

VII. CONCLUSIONES

La presente investigación identifica que el algoritmo K-Means provee una forma viable de clasificación del suelo urbano usando variables heterogéneas y que la diferencia entre los clústeres generados puede probarse como multivariada y diferenciada mediante datos abiertos. A partir de esta clasificación, en la ciudad de Huancayo se identifica un fraccionamiento espacial que, principalmente, se determina por las variables proporción de suelo ocupado, precio ofertado y distancia al sistema vial principal.

Aunque la caracterización no asistida puede someterse a opiniones, se debe poner en alerta la situación que viene experimentando la expansión urbana en Huancayo y plantear alternativas para un análisis más objetivo de su ocupación, aprovechando los medios de análisis disponibles.

VIII. REFERENCIAS BIBLIOGRÁFICAS

Aceid y Fundación ACS (2018). *Encuentro Iberoamericano sobre Prácticas Urbanas*. Recuperado de <https://www.aecid.ht/es/convocatorias/encuentro-iberoamericano-sobre-practicas-urbanas-innovadoras-hacia-ciudad-2030>

Alfasi, N. y Migdalovich, E. (2020). Losing faith in planning. *Land Use Policy*, 97. DOI: <https://doi.org/10.1016/j.landusepol.2020.104790>

Amézquita, L., Rodríguez, L. y Murillo, H. (2015). Los precios del suelo en Bogotá. El barrio Veinte de Julio. *Bitácora Urbano Territorial*, 25(1), 43-50. DOI: <https://dx.doi.org/10.15446/bitacora.v1n25.40236>

Araque Solano, A. y Caballero Quintero, Y. (2009). La encrucijada de la Vivienda de Interés Social en Bogotá: Los precios del suelo. *Civilizar Ciencias Sociales y Humanas*, 9(16), 127-152. Recuperado de http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-89532009000100009&lng=en&tying=es

Baer, L. y Kauw, M. (2016). Mercado inmobiliario y acceso a la vivienda formal en la Ciudad de Buenos Aires, y su contexto metropolitano, entre 2003 y 2013. *Revista EURE - Revista de Estudios Urbano Regionales*, 42(126). Recuperado de <https://www.eure.cl/index.php/eure/article/view/1676>

Belhadia, A., Djenourib, Y., Norvagc, K., Ramampiaroc, H., Masseglied, F. y Lin, J. (2020). Space-time series clustering: Algorithms, taxonomy and case study on urbansmart cities. *Engineering Applications of Artificial Intelligence*, 95. DOI: <https://doi.org/10.1016/j.engappai.2020.103857>

Calderón, J. (2015). Programas de vivienda social nueva y mercados de suelo urbano en el Perú. *Revista EURE - Revista de Estudios Urbano Regionales*, 41(122). Recuperado de <https://www.eure.cl/index.php/eure/article/view/654>

Campesato, O. (2020). *Artificial Intelligence Machine Learning and Deep Learning*. New Delhi: Mercury Learning and Information.

Cardó, P. (2017). Estancamiento y crecimiento del área urbana en 2012. Oferta del suelo en el Gran San Juan. *Bitácora Urbano Territorial*, 27(1), 27-34. DOI: dx.doi.org/10.15446/bitacora.v27n1.40348

Carvajal Mascaró, F., Moreira Muñoz, A., Salazar Burrows, A., Leguía Cruz, M. y Jorquera-Guajardo, F. (2019). Divergencias y contradicciones en la planificación sustentable del periurbano rural metropolitano de Valparaíso. Caso Reserva de la Biosfera La Campana-Peñuelas, Chile central. *Urbano*, 22(39), 64-87. DOI: <https://doi.org/10.22320/07183607.2019.22.39.04>

Castillo-García R. (2021). Evolución de la Planificación Urbana en el Perú 1946 - 2021: De la Planificación Urbana Normativa a la Planificación del Desarrollo Urbano Sostenible. *PAIDEIA XXI*, 11(1), 31-79. DOI: <https://doi.org/10.31381/paideia%20xxi.v11i1.3783>

Delgadillo, V. (2016). Ciudades iletradas: orden urbano y asentamientos populares irregulares en la ciudad de México. *Territorios*, 35, 81-99. DOI: <https://dx.doi.org/10.12804/territ35.2016.04>

Durduran, S. (2015). Automatic classification of high-resolution land cover using a new data weighting procedure: The combination of K-Means clustering algorithm and central tendency measures (KMC-CTM). *Applied Soft Computing Journal*, 35, 136-150. DOI: <https://doi.org/10.1016/j.asoc.2015.06.025>

Espinoza, A. y Fort, R. (2017). *Mapeo y Tipología de la Expansión Urbana en el Perú*. Lima: ADI Perú.

Feng, Z., Peng, J. y Wu, J. (2020). Using DMSP/OLS nighttime light data and K-means method to identify urban-rural fringe of megacities. *Habitat International*, 103. DOI: <https://doi.org/10.1016/j.habitatint.2020.102227>

Foldvary, F. y Minola, L. (2017). The taxation of land value as the means towards optimal urban development and the extirpation of excessive economic inequality. *Land Use Policy*, 69, 331-337. DOI: <http://dx.doi.org/10.1016/j.landusepol.2017.09.022>

Fondo Mivivienda [FMV] (2018a). Miles se benefician con subsidios. *La revista inmobiliaria del Perú: MIVIVIENDA*, 126(14), 4-5. Lima: Fondo MIVIVIENDA.

Fondo Mivivienda [FMV] (2018b). Mejorarán planes de desarrollo urbano. *La revista inmobiliaria del Perú: MIVIVIENDA*, 126(14), 6-7. Lima: Fondo MIVIVIENDA.

Fondo Mivivienda [FMV] (2018c). Tienen capacidad de compra. *La revista inmobiliaria del Perú: MIVIVIENDA*, 127(14), 6-9. Lima: Fondo MIVIVIENDA.

Forestier, G. y Wemmer, C. (2016). Semi-supervised learning using multiple clusterings with limited labeled data. *Information Sciences*, (361-362), 48-65. DOI: <https://doi.org/10.1016/j.ins.2016.04.040>

Gaete, H. (2021). Tendencias del mercado de suelo urbano en periodo largo. Concepción, Chile. 1989-2018. *ACE: architecture, city and environment*, 16(46), 9946. DOI: <http://dx.doi.org/10.5821/ace.16.46.9946>

García, F. y Peralta, M. (2016). Las urbanizaciones multifamiliares cerradas y su entorno urbano: una nueva geografía simbólica en la ciudad de Cali (Colombia). *Revista EURE - Revista de Estudios Urbano Regionales*, 42(126). Recuperado de <https://www.eure.cl/index.php/eure/article/view/1548/880>

Garza Puentes, N. y Tovar Vanegas, R. (2009). El mercado de vivienda en Barranquilla y el sector externo de la economía. *Revista de Economía del Caribe*, (4), 181-209.

Gasic, I. (2018). Inversiones e intermediaciones financieras en el mercado del suelo urbano. Principales hallazgos a partir del estudio de transacciones de terrenos en Santiago de Chile, 2010-2015. *Revista EURE - Revista de Estudios Urbano Regionales*, 44(133). Recuperado de <https://www.eure.cl/index.php/eure/article/view/2403/1108>

Gasparyniene L., Venclauskiene, D. y Remeikiene, R. (2014). Critical review of selected housing market models concerning the factors that make influence on housing price level formation in the countries with

transition economy. *Procedia - Social and Behavioral Sciences* 110, 419 – 427. DOI: <https://doi.org/10.1016/j.sbspro.2013.12.886>

Gedal, M. y Ellen, I. (2018). Valuing urban land: Comparing the use of teardown and vacant land sales. *Regional Science and Urban Economics*, 70, 190–203. DOI: <https://doi.org/10.1016/j.regsciurbeco.2018.03.006>

Glaeser, E., Gyourko, J. (2003). The impact of building restrictions on housing affordability. *Federal Reserve Bank of New York Economic Policy Review* 9(2), 21–39.

Glaeser, L. y Ward, A. (2009). The causes and consequences of land use regulation: Evidence from Greater Boston. *Journal of Urban Economics*, 65(3), 265–278. DOI: <https://doi.org/10.1016/j.jue.2008.06.003>

GRADE (2020). *Hacia una nueva política de vivienda en el Perú: Problemas y posibilidades*. Lima: Grupo de Análisis para el Desarrollo GRADE.

Graells-Garrido, E., Serra Burriel, F., Rowe, F., Cucchiatti, F. y Reyes, P. (2021). A city of cities: Measuring how 15-minutes urban accessibility shapes human mobility in Barcelona. *PLOS ONE* 16(5). DOI: <https://doi.org/10.1371/journal.pone.0250080>

Hindi, T. de M. C., Moreira, T. y Rossi, A. (2020). Fornas alternativas de propriedade: cooperativismo, produção capitalista do espaço urbano e mercado imobiliário. A: Seminario Internacional de Investigación en Urbanismo. XII Seminario Internacional de Investigación en Urbanismo, São Paulo-Lisboa, 2020. São Paulo: Faculdade de Arquitetura da Universidade de Lisboa. DOI: <https://doi.org/10.5821/siiu.10040>.

Hwang, S., Park, M. y Lee, H. (2013). Dynamic analysis of the effects of mortgage-lending policies in a real estate market. *Mathematical and Computer Modelling*, 57(9–10), 2106–2120. DOI: <https://doi.org/10.1016/j.mcm.2011.06.023>

Klaufus, C., Van Lindert, P., Van Noorloos, F. y Steel, G. (2017). All-Inclusiveness versus Exclusion: Urban Project Development in Latin America and Africa. *Sustainability*, 9(11). DOI: <https://doi.org/10.3390/su9112038>

Kok, N., Monkkonen, P. y Quigley, J. (2014). Land use regulations and the value of land and housing: An intra-metropolitan analysis. *Journal of Urban Economics*, 81, 136–148. DOI: <https://doi.org/10.1016/j.jue.2014.03.004>

Joshi, A. (2020). *Machine Learning and Artificial Intelligence*. Springer.

Li, X., Sun, M. y Boersma, K. (2019). Policy Spillover and Regional Linkage Characteristics of the Real Estate Market in China's Urban Agglomerations. *Journal of Management Science and Engineering*, 4(3), 189–210. DOI: <https://doi.org/10.1016/j.jmse.2019.05.004>

Liu, J., Jiao, L., Zhang, B., Xu, G., Yang, L., Dong, T., Xu, Z., Zhong, J. y Zhou, Z. (2021). New indices to capture the evolution characteristics of urban expansion structure and form. *Ecological Indicators*, 122(1). DOI: <https://doi.org/10.1016/j.ecolind.2020.107302>

Liu, L., Peng, Z., Wu, H., Jiao, H., Yu, Y. y Zhao, J. (2018). Fast identification of urban sprawl based on K-Means clustering with population density and local spatial entropy. *Sustainability (Switzerland)*, 10(8). DOI: <https://doi.org/10.3390/su10082683>

López Navarrete, J. y Peña Medina, S. (2017). La segregación socioespacial en Ciudad Juárez, Chihuahua, 1990–2010. *Región y sociedad*, 29(68), 115–152. DOI: <https://dx.doi.org/10.22198/rys.2017.68.a210>

Migueltorena, A. y Lan, D. (2013). Racionalidades y contrarrazones, a partir de la vivienda, en la producción del espacio urbano de Tandil, Argentina. Cuadernos de Geografía: *Revista Colombiana de Geografía*, 22(1), 109–125. Recuperado de <https://www.redalyc.org/articulo.oa?id=281825518011>

Ministerio de Vivienda Construcción y Saneamiento [MVCS] (2021). *Resumen de la Política Nacional de Vivienda y Urbanismo*. Lima: MVCS. Recuperado de <https://www.gob.pe/institucion/vivienda/informes-publicaciones/2027198-resumen-de-la-politica-nacional-de-vivienda-y-urbanismo>

Moreno, C., Allam, Z., Chabaud, D., Gall, C. y Pratlong, F. (2021). Introducing the “15-Minute City”: Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. *Smart Cities*, 4(1), 93–111. DOI: <http://dx.doi.org/10.3390/smartcities4010006>

Mouratidis, K. (2021). Urban planning and quality of life: A review of pathways linking the built environment to subjective well-being. *Cities*, 115, 103229. <https://doi.org/10.1016/j.cities.2021.103229>

Municipalidad Provincial de Huancayo (2016). *Plan de Desarrollo Metropolitano – Provincia de Huancayo 2015 – 2035*. Huancayo: Municipalidad Provincial de Huancayo.

Nabil, N. y Eldayem, G. (2015). Influence of mixed land-use on realizing the social capital. *HBRC Journal*, 11(2), 285–298. DOI: <https://doi.org/10.1016/j.hbrj.2014.03.009>

Naciones Unidas (2018). *Plan de Acción Regional para la implementación de la Nueva Agenda Urbana en América Latina y el Caribe 2016–2036*. Santiago: Naciones Unidas.

Parias, A. (2008). El mercado de arrendamiento en los barrios informales en Bogotá, un mercado estructural. *Territorios*, (18–19), 75–101. Recuperado de <https://revistas.urosario.edu.co/index.php/territorios/article/view/828>

Peña-Zamalloa, G. (2018). Proyección del cambio de uso de suelo urbano mediante técnicas de microsimulación, bajo un escenario de escasez de datos en el sector de San Carlos, Huancayo, Perú 2018–2028. *Espacio y Desarrollo*, 32, 99–124. DOI: <https://doi.org/10.18800/espacioydesarrollo.201802.005>

Pickard, B. y Meentemeyer, R. (2019). Validating land change models based on configuration disagreement. *Computers, Environment and Urban Systems*, 77. DOI: <https://doi.org/10.1016/j.compenvurbsys.2019.101366>

Pimentel Sánchez, N. (2020). ¿Tomar lotes para vivir o para vender? Tráfico de tierras y práctica clientelar en la periferia urbana. *Revista de Sociología*, (31), 133–159. DOI: <https://doi.org/10.15381/rsoc.v0i31.19279>

Sabatini, F. y Arenas, F. (2000). Entre el Estado y el mercado: resonancias geográficas y sustentabilidad social en Santiago de Chile. *Revista EURE - Revista de Estudios Urbano Regionales*, 26(79). Recuperado de <https://www.eure.cl/index.php/eure/article/view/1212>

Salazar, C. (2014). “El puño invisible” de la privatización. *Territorios*, 30, 69–90. DOI: <https://dx.doi.org/10.12804/territ30.2014.03>

Saleh, A., Hwa, T. y Majid, R. (2016). Housing Mismatch Model in Suburban Areas. *Procedia - Social and Behavioral Sciences*, 234, 442–451. DOI: <https://doi.org/10.1016/j.sbspro.2016.10.262>

Scotiabank (2015). *El mercado inmobiliario en perspectiva*. Lima: Estudios Económicos Scotiabank.

Steurer, M. y Bayr, C. (2020). Measuring urban sprawl using land use data. *Land Use Policy*, 97. DOI: <https://doi.org/10.1016/j.landusepol.2020.104799>

Tellier, L. (2020). Characterizing urban form by means of the Urban Metric System. *Land Use Policy*. DOI: <https://doi.org/10.1016/j.landusepol.2020.104672>

Tricaricoa, L. y de Vidovich, L. (2021). Proximity and post-COVID-19 urban development: Reflections from Milan, Italy. *Journal of Urban Management*, 10(3), 302-310. DOI: <https://doi.org/10.1016/j.jum.2021.03.005>

Vorontsova, A., Vorontsova, V. y Salimgareev, D. (2016). The development of Urban Areas and Spaces with the Mixed Functional Use. *Procedia Engineering*, 150. DOI: <https://dx.doi.org/10.1016/j.proeng.2016.07.277>

Yu, J., Zhou, K. y Yang, S. (2019). Land use efficiency and influencing factors of urban agglomerations in China. *Land Use Policy*, 88. DOI: <https://doi.org/10.1016/j.landusepol.2019.104143>

Zhou, X., Gu, J., Shen, S., Ma, H., Miao, F., Zhang, H. y Gong, H. (2017). An automatic K-Means clustering algorithm of GPS data combining a novel niche genetic algorithm with noise and density. *ISPRS International Journal of Geo-Information*, 6(12), 392. DOI: <https://doi.org/10.3390/ijgi6120392>