

CIUDAD Y TERRITORIO

ESTUDIOS TERRITORIALES

ISSN(P): 1133-4762; ISSN(E): 2659-3254

Vol. LVI, Nº 221, otoño 2024

Págs. 877-896

<https://doi.org/10.37230/CyTET.2024.221.8>

CC BY-NC-ND



Impacto de la morfología urbana en la segmentación del mercado de la vivienda: caso de estudio Madrid

David REY-BLANCO ⁽¹⁾Juan Ramón SELVA-ROYO ⁽²⁾Julio GONZÁLEZ-ARIAS ⁽³⁾⁽¹⁾ Chief Data Officer, Idealista⁽²⁾ Spatial Data Scientist, Idealista⁽³⁾ Profesor de Economía Financiera. Universidad Nacional de Educación a Distancia (UNED)

Resumen: Los submercados de viviendas representan segmentos inmobiliarios con características similares, a menudo confundidos con las áreas administrativas existentes, como barrios o distritos, que no reflejan la estructura inmobiliaria. Esto deriva en medidas de precios medios no homogéneas, difíciles de interpretar y poco representativas. Se propone la combinación de un seccionado geográfico discreto y jerárquico sobre segmentos tipomorfológicos con la división administrativa. Se definen 8 tipos urbanos en la ciudad de Madrid, aplicables a otros contextos urbanos, calculados a partir de datos catastrales. El análisis experimental confirma la reducción de la incertidumbre de los precios medios en 124 de los 136 barrios de Madrid, anticipando su potencial aplicación en mejora de índices de precios o en la identificación de áreas “santuario” desde un punto de vista inmobiliario.

Palabras clave: Índice de precios de la vivienda; Submercados inmobiliarios; Morfología urbana; Vivienda; Tejidos urbanos.

Recibido: 31.03.2023; Revisado: 09.02.2024

Correo electrónico: drey@idealista.com N. ORCID: <https://orcid.org/0000-0002-3549-1714>

Correo electrónico: jrselva@idealista.com N. ORCID: <https://orcid.org/0000-0001-6617-5741>

Correo electrónico: jglez@cee.uned.es N. ORCID: <https://orcid.org/0000-0003-4993-4739>

Los autores agradecen los comentarios y sugerencias realizados por las personas evaluadoras anónimas, que han contribuido a mejorar y enriquecer el manuscrito original.

También agradecen a Idealista y a la Administración, por la puesta a disposición de los datos necesarios para la investigación.

Impact of urban morphology on real estate segmentation: Madrid case study

Abstract: Housing submarkets represent real estate segments with similar characteristics, often confused with existing administrative areas such as neighborhoods or districts, which do not reflect the real estate structure. This results in non-homogeneous average price measures that are unrepresentative and difficult to interpret. Our proposal combines a discrete and hierarchical geographical segmentation of typomorphological segments along with administrative division. Eight urban types are defined in the city of Madrid, applicable to other urban contexts, and calculated from cadastral data. Experimental analysis confirms the reduction of uncertainty in average prices in 124 out of 136 neighborhoods in Madrid, anticipating its potential application in improving price indices or identifying “sanctuary” areas from a real estate standpoint.

Keywords: House Price Index; Submarket segmentation; Housing; Typomorphology; Urban tissues.

1. Introducción

En los últimos años, la información espacial y los datos masivos se han convertido en un lugar común para la investigación y el conocimiento profundo de la morfología urbana (ARRIBAS-BEL, 2014). Si bien es cierto que los investigadores concentraron inicialmente sus esfuerzos en el estudio del tejido urbano, como instrumento clave para comprender el contexto físico de la ciudad (CANIGGIA & MAFFEI, 1979; HAYWARD, 1987), la irrupción de las nuevas tecnologías y plataformas de datos abiertos ha evidenciado recientemente el potencial del análisis de los patrones espaciales para mejorar la comprensión de la complejidad de la ciudad contemporánea (BOEING, 2019; CARPIO-PINEDO & al., 2020; FLEISCHMANN & al., 2021).

La investigación académica ha adquirido recientemente gran peso en la caracterización de las ciudades, con contribuciones muy significativas, generalmente basadas en un análisis de parámetros de densidad, también conocidas como el enfoque Spacemate (BERGHAUSER-PONT & HAUPT, 2007, 2009), y en procesos de identificación remota de tejido urbano. En cuanto a este último, se han logrado nuevos conocimientos tanto a través de un enfoque vinculado a patrones morfométricos (DIBBLE & al., 2019), como a estudios de sostenibilidad energética (RODE & al., 2014; SUBDIRECCIÓN GENERAL DE POLÍTICAS URBANAS, VIVIENDA Y SUELO, 2019; MARTÍN-CONSUEGRA & al., 2021). Estos avances han confirmado, entre otras cuestiones, que la identificación global del tejido urbano es una tarea difícil de separar de la realidad cultural e histórica (PANERAI & al., 2012).

Disciplinas como la tipomorfología urbana y el estudio taxonómico de la forma urbana según patrones físicos y culturales, han demostrado

ser especialmente útiles para complementar enfoques clásicos como la metodología de Space Syntax (LAW & al., 2013; STONOR, 2014; BERGHAUSER-PONT & MARCUS, 2015), en términos de redes de calles y conectividad urbana, o bien a través de métodos de aprendizaje automático para identificar regiones homogéneas como las Spatial Signatures, propuestas por FLEISCHMANN & ARIBAS-BEL (2022), y que permiten una interpretación exhaustiva del tejido urbano (ARRIBAS-BEL & FLEISCHMANN, 2022). No hay duda de que abren grandes posibilidades a la investigación urbana, cuando los conjuntos de datos masivos espaciales se someten a un análisis adecuado (BOEING, 2018; SAMARDZHIEV & al., 2022).

Por otro lado, disponer de herramientas adecuadas para controlar la evolución de los precios de la vivienda es esencial para desarrollar una política económica eficaz, dado el nivel de exposición de las economías avanzadas al sector inmobiliario. Particularmente para el mercado inmobiliario español, que en los últimos años ha mostrado una gran fortaleza, con crecimientos hasta el primer semestre de 2022, pero expuesto a potenciales riesgos derivados de una alta inflación, alza de tipos de interés y gran incertidumbre geopolítica (RODRÍGUEZ, 2022b, 2022a). El análisis de los precios de la vivienda es una herramienta útil para el estudio de otros fenómenos, por ejemplo, ECHAVES & MARTÍNEZ (2021) determinan que la proporción de precios de compraventa y alquiler está relacionada con la tasa de vivienda en alquiler y, también, con el grado de emancipación de los jóvenes.

Las herramientas de análisis de los fenómenos inmobiliarios han sido tradicionalmente los índices de precios de la vivienda. Para el caso español, el INSTITUTO NACIONAL DE ESTADÍSTICA, INE (2022) publica con carácter trimestral,

complementarios a los datos estadísticos publicados por el Ministerio de Transportes, Movilidad y Agenda Urbana (MITMA) que contienen, entre otros, precios de escrituración o costes de construcción. También es posible consultar estadísticas de precios de oferta de los portales inmobiliarios Idealista (IDEALISTA 2023) y FOTOCASA (2021). Sin embargo, el precio de la vivienda es muy sensible al precio del suelo, por tanto, todos estos instrumentos de análisis deben tener muy en consideración el aspecto espacial (HILL & SCHOLZ, 2018), debido a los fenómenos de heterogeneidad¹ y dependencia espacial² en el ámbito de la vivienda (ANSELIN & REY, 2014).

En ninguno de las estadísticas o índices de precios anteriores se utiliza la segmentación tipomorfológica para definir los estratos muestrales. Por ejemplo, en el caso del Índice de Precios de la Vivienda (IPV) del INE (2022), que es la estadística más importante de la vivienda residencial en España, solo se tienen en consideración las características económicas del entorno, las características funcionales de la vivienda, como la superficie o si tiene aparcamiento, y la información de ámbito geográfico de alto nivel³, pero no el tejido urbano en el que se enmarca.

Las medidas de precios medios están muy expuestas a dos cuestiones: la primera, que cada vivienda es un activo singular y por tanto no es comparable completamente con el resto; y la segunda, la presencia de aspectos desconocidos en la operación⁴. La falta de control de este factor, al calcular los índices de precios, puede incorporar una fuente de error en las estadísticas de precios. Esta puede describirse como el rango de precios distintos que una misma vivienda podría tomar en el momento de venta, y es un fenómeno reconocido en varios campos empíricos de la economía (KIRMAN & VRIEND, 2001).

La idea planteada en este artículo parte del supuesto de que la segmentación del mercado inmobiliario en submercados puede ser más eficiente si se consideran las características tipológicas de las viviendas, en lugar de simplemente dividir por zonas geográficas. Para contrastar esta hipótesis, se utiliza el ejemplo de la ciudad

de Madrid, que cuenta con 21 distritos y 136 barrios según el criterio de IDEALISTA⁵, para proponer un nuevo método de segmentación espacial basado en un índice espacial de tipo hexagonal. El objetivo es identificar grupos de viviendas con características tipológicas similares, para una definición más precisa de los estratos muestrales de bienes inmobiliarios sobre los que calcular estadísticas o índices de precios. Además, se ha elaborado y divulgado un conjunto de datos abiertos exhaustivo y documentado⁶.

2. Morfología urbana y submercados de la vivienda

Tal y como hemos visto, la morfología urbana ha recibido un fuerte impulso en esta última década, motivada por el desarrollo de nuevas capacidades de identificación de parámetros relevantes y la implementación de agrupación se zonas mediante métodos sin intervención manual (BOBKOVA & al., 2017; BOBKOVA & BERGHAUSSER-PONT & MARCUS, 2019).

La aparición de estos nuevos y masivos métodos cuantitativos, se puede aplicar no solo a la configuración morfológica de la ciudad, sino también a otras disciplinas vinculadas a los estudios urbanos. Como muchos autores han destacado (HILLIER & HANSON, 1988; KOSTOF, 1991; BATTY, 2013), una adecuada identificación de la variedad de procesos urbanos, incluidos los submercados de la vivienda, podría lograrse a través de este compromiso espacial y transdisciplinario. Por esta razón, parece pertinente considerar si la tipología también puede explicar y ser decisiva en las propuestas de segmentación de submercados de vivienda.

El creciente interés en la definición de patrones de configuración espacial que identifiquen fácilmente áreas homogéneas en la ciudad (FLEISCHMANN & ARRIBAS-BEL, 2022), tiene precedentes inmediatos y exitosos relacionados con técnicas de análisis de agrupación (clusterización). Estos métodos establecen que puede haber un conjunto de variables explicativas en calles,

¹ La heterogeneidad espacial (ANSELIN & GRIFFITH, 1988) consiste en la falta de uniformidad de los efectos (incluida la dependencia espacial) sobre el espacio geográfico.

² En términos de asociación (correlación), la dependencia espacial implica precios parecidos para lugares cercanos, en términos más amplios se refiere a la influencia del entorno en las variaciones del precio.

³ La zona define según el tipo de entorno sociodemográfico, población municipal o grupo de provincias.

⁴ Formalmente definidas como variables omitidas, consecuencia de aspectos específicos de la operación que no

están controlados por el modelo aplicado, particulares de los agentes que la llevan a cabo, como, por ejemplo, el nivel de descuento máximo que está dispuesto a aplicar el vendedor.

⁵ Oficialmente el ayuntamiento de Madrid divide la ciudad en 131 barrios <https://datos.gob.es/es/catalogo/101280796-barrios-municipales-de-madrid>.

⁶ Los datos (REY & SELVA-ROYO, 2023) se acompañan del material de análisis, de acuerdo con las directrices sugeridas por PÁEZ (2021) para trabajos de investigación abiertos y reproducibles (ARRIBAS-BEL & al., 2021).

parcelas y edificios que, agrupados adecuadamente, permiten describir los diferentes tejidos urbanos.

Por lo tanto, la investigación se centra en aquellos principios en los que se basa la zonificación urbana, en consonancia con las características morfológicas de los edificios. Aunque en España cada administración local tiene la competencia urbanística para definir los criterios con los que define estas zonas, a menudo con un nombre muy variado, suele ser común distinguir entre grupos fácilmente reconocibles: viviendas colectivas frente a unifamiliares, bloques compactos frente a viviendas en hilera, etcétera. Junto con esta situación no estandarizada, también vale la pena señalar que estas categorías tienden a agruparse finalmente dentro de los procesos de desarrollo urbano.

Derivado de lo anterior, generalmente no se tienen suficientemente en cuenta los criterios tipomorfológicos en los límites espaciales administrativos y estadísticos. Estos dos desafíos —una identificación adecuada y la correspondiente agrupación espacial— han dificultado históricamente la vinculación de los estudios urbanos cuantitativos con las categorías tipológicas. Dando un paso más, en este trabajo nuestro objetivo es identificar de forma remota los tejidos urbanos, tratando de automatizar la captura del carácter compuesto y multi-escala de este concepto morfológico más allá de las actuales fronteras administrativas.

Las áreas urbanas desde la perspectiva del mercado inmobiliario, en general, no suelen mostrarse como una entidad uniforme, sino que pueden considerarse como la agregación de segmentos distintivos en el espacio llamados submercados (GOODMAN & THIBODEAU, 1998), de modo que la segmentación de la demanda y la oferta en la geografía podría llevar a un desequilibrio espacial. Aunque hay un acuerdo general entre los investigadores acerca de la existencia de subconjuntos, no existe una definición única ni coherente de lo que se debería considerar como un submercado de vivienda (ADAIR & al., 1996; WATKINS, 2001), aunque se han probado distintas aproximaciones⁷ aún no se ha logrado un canónico para definirlos (BOURASSA & al., 2021).

La idea económica del equilibrio del mercado subyace en el modelo de precios hedónicos con competencia perfecta, que fue introducido inicialmente por ROSEN (1974). De ahí se deriva

que el precio marginal implícito estimado para las características de la propiedad es igual a la disposición de los individuos marginales a pagar por estos atributos de la vivienda. Gran parte de estos atributos están relacionados con la morfología de la casa, otros se deben al estado de conservación o a especificidades de la ubicación, como comodidades cercanas.

La literatura especializada maneja dos categorías de especificaciones para definir los submercados, ya sea a través de métodos espaciales o no espaciales. Estas enfatizan en un área geográfica predefinida con las preferencias de elección homogéneas de las personas como el principal factor impulsor (por ejemplo, ciudad interior / exterior, norte / sur, distritos políticos y distritos postales) (MACLENNAN, 1987; GABRIEL & WOLCH, 1984; HANCOCK & al., 1991). La definición espacial se basa en que la ubicación es un elemento portador de utilidad (GRILICHES, 1961), por lo tanto, los elementos localizados cercanos a una vivienda generarán una serie de beneficios (utilidad) o perjuicios (desutilidad) para el propietario de dicha vivienda.

Cuando el tipo de construcción es diverso y su segregación espacial es diferente, el submercado de viviendas también puede especificarse por los tipos de estructura de construcción (por ejemplo, casa independiente, adosada, en hilera), o alternativamente por características de la estructura (por ejemplo, la edad de la propiedad o el número de habitaciones) (HEIKKILA & al., 1989; D'ACCI, 2019). Además, algunos investigadores han encontrado una especificación con espacio y construcción anidados, en la que las características de la estructura tienen un mejor rendimiento (GOODMAN, 1981; MACLENNAN & TU, 1996; BOURASSA & al., 2003). Sin embargo, no hay un consenso sobre la cuestión de cómo definir la división espacial, lo que lleva a elegir áreas existentes en detrimento de una división más elaborada (HEYMAN & al., 2018).

3. Metodología

El método propuesto toma el caso de estudio de la ciudad de Madrid, que es uno de los mercados de viviendas residenciales más activos y poblados de España. El municipio cubre un área de más de 600 kilómetros cuadrados, y está dividido oficialmente en 21 distritos y 136 barrios, que representan la división administrativa de la ciudad que es de uso común por los diferentes

⁷ Por ejemplo, el propuesto por Wu (2020) basado en apren-

dizaje automático.

agentes del mercado inmobiliario: administración pública, ciudadanos, agencias y medios de comunicación

Las fuentes de información utilizadas para el estudio son, por una parte, la información catastral de la ciudad procedente de la sede electrónica del Registro Central del Catastro; y, por otro lado, datos de 115 895 pisos en venta en la ciudad de Madrid, publicados en la web Idealista (IDEALISTA 2022), y correspondientes a los cuatro trimestres de 2018⁸. El fichero de entrada cuenta con varios campos: coordenadas del inmueble (de las que se deriva el barrio en el que está situado), fecha de publicación, precio, área construida y precio por metros cuadrado.

La división en barrios y distritos, anteriormente mencionada, fue creada con fines administrativos, pero se utilizará para el análisis empírico de los precios. En un primer análisis empírico del conjunto de datos de precios de idealista puede apreciarse que la ciudad no es homogénea ni en términos de precios, ni tampoco lo es en términos de su variabilidad. Se observa que los precios medios por barrio toman valores entre los 1337 €/m² y los 4335 €/m², concentrándose los precios más altos en torno al centro y norte de la ciudad, y los más bajos en el sur y este. De la misma manera, la variabilidad de los precios también varía sensiblemente por zona, tomando desviaciones típicas por barrio entre los 34 €/m² y los 114 €/m². Las desviaciones más altas se concentran en la zona oeste y área central (dentro de la M-30).

El dato de precios inmobiliario procede de los datos publicados por el portal Idealista en su política de datos abiertos y que se han utilizado en otros artículos científicos (REY-BLANCO & al., 2023). Al ser un dato generado por los usuarios de la plataforma de internet, está expuesto a problemas de errores de datos, sesgos o duplicidades (LOBERTO & al., 2018), lo cual ha requerido un proceso de limpieza de valores atípicos y deduplicación⁹ antes del cálculo de las medidas de precios.

En general, en el ámbito de la investigación de la vivienda, no existen muchas fuentes que aseguren la suficiente desagregación geográfica y

funcional para realizar análisis en profundidad, lo que dificulta la reproducibilidad de este tipo de estudios. Cuestión que, afortunadamente, parece haber mejorado recientemente con la aparición de fuentes públicas desagregadas, como el índice de precios del alquiler del MITMA, o la propia fuente utilizada en el presente estudio.

El método presentado se estructura en dos pasos: (1) el análisis de los distintos tipos morfológicos que conforman cada uno de los tejidos urbanos de tipo residencial, y su distribución espacial en el municipio de Madrid; (2) el estudio de las diferencias en variabilidad, en los precios, entre las zonas administrativas y las zonas tipomorfológicas obtenidas en el punto anterior.

4. Segmentación tipomorfológica

La propuesta de segmentación intenta lograr un proceso que construya una caracterización tipomorfológica adecuada de forma sistemática, automática, y fácilmente extrapolable a cualquier ciudad¹⁰, que permita, además, identificar e interpretar los diferentes tejidos urbanos de tipo residencial identificados en base a tres condiciones fundamentales:

- Acceso a una fuente de información espacial abierta y homogénea.
- Uso de un índice espacial global y escalable para una captura adecuada de datos y un posterior procesamiento.
- Establecimiento de parámetros lo suficientemente representativos de toda la posible variedad tipomorfológica.

Es importante poner de manifiesto que esta primera aproximación ha utilizado únicamente características tipomorfológicas de fincas catastrales para construir las agrupaciones, con el objetivo de evaluar si este único aspecto es suficiente para explicar la variabilidad de los precios. Sin embargo, esto puede limitar la capacidad de segmentación del método propuesto ya que los tejidos urbanos identificados están insertos en un mercado inmobiliario, sobre el

⁸ La extracción de anuncios corresponde a un conjunto abierto proporcionado por el portal Idealista para las ciudades de Madrid, Barcelona y Valencia, compuesto de datos desagregados del año 2018, y destinado al desarrollo de investigación reproducible del mercado inmobiliario (REY & SELVA-ROYO, 2023).

⁹ Es común encontrar múltiples instancias de la misma propiedad, ya que es una práctica común vender propiedades a través de varios agentes inmobiliarios.

¹⁰ Recientemente, el MITMA ha publicado la plataforma Urban3r (MITMA, 2023c), que clasifica cada edificio catastral en 18 categorías tipológicas, segmentadas a partir del tipo de vivienda, número de plantas en viviendas plurifamiliares y año de construcción. El objetivo de esta fuente abierta es impulsar la regeneración urbana en España, incentivando mejoras en las condiciones energéticas de los edificios (MITMA 2023b, pp 7, 14-15).

que actúan otras variables cuya influencia en los precios y en la formación de la demanda es conocida (presencia de servicios públicos, comunicaciones o áreas comerciales, nivel de ingresos) (BOWES & IHLANFELDT, 2001; REY-BLANCO & al., 2024; GREYER & MIESZKOWSKI, 1980; CAO, 2015).

Para la primera condición, se opta por utilizar el conjunto de datos catastrales españoles en formato GML¹¹ de acuerdo con la Directiva INSPIRE europea 2007/2/EC, 14 de marzo de 2007, L108/1. Específicamente, la información vinculada a las parcelas catastrales (CP) y a los edificios (BU). Distribuidos a través de archivos con un formato homogeneizado, cuenta con información geométrica de las construcciones, por tanto, muy adecuado para nuestro propósito.

En cuanto al segundo aspecto, nuestro enfoque es utilizar un sistema geográfico reticular global discreto (*Discrete Global Grid System, DGGS*) ya existente, aplica una malla compuesta por formas poligonales que cumple con los requisitos de la especificación del *Open Geospatial Consortium* (OGC). En este sentido, sobre el análisis realizado por BONDARUK (2020), en el que se comparan distintas aproximaciones, se decide usar el sistema global H3 desarrollado por la empresa UBER (2018), por su conformidad con OGC y su rapidez de cálculo.

H3 es un índice espacial de carácter discreto y jerárquico, que permite una cobertura eficiente de toda la Tierra y la agrupación de información en diferentes niveles de escala. Su geometría hexagonal también permite una equidistancia eficiente dentro de las celdas vecinas, favoreciendo los procesos de suavizado iterativo. Así,

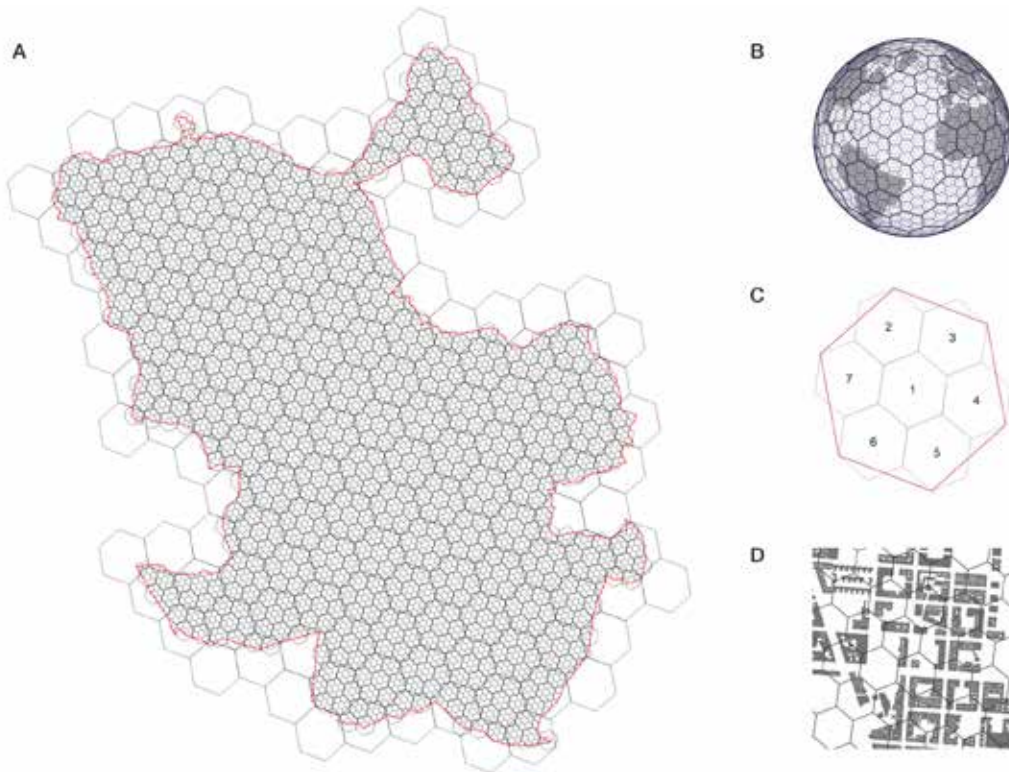


FIG. 1 / Sistema discreto H3¹²

Fuente: Elaboración propia, excepto B (BRODSKY, 2018)

¹¹ *Geographic Markup Language* es un formato de fichero para representar datos geográficos.

¹² A) Límites municipales de la ciudad de Madrid cubierta por los niveles H3: 7, 8 y 9; B) Sistema global discreto; C)

Carácter jerárquico de las celdas hexagonales; D) nivel H3 10: una escala adecuada para la detección de tejidos urbanos.

las dimensiones discrecionales que suelen dibujarse, como la cuadrícula de 500 o 600 metros, podrían mejorarse y ajustarse considerablemente a la realidad morfológica urbana (Fig. 1).

Los parámetros implicados en los diferentes sistemas de caracterización tipomorfológica se han establecido como parte de un conocido proceso de minería de datos en la agrupación urbana (GIL & al. 2012; MERCADÉ & al., 2018; ARRIBAS-BEL & FLEISCHMANN, 2022). Entre los parámetros habituales, RODE & al. (2014) identifican hasta 11, mientras que FLEISCHMANN & al. (2020) identifica en su taxonomía hasta 35 características relevantes. En la Fig. 2 se ha resumido el catálogo de variables tomadas en los distintos estudios previos.

Tras un exhaustivo análisis del estado del arte, de todos ellos hemos elegido finalmente 7 parámetros¹³, suficientemente representativos en vista de los resultados de la agrupación:

- Índice de edificabilidad: referido a los metros de techo construidos por superficie o edificabilidad (i_{edif}). Calculado como la superficie construida multiplicada por la altura, con respecto a la superficie en planta de la parcela, referida a la celda H3 mediante el centroide del elemento constructivo catastral.
- Índice de ocupación: cobertura del suelo por parte de la edificación (i_{ocup}). Calculado como la superficie construida de la planta con respecto al total de la parcela, referida a la celda H3 mediante el centroide de la parcela.
- Densidad de alineación: identifica si las partes del edificio llegan a la línea divisoria entre los espacios parcelados y los continuos (i_{alin}). Calculada como los metros lineales de los perímetros de elementos construidos coincidentes con los límites de la parcela, recayentes a espacios no parcelados (referida

	Este estudio	MERCADÉ & al. (2018)	BERGHAU-SER & al. (2019)	RODE & al. (2014)	ARRIBAS-BEL & FLEISCHMANN (2022)
Edificabilidad	X	X	X	X	X
Ocupación	X	X	X	X	X
Alineaciones	X	X		X	X
Medianeras	X	X		X	X
Compacidad	X	X		X	X
Densidad viaria	X		X	X	X
Antigüedad	X		X		X
Accesibilidad (intensidad)		X			X
Accesibilidad (ocupación)		X			X
Contexto urbano			X		X
Usos del suelo			X		X
Número de alturas				X	X
Proporción de la parcela				X	X
Proporción de la fachada				X	X
Ancho de calle				X	X

Fig. 2 / Variables de forma urbana utilizadas en estudios previos y la metodología propuesta

Fuente: Elaboración propia

¹³ Se reconocen las limitaciones de la parcela catastral en términos de consistencia, como unidad de cálculo de varios parámetros. Si bien es cierto que en situaciones de *edificación abierta*, el registro catastral no es uniforme en sus criterios (las zonas no edificadas pueden recaer sobre parcela independiente o común a la de la edificación— o bien diluirse en el espacio continuo no parcelado, debido a factores

urbanísticos o registrales de difícil identificación), no hay que olvidar que la metodología propuesta se basa en técnicas estadísticas vinculadas a la multidimensionalidad, y que el dato se vuelca en celdas de un tamaño bastante mayor que la parcela. En todo caso, para nuestro propósito, la base de datos catastral es sin duda la fuente mejor informada en términos de cobertura, actualización y homogeneidad.

a la celda H3 mediante el centroide de los correspondientes segmentos lineales).

- Densidad de paredes medianeras: referidas solo a las que tienen más de un piso (i_{medi}). Calculada como los metros lineales de los perímetros de elementos construidos coincidentes con los límites de la parcela, recayentes a espacios parcelados (referida a la celda H3 mediante el centroide de los correspondientes segmentos lineales).
- Compacidad del edificio: relación entre su volumen y envoltorio (i_{comp}). Calculada como relación entre el volumen construido y la superficie de la envoltorio de los edificios por parcela, referida a la celda H3 mediante el centroide de la parcela catastral.
- Densidad de la red viaria: suma de las longitudes de los ejes del viario contenida en cada celda (i_{dv}). Calculada como los metros lineales de segmentos viarios informados por *Open Street Map* dentro de cada celda H3.
- Edad de la construcción: promedio de antigüedad de los edificios involucrados (i_{antig}). Calculada como la mediana del número de años con respecto a la fecha del dato catastral de las parcelas cuyo centroide recae en la celda H3.

Todas estas variables se obtienen después un proceso previo del conjunto de datos cartográficos catastrales. Durante el proceso de supervisión, se evaluó un octavo parámetro consistente en el número de parcelas incluidas en cada celda H3 de nivel 10, pero se omitió por restar estabilidad al proceso de agrupación.

Para el proceso de captura y procesamiento, se utilizó una base de datos en un entorno PostgreSQL con la extensión espacial PostGIS.

Después de cargar las tablas de información y filtrar los identificadores catastrales de las parcelas por uso residencial, se vincularon los valores de los parámetros a los centroides de sus geometrías correspondientes, agrupándose después en celdas H3 para construir los índices respectivos.

Es importante destacar que la elección de una escala adecuada del índice H3 es decisiva para agrupar correctamente la información. La dimensión de las crujiás de los edificios residenciales se mueve en un pequeño rango independientemente de la tipología: las necesidades básicas de ventilación e iluminación de los espacios habitables requieren que las distancias entre fachadas rara vez superen los 25 m, aproximadamente. Por esta razón, se ha considerado que el nivel 10 del índice H3 –de 1,5 ha de superficie, con un radio de hexágono de aproximadamente 75 m– es el más adecuado para una primera agrupación representativa de parámetros tipológicos (FIG. 3), a una escala muy fina, para luego avanzar a diferentes niveles de esta rejilla espacial jerárquica. El objetivo es superar los enfoques variados en la caracterización de la forma urbana, facilitando su delimitación con un sistema de rejilla altamente operativo (BONDARUK & al., 2020) que permita después una amplia gama de escalas de agregación: barrio, distrito, municipio y territorio.

Una vez obtenida la información espacial de los parámetros, se realizó una clasificación por agrupación utilizando el algoritmo de clasificación semi-supervisada K-medias. Como muchos investigadores han observado, el método es muy sensible a su parametrización (GIL & al., 2012; BERNABÉ & al., 2013; VIALARD, 2014; SAMARDZHIEV & al., 2022), y las clases generadas por el método pueden ser muy diferentes en función de la configuración original. El método heurístico K-medias, estandarizado por LLOYD



FIG. 3 / Agrupación con las mismas observaciones con dos niveles distintos H3, 11 (izquierda) y 10 (derecha)

Fuente: Elaboración propia

(1982) y revisado recientemente en el manual de WITTEN (2011), requiere un número inicial de grupos, k (siendo $k \leq n$), estableciendo una partición de las observaciones previas (x_1, x_2, \dots, x_n) de acuerdo con la distancia euclidiana que minimiza la suma de los cuadrados dentro de cada grupo, $S = \{S_1, S_2, \dots, S_k\}$, como se muestra en la siguiente expresión:

$$\text{arg}_S \min \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (1)$$

En nuestro caso, se asume que el rango del número de grupos (k) se encuentra entre 6 y 12, en base a dos de los estudios más significativos recientes (MERCADÉ & al., 2018; BERHAUSER-PONT & al., 2019). En el primero, el método *Spacemate* tiene en cuenta dos medidas clave de densidad: el índice de superficie construida y el índice de superficie de planta baja (FSI y GSI, por sus siglas en inglés), que están relacionados con la altura. Por esta razón, los grupos se identifican con un doble atributo (espacioso, compacto y denso, seguido de bajo, medio o alto) según el caso. Por otro lado, el límite superior del rango, 12, apuntado por MERCADÉ & al., (2018) es mucho más alto debido a su alcance único y más amplio (el Área Metropolitana de Barcelona, con 36 municipios y un total de 636 kilómetros cuadrados). En cualquier caso, es un número mucho más preciso que los simples tres o cuatro grupos habituales de la clasificación detallada de planificación urbana, o las cinco categorías diferentes por tipología establecidas por criterios catastrales (DIRECCIÓN GENERAL DEL CATASTRO, 1993). En este sentido, se han evaluado diferentes configuraciones del método, determinando cual es el número de grupos óptimo para la división, y el número de características tipológicas usadas para definir el tejido urbano¹⁴.

5. Medidas de variabilidad en los precios de zona

Para estimar el grado de reducción en la variabilidad de los precios, se ha comparado una medida estadística de dispersión de los precios para la división administrativa y tipomorfológica. La dispersión representa el grado de desviación media esperable entre un precio medio de zona

y los precios de los inmuebles en el submercado inmobiliario al que pertenece. Por tanto, una dispersión nula expresa una segmentación perfecta, en la que el precio medio se corresponde al precio de cualquier inmueble de la zona. En cambio, una alta dispersión indica que el precio de zona no representa a los inmuebles de dicha zona¹⁵.

El uso de la morfología en la selección de submercados se justifica, desde un punto de vista inmobiliario, por la frecuente heterogeneidad constructiva en los barrios (debido a: diversos años de construcción o calidades constructivas.). Esta ausencia de homogeneidad produce índices de precios menos representativos, por ejemplo, un barrio compuesto por una mitad de edificios antiguos y la otra mitad por obra nueva, podría tener un precio medio que no es aplicable ni al primer segmento de casas ni al segundo.

No obstante, una segmentación del mercado inmobiliario únicamente basada en aspectos constructivos o geográficos no recoge otras fuentes de variabilidad en los precios relacionadas con aspectos como la influencia de la mediación, las distintas elasticidades de precios en los tramos altos y bajos de ingresos (por ejemplo, los inmuebles de tipo premium están sujetos a lógicas de financiación y de mercado que el grueso del mercado). Tampoco recoge la interacción en los precios con las políticas públicas de la vivienda o las motivaciones individuales en la adquisición, que no están sometidas a las reglas de reparto mediante el mercado.

La representación numérica de la dispersión se puede encontrar de muchas formas en la literatura, entre las más comunes: la desviación estándar, el coeficiente de variación, o las relaciones entre cuantiles como el rango intercuartílico (LACH, 2002), a continuación (IQR). La primera, de uso muy habitual, no es recomendable por la naturaleza asimétrica de la distribución de precios (LEUNG, LEONG & WONG, 2006), por tanto, es preferible usar un coeficiente de variación robusto. En nuestro caso, basado en la media geométrica y la media de las semidesviaciones típicas.

Se ha decidido utilizar los precios medios por unidad de superficie (€/m²), en lugar del precio total, por su capacidad de reducir el sesgo

¹⁴ Para este fin se han utilizado los siguientes paquetes del lenguaje python: h3, kmeans, numpy, matplotlib y geopandas.

¹⁵ Esta situación es frecuente en zonas altamente

heterogéneas, en las que el precio medio recoge diferentes segmentos de precio, y cuyo valor no representa correctamente a ninguno de estos segmentos.

introducido al utilizar inmuebles de distinta área en la muestra. Esta aproximación se puede asimilar a un proceso de control de calidad por ajuste mixto (HILL, 2013).

La variabilidad se calcula como el Coeficiente de Variación (CV), cuyo valor se puede inter-

donde P_B es el precio de cualquier inmueble en el barrio B , $\mu(P_B)$ es la media geométrica de precios del barrio y $\sigma_-(P_B)$ su semidesviación típica negativa (desviaciones sobre valores por debajo de la media), y $\sigma_+(P_B)$ la semidesviación positiva.

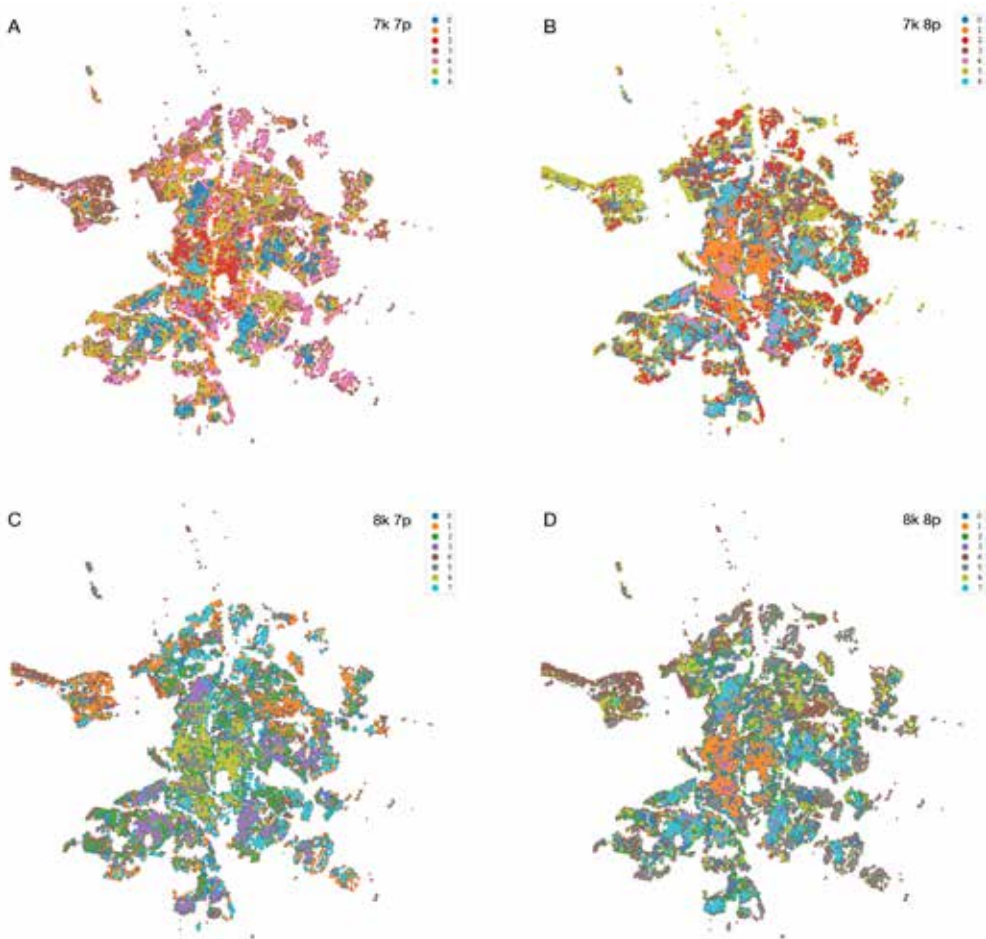


FIG. 4 / Muestras de escenarios de agrupación de K-medias.⁴⁵

Fuente: Elaboración propia

pretar como la diferencia media, en porcentaje, de los distintos inmuebles con respecto al precio medio de la zona. El valor para cada barrio se estima según la expresión:

$$CV_B = \frac{\frac{1}{2}[\sigma_-(P_B) + \sigma_+(P_B)]}{\mu(P_B)} \quad (2)$$

Para comparar la variabilidad dentro de cada barrio incorporando las áreas y los grupos tipomorfológicos, se utiliza una expresión análoga (CV), aplicando como media ponderada el coeficiente de variación de todas las áreas N contenidas en cada área.

$$CV_B^* = \frac{1}{N_B} \sum_{c=1}^C N_{B,c} \frac{\frac{1}{2}[\sigma_-(P_{B,c}) + \sigma_+(P_{B,c})]}{\mu(P_{B,c})} \quad (3)$$

donde $\mu(P_{B,C})$ representa la media del precio de los inmuebles para cada grupo tipomorfológico de los existentes C en el barrio B , y las σ^- y σ^+ las semidesviaciones negativa y positiva para dicho estrato. El ponderador vendría dado por la proporción de inmuebles del grupo en el barrio $N_{B,C}$ con respecto al tamaño total de la muestra N_B .

6. Resultados

La sección de resultados presenta en dos apartados, de acuerdo con estructura de la metodología adoptada. Inicialmente, se muestran los hallazgos concernientes a la agrupación tipomorfológica de las viviendas. Este análisis preliminar sienta las bases para la segunda parte, donde se examina cómo estas clasificaciones tipomorfológicas influyen y reflejan variaciones en los precios inmobiliarios por zonas.

a. Visualización y justificación de los agrupamientos tipomorfológicos

En base a la metodología presentada, se ha generado un gran número de escenarios en función de los parámetros predefinidos (de 6 a 8) y el número de grupos (de 6 a 12), para encontrar el modelo más confiable. Durante el análisis de los diferentes modelos, se detectó que las agrupaciones más adecuadas se lograban cuando el número de grupos se situaba en el rango de 7 a 8 (Fig. 4). Para determinar este número óptimo, se utilizaron tablas descriptivas de los valores medianos por grupo, así como diferentes métricas comparativas como las medidas de Silhouette, Calinski-Harabasz y Davies-Bouldin (LORD & al., 2017).

Se encontró que, aunque los modelos de 7 clústeres (Fig. 4, A y B) obtuvieron métricas ligeramente mejores, la descomposición tipológica se ajustó más a la realidad de los edificios si se consideraban 8 categorías tipomorfológicas diferentes. Una vez tomada esta decisión, los índices correspondientes al modelo de los 7 parámetros descritos anteriormente fueron algo más robustos, siendo más clara su correlación interna. Por lo tanto, el escenario de 8 grupos generado a partir de los 7 parámetros tipomorfológicos mencionados anteriormente, 8k7p (Fig. 4 C), fue finalmente seleccionado como el más adecuado.

Estos grupos recogen, con un alto grado de fidelidad, ocho tejidos urbanos diferentes y fácilmente identificables relacionados con el caso de la ciudad de Madrid (Fig. 5), que a su vez se ajustan a las cuatro grandes familias tipológicas del sur de Europa: tejidos consolidados (divididos según la edad y la intensidad), edificios en manzana (divididos según la ocupación y el número de paredes medianeras), construcción abierta (dividida por compactidad) y viviendas unifamiliares (divididas según la intensidad y la ocupación).

Las categorías finalmente asignadas se enumeran de la siguiente manera:

- Tejido compacto – centros históricos.
- Tejido consolidado – barrios antiguos.
- Manzana compacta – cuadrículas urbanas.
- Manzana abierta – nuevos ensanches.
- Edificación abierta / bloques (grano grande).
- Edificación abierta / bloques (grano fino).
- Viviendas unifamiliares aisladas (intensidad intermedia).
- Viviendas unifamiliares aisladas (baja intensidad).

	tejido compacto CENTRO HISTÓRICO	tejido consolidado BARRIOS ANTIGUOS	manzana compacta ENSANCHE	manzana abierta NUEVOS ENSANCHES	edificac. abierta 1 COMPACTA	edificac. abierta 2 GRANO FINO	vv. unifamiliar 1 MEDIA INTENSIDAD	vv. unifamiliar 2 BAJA INTENSIDAD
kBcls								
i_edif	2.212023	1.595455	2.911873	1.342926	0.639934	1.272083	0.518470	0.339338
i_ocup	0.522468	0.400685	0.467399	0.209273	0.115022	0.242661	0.136010	0.079068
i_alin	0.035695	0.037926	0.021486	0.008334	0.008176	0.011463	0.005960	0.006731
i_medi	0.265717	0.113060	0.116070	0.034030	0.032995	0.071467	0.034959	0.032991
i_comp	0.877686	1.022031	1.288248	2.062925	1.164698	1.042767	1.178011	1.097897
i_dv	0.027021	0.030093	0.021304	0.018837	0.037661	0.021451	0.019843	0.014968
i_antig	85.112853	54.076670	74.325077	26.891230	54.845258	56.569301	23.262619	53.671071

Fig. 5 / Análisis descriptivo de los 8 grupos propuestos, incluye los valores medianos para cada parámetro

Fuente: Elaboración propia

¹⁶ A y B muestran modelos con 7 clústeres diferentes obtenidos a partir de 7 y 8 parámetros respectivamente (7k7p,

7k8p); C y D correspondientes a 8 modelos de clúster (8k7p, 8k8p).

Para medir la contribución de las diferentes variables, se ha estimado el coeficiente de determinación R^2 para cada una de ellas (ESRI, 2023), esta métrica ofrece una estimación numérica de la contribución de cada una de ellas a la construcción de los grupos. Los resultados indican que las variables Ocupación (0,521), Alineaciones (0,477) y Edificabilidad (0,474) son las más determinantes para obtener tejidos urbanos más diferenciados, en cambio la medida de Medianeras (0,063) es la que menos carácter diferencial ofrece. El R^2 adolece de algunas limitaciones en cuanto a su interpretabilidad y comparabilidad, como describen LOPERFIDO & TARPEY (2018), además de no incorporar la influencia de la interacción conjunta de dos o más. Esto puede demostrarse en este caso con la medida “Medianeras”, que, aunque tiene un coeficiente bajo es crucial para distinguir el grupo “Edificación abierta / bloques (grano fino)” del resto.

La nomenclatura de grupos se ha realizado atendiendo principalmente a la descriptiva de los parámetros, tras supervisar más de 20

combinaciones de número de grupos y parámetros implicados. Se ha partido de manera auto-condicionada por las grandes cuatro familias de tipologías, habitualmente presentes en las normativas urbanísticas españolas (manzanas, bloques abiertos y vivienda unifamiliar, además de los cascos históricos), para establecer las calificaciones urbanísticas residenciales. No obstante, es necesario destacar que el proceso se basa en reglas generales combinadas con el criterio experto de los investigadores, por tanto, en la asignación de la nomenclatura, en ocasiones ha primado la importancia relativa de la antigüedad (centro histórico vs barrios antiguos) y en otras la ocupación de suelo (edificación abierta 1 contra abierta 2).

En base a la visualización de la propuesta de regionalización del espacio según las categorías tipológicas, se procedió a refinar los datos (suavizarlos) a través de un criterio de contigüidad celular. Específicamente, se utilizó la función de anillos hexagonales “kRing” en la etapa 1, es decir, se agrupan zonas en base al estudio de

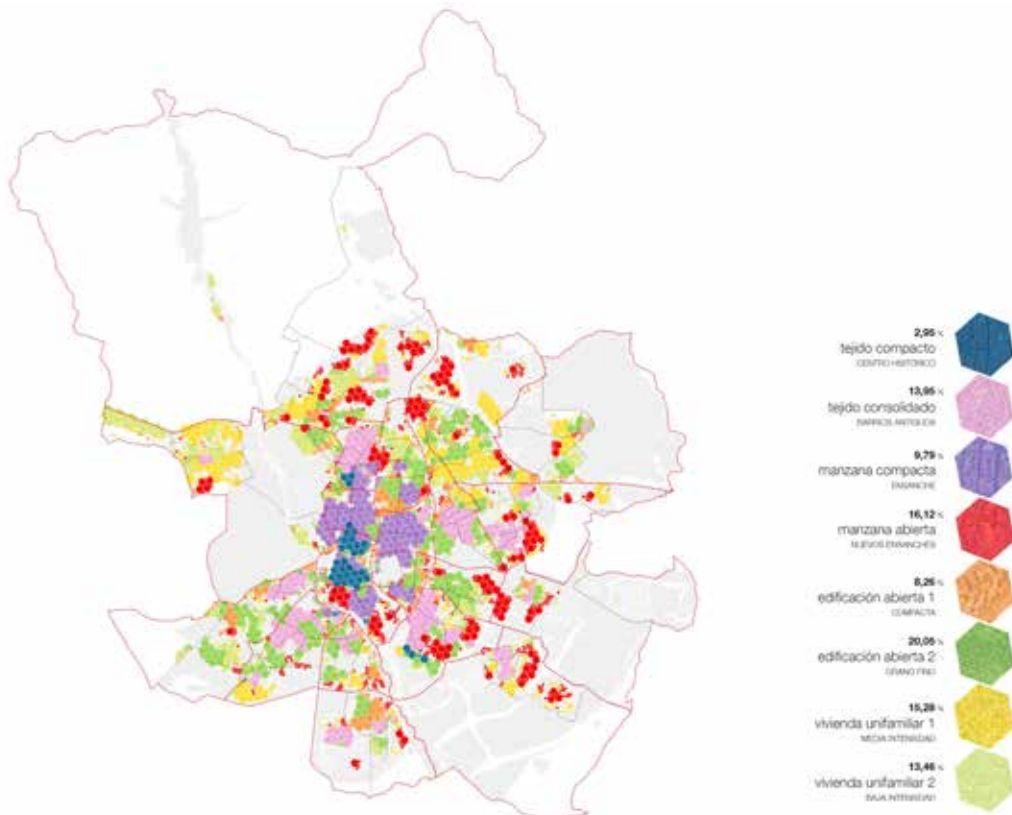


FIG. 6 / Visualización de la agrupación tipomorfológica H3 nivel 10 después del proceso de suavizado (código de colores conforme a la FIG. 5)

Fuente: Elaboración propia

los 6 hexágonos inmediatamente adyacentes, aplicando condiciones de anillo y de borde iterativas, agrupando en hexágonos de nivel superior aquellas zonas adyacentes con los mismos tejidos urbanos.

Este refinamiento permite identificar las características tipológicas dominantes para cada región hexagonal, ya que los usos en la ciudad pueden no estar distribuidos de manera uniforme. Por lo tanto, es relativamente común que la ubicación de equipamientos urbanos, oficinas y otros usos aparte de la vivienda, incluso dentro de un tejido eminentemente residencial, pueda distorsionar las observaciones capturadas a pequeña escala de grano. Adicionalmente, después de aplicar condiciones de suavizado sucesivas y alternas, se completaron las áreas con muestra insuficiente en el nivel 11 (aquellas con solo uno o dos hexágonos con un índice de intensidad residencial no nulo)¹⁷.

Finalmente, el nivel 10 resultante se agrega en celdas de nivel 9, de acuerdo con los criterios mencionados sobre la moda, número de hexágonos vecinos y la información representativa mínima. Además, la biblioteca software H3 proporciona formas de comprimir las geometrías de mallado en resoluciones multinivel (BONDARUK & al., 2020), por lo que se agrupan regiones hexagonales contiguas de la misma clase hasta el nivel 7 (la subdivisión más gruesa de un hexágono padre implica 7 celdas hijas). Por lo tanto, el proceso de compactación de zonas similares cercanas permite identificar agrupaciones zonales de los tejidos urbanos dentro de la ciudad, apuntando nuevas evidencias a favor de futuras

técnicas de regionalización basadas en la tipología (ARRIBAS-BEL & FLEISCHMANN, 2022).

Como se aprecia en la FIG. 6, la ciudad se compone de una gran variedad de tejidos urbanos (en colores distintos), por tanto, las divisiones administrativas engloban realidades morfológicas diversas. Una ventaja asociada de utilizar el enfoque de agrupación en pequeñas regiones hexagonales es que son fácilmente identificables, por ejemplo: “Barrio: Goya, tipo: Manzana compacta (ensanche)”.

b. Variabilidad en los precios

De forma general, los coeficientes de variación (CV_B^*) para barrio y distrito se reducen, como se observa en la FIG. 6. Dicha disminución se produce en todos los distritos y en el 91% de los barrios (124 de 136), con un decremento medio de CV, en términos absolutos, del 2,8%, que representa una mejora del 18%¹⁸.

A modo de ejemplo se muestran, en la FIG. 7, los valores del barrio Simancas dentro del distrito de San Blas, para los 7 grupos generados automáticamente y las medias para el barrio. Se puede comprobar como el efecto de composición genera un precio medio de 2397 €/m², con una alta dispersión de precios (1260 €/m²). En cambio, cuando se desglosa en segmentos, el precio medio es más representativo, y la dispersión individual también se reduce, a excepción de la “Edificación abierta 1 (compacta)”.

En los 12 barrios en los que la dispersión aumenta, esta variación es siempre menor del 5,2%, y en 8 de ellos se incrementa en menos

Área	N	Precio €/m ²	IQR	CV*
Barrio completo	932	2397	1260	18%
Manzana abierta (nuevos ensanches)	539	2051	596	16%
Vivienda unifamiliar 2 (baja intensidad)	169	3479	574	10%
Manzana compacta (ensanche)	100	2352	589	18%
Edificación abierta 1 (compacta)	79	2874	1372	15%
Tejido compacto (centro histórico)	21	2680	549	9%
Tejido consolidado (barrios antiguos)	18	2837	463	6%
Vivienda unifamiliar 1 (media intensidad)	6	4138	494	5%

FIG. 7 / Datos barrio de Simancas (distrito de San Blas) en Madrid

Fuente: Elaboración propia

¹⁷ El orden de aplicación de los pasos afecta significativamente al refinado, en este caso, se decidió aplicar dos rondas consecutivas de dos condiciones de anillo y de borde, y posteriormente, una ronda con una condición de anillo y otra de borde (REY & SELVA-ROYO, 2023).

¹⁸ Ambas magnitudes se expresan en porcentaje, la mejora indica el porcentaje de coeficiente de variación que se reduce.

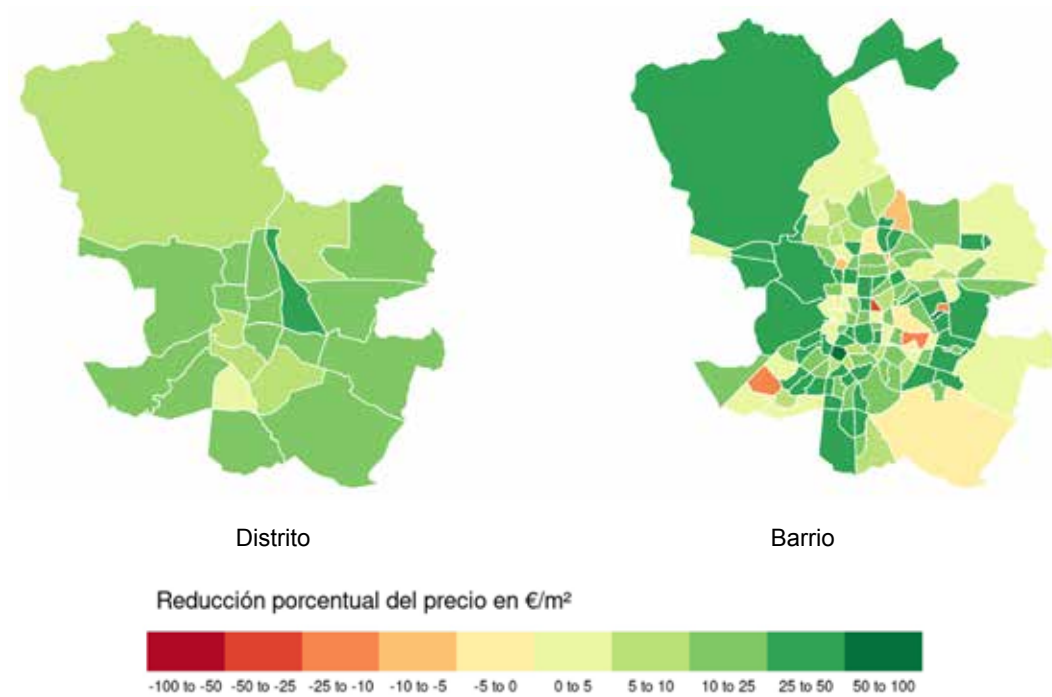


FIG. 8 / Reducción, en porcentaje, del coeficiente de variación en barrio y distrito

Fuente: Elaboración propia

Distrito	Barrios	N	CV* B original	CV* B agrupado	Reducción	Mejora (%)
Ciudad Lineal	9	7273	19,8%	14,6%	5,2%	26,4%
San Blas	8	4360	18,2%	14,0%	4,1%	22,7%
Salamanca	6	10 488	18,6%	14,6%	4,0%	21,4%
Latina	7	5730	18,4%	14,4%	3,9%	21,4%
Vicálvaro	4	1056	17,4%	13,7%	3,6%	21,0%
Moratalaz	6	1638	13,8%	10,5%	3,3%	23,9%
Villaverde	5	4555	20,1%	17,0%	3,1%	15,4%
Retiro	6	4697	20,0%	17,1%	2,8%	14,2%
Moncloa	7	4574	20,1%	17,5%	2,5%	12,6%
Chamberí	6	7295	15,6%	13,1%	2,5%	15,8%
Barajas	5	1172	13,6%	11,2%	2,4%	17,9%
Tetuán	6	6720	20,3%	18,0%	2,2%	10,9%
Carabanchel	8	8312	19,0%	16,9%	2,1%	11,0%
Villa de Vallecas	3	3939	16,9%	15,2%	1,8%	10,5%
Chamartín	6	6424	17,4%	15,7%	1,8%	10,2%
Hortaleza	8	3958	18,5%	17,0%	1,4%	7,8%
Fuencarral	10	4331	16,3%	14,9%	1,4%	8,5%
Centro	6	10 774	16,8%	15,6%	1,2%	7,4%
Puente de Vallecas	6	6921	18,3%	17,2%	1,1%	6,3%
Arganzuela	6	4595	13,7%	12,8%	0,9%	6,4%
Usera	7	3587	20,2%	19,5%	0,6%	3,2%

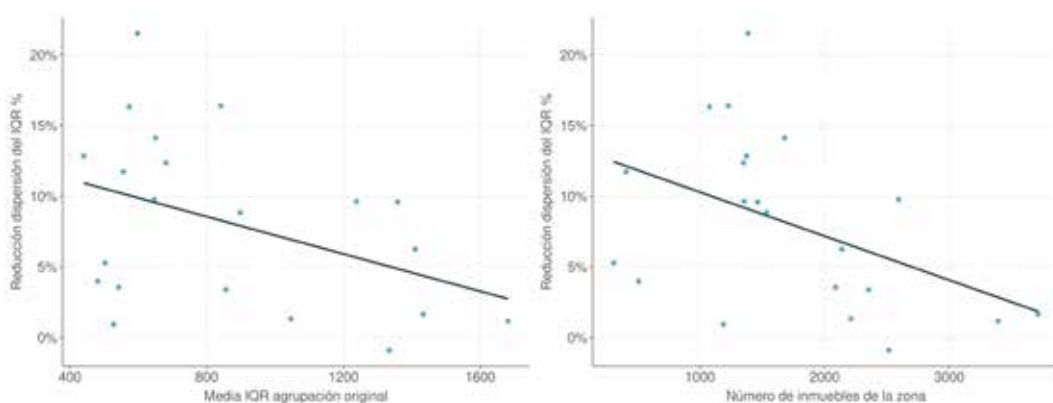
FIG. 9 / Reducción de la dispersión de precios por distritos, calculados sobre precios en Idealista en 2018

Fuente: Elaboración propia

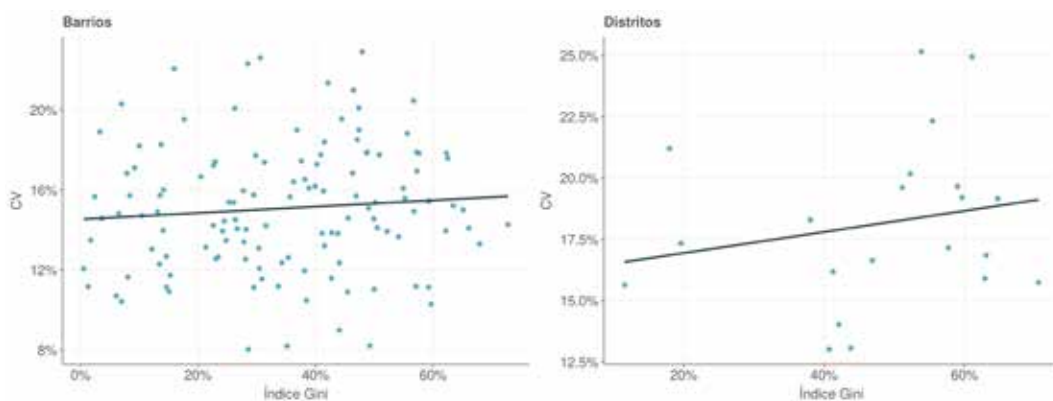
Barrio	Área	N	Precio €/m ²	IQR	CV
Lista	Barrio completo	1678	5487	1696	13,5%
	Vivienda unifamiliar 1 (media intensidad)	1649	5486	1696	13,2%
	Tejido compacto (centro histórico)	29	5526	1002	24,1%
Marroquina	Barrio completo	342	2662	600	12,5%
	Manzana compacta (ensanche)	233	2603	444	9,8%
	Vivienda unifamiliar 2 (baja intensidad)	96	2877	1055	12,1%
	Edificación abierta 1 (compacta)	13	2239	970	22,6%
El Pardo	Barrio completo	45	2315	486	19,5%
	Tejido compacto (centro histórico)	20	2771	1053	7,0%
	Edificación abierta 1 (compacta)	25	2004	736	13,2%

FIG. 10 / Precios medios y variabilidad en los barrios de Lista, Marroquina y El Pardo

Fuente: Elaboración propia



(a) Relación de reducción de variabilidad con el coeficiente de variación original y el precio medio de zona



(b) Relación Índice Gini de grupos y coeficientes de dispersión de precios originales (barrio y distrito)

FIG. 11 / Reducción de la variabilidad en función distintos factores

Fuente: Elaboración propia

del 1%. Los casos que más aumentan en dispersión son, en términos absolutos, Lista (-5,2%), Marroquina (-2,3%) y Águilas (-2,2%); y los más favorables: El Pardo (9,4%), Ambroz (9,2%) y Santa Eugenia (8,4%). (FIG. 8)

Por distritos, como muestra la FIG. 9, la reducción del rango intercuartílico es generalizada, con una media del 2,5% que representa una mejora del 14%.

A nivel de barrio, el lugar donde la segmentación es menos eficaz es Lista, en el distrito de Salamanca. Se aprecia un aumento de la variabilidad de CV del 6% (FIG. 10), esta variación no es muy relevante y se debe por la minoría tipológica "Tejido compacto (centro histórico)". Se comprueba que las diferencias en variabilidad son mínimas cuando el número de segmentos distintos es bajo. De la misma manera, para el caso de Marroquina, en el distrito de Moratalaz, existe un empeoramiento del CV del 13% al 15% (FIG. 10), y se observa que también un segmento minoritario afecta la medida de dispersión.

En el caso contrario, el barrio de El Pardo es el que más mejora, con una reducción del 10% de la dispersión original (FIG. 10). Se puede apreciar cómo la dispersión global de barrio está en un orden de magnitud mucho mayor que el de los segmentos, y es el efecto de composición de distintos estratos el que hace que el precio medio de barrio tenga una alta incertidumbre.

El análisis de la reducción de la variabilidad final con respecto a la dispersión original y el precio medio por zona (FIG. 11 a), muestra que la tasa de reducción es mayor en las zonas con menor variabilidad original y en aquellas que tienen precios menores. No obstante, en cualquiera de los dos casos, estas medidas cuentan con alto grado de incertidumbre y cuyo estudio requerirá un análisis en mayor profundidad.

La distribución de los tamaños de los grupos tipomorfológicos en cada área contribuyen a la dispersión, aunque no es el factor determinante, puesto esta no está relacionada solo con la distribución de sus pesos, sino por las diferencias que tienen entre ellos. Esto se puede derivar de la FIG.11b, que relaciona el coeficiente de variación con el coeficiente de Gini¹⁹ para barrio y distrito. Se observa una ligera relación positiva entre una mayor variabilidad de los precios y la desigualdad en los pesos relativos de los grupos más evidente a nivel de distrito. Pero esta relación no es excesivamente fuerte, por lo que se puede concluir

que, aunque tiene influencia, la contribución al desorden de precios no viene determinada solamente por el tamaño de los grupos, sino que también por las diferencias entre ellos.

7. Discusión

En base a los resultados, puede comprobarse que una parte relevante de la dispersión en los precios medios de la vivienda puede atribuirse a la heterogeneidad de las viviendas en la zona, que coincide con lo propuesto por LEUNG & al. (2006). Además, se comprueba que cuanto menores son los precios, mayor es la exposición a la variabilidad. Como hipótesis adicional, se plantea que una segunda fuente de variabilidad podría ser la desigualdad en las condiciones de accesibilidad a servicios de la zona, ya que estos aspectos son determinantes en el precio, como en el caso de la cercanía a medios de transporte (BOWES & IHLANFELDT, 2001; REY-BLANCO & al., 2024), el acceso a servicios públicos (GREYER & MIESZKOWSKI, 1980) o la proximidad de áreas comerciales (CAO, 2015).

El estudio demuestra, además, que la división administrativa, en una ciudad como Madrid, considera las unidades como un conglomerado de elementos inmobiliariamente muy diferentes. Estas divisiones se toman como base en el cálculo de los índices de la vivienda INE (2022), pero no representan unidades homogéneas en términos de precio como se describe en el apartado anterior. Por tanto, la ausencia de control es un potencial inductor de distorsiones de estimación en los precios medios e índices de precios desarrollados por entidades públicas o privadas (FOTOCASA, 2021; IDEALISTA, 2022; MITMA, 2020, 2023a).

Una adecuada estimación de los precios y su análisis en el tiempo ha adquirido un gran interés después de la crisis del COVID-19 (KEENAN, 2020). La importancia de las características de la casa, después de las experiencias de confinamiento, señala que el mercado inmobiliario podría ser cada vez más sensible a este tipo de segmentación (ROGERS & POWER, 2020).

En cuanto al nivel de detalle, y en particular sobre el proceso de suavizado, se ha encontrado que la condición de suavizado en tres rondas es suficiente para eliminar el ruido de agrupación. De hecho, en ese nivel de H3 se encuentra la reducción

¹⁹ El coeficiente Gini es una medida de desigualdad entre grupos, tiene diversas aplicaciones, entre ellas, es de uso

común para medir la desigualdad de la riqueza en la población.

mayor y más estable de la variabilidad de precios, que difumina la composición de las áreas.

Es posible que un filtrado insuficiente de los usos residenciales o errores en los datos de entrada catastral, tanto cualitativos como cartográficos, puedan distorsionar ligeramente los resultados del *clustering* tipomorfológico. El caso de uso determina el nivel de detalle o calidad de datos mínima aceptable: por ejemplo, para el desarrollo de análisis cuantitativos a escala territorial sería recomendable un proceso de suavizado y compactación previa. Sin embargo, debido a la naturaleza discreta de las observaciones de precios, una aproximación directa de los modelos propuestos sería adecuada para la identificación de tendencias en los submercados residenciales.

Conviene tener en cuenta una eventual menor fiabilidad en el grano fino (celdas H3 aisladas de nivel 10) en las zonas con mayor presencia de tejidos de bloque abierto, al hilo de la ya comentada falta de consistencia de la fuente del dato catastral en el tratamiento parcelario. Si bien, dicha cuestión se mitigaría, en parte, por la capacidad de las técnicas de *clustering* de suavizar los efectos de la heterogeneidad de criterios. En cualquier caso, a la vista de los resultados obtenidos²⁰, puede considerarse una cuestión relativamente menor, máxime cuando el estudio realiza una agregación posterior en niveles 9 de H3 y un tratamiento estadístico de la dispersión con relación a la segmentación administrativa.

En este caso específico, el método de agrupación se aplica para refinar los índices de precios de la vivienda, lo que apunta a nuevas posibilidades de investigación en el campo de la economía urbana más allá de la caracterización tipológica. De hecho, la superación de las fronteras administrativas preestablecidas facilitará una comprensión más precisa del mercado inmobiliario, resolviendo los sesgos de selección vinculados al problema de la unidad de área modificable (MANLEY, 2014).

En ese sentido, la agrupación propuesta en el nivel 9 de H3 después del proceso de suavizado iterativo también podría utilizarse como criterio para determinar zonas homogéneas en la ciudad, también descritas como áreas santuario (una parte reconocible del tejido urbano rodeada por calles principales), siguiendo el término introducido por DIBBLE (2019) y MEHAFFY (2010). Así, la segmentación de submercados podría evolucionar desde un carácter espacialmente

discreto a un conjunto claramente identificable de nuevas fronteras intra-distrito que faciliten su adopción por el público en general.

Además, debido a que el enfoque geométrico muestra limitaciones en las áreas frontera, se apunta a la conveniencia de incorporar a futuro el concepto de *Spatial Signature* (ARRIBAS-BEL & FLEISCHMANN, 2022) para crear áreas geométricas específicas, y determinar de forma dinámica los atributos utilizados en la segmentación. Y es que, dada la existencia de heterogeneidad espacial en la información inmobiliaria, una clasificación única podría ser incompatible con regiones con características diversas. En esta línea, para los casos de estudio, se comparará el método automático con la segmentación tipológica de Urban3r (MITMA, 2023c), para evaluar su potencial aplicación en la zonificación de precios.

Finalmente, debemos poner de manifiesto las limitaciones derivadas de haber caracterizado la realidad inmobiliaria solo desde el ángulo residencial, y que, aunque es la más numerosa en el ámbito de estudio, está interrelacionada con la presencia y comportamiento de mercado de otros inmuebles como, por ejemplo, los de tipo terciario o espacios de uso público. Esta cuestión, junto con la posición en la estructura urbana y de comunicaciones (centralidad, conectividad, etc.), la influencia de aspectos demográficos y socioeconómico serán objeto de posteriores investigaciones.

8. Conclusiones

En primer lugar, se pueden apreciar algunos avances en la investigación morfológica a través de la conexión entre los procesos de agrupación de minería urbana descritos y la adopción de un índice espacial global y versátil, como ha demostrado ser H3. Así, en lugar de calcular grupos a partir de la información de parcelas y visualizarlos por edificios (BERGHAUSER-PONT & al., 2019), la agrupación propuesta permite identificar fácilmente tejidos urbanos más complejos centrados en la función residencial.

De hecho, debido a su carácter principal en el crecimiento y transformación urbana, el concepto de tejido urbano atrae actualmente las miradas de los investigadores, adoptándose cada vez más en oposición a la zonificación basada en el uso (KROPF, 2017). De acuerdo con RODE

²⁰ Se observa que zonas de bloque abierto con una definición parcelaria muy dispar por parte de Catastro (como, por ejemplo, el barrio de Colonia Jardín y algunos sectores de

San Blas), reciben tras el *clustering* una idéntica asignación de 'edificación abierta de grano fino' incluso en el nivel 10.

& al. (2014), entendemos que una identificación correcta del tejido urbano puede arrojar más luz sobre el carácter urbano y las expectativas económicas que un enfoque tipológico simple. Para el caso particular de la ciudad de Madrid, la aplicación de nuestro método ayudaría a perfeccionar las políticas públicas, facilitando la asignación eficiente de recursos y la atención precisa a las necesidades de cada área, con un control continuo cualquier desequilibrio en el mercado que pueda afectar a las condiciones de accesibilidad a la vivienda.

Las líneas de trabajo futuras involucrarán la desagregación de la segmentación sobre el nivel de finca catastral, incorporando el modelo de precios hedónicos en el proceso, incluir nuevas variables, además, de extender el análisis en otras ciudades españolas y áreas rurales. En particular se incorporarán los aspectos limitantes recogidos en el texto como son la incorporación de tipologías no residenciales, la estructura espacial de la ciudad, accesibilidad a servicios y comunicaciones, aspectos sociodemográficos, económicos.

Por último, es relevante señalar que se ha establecido un conjunto homogéneo de tipologías urbanas basadas en parámetros morfológicos del dato espacial catastral, todos ellos disponibles en el formato de la directiva europea INSPIRE. Y aunque el estudio de caso ha tenido en cuenta la realidad constructiva española, esto podría favorecer estudios de replicabilidad entre ciudades de diferentes países que podrían abordarse en el futuro, y abrir nuevas posibilidades para la investigación cuantitativa urbana y su aplicación a una amplia gama de disciplinas como los estudios inmobiliarios, de sostenibilidad o de planificación urbana.

9. Bibliografía

- ADAIR, A. S. & BERRY, J. N. & MCGREAL, W. S. (1996): Hedonic modelling, housing submarkets and residential valuation. *Journal of Property Research*, 13 (1): 67-83. <https://doi.org/10.1080/095999196368899>
- ANSELIN, L. & GRIFFITH, D.A. (1988): Do spatial effects really matter in regression analysis? *Papers in Regional Science*, 65 (1): 11-34.
- ARRIBAS BEL, D. (2014): Accidental, open and everywhere: Emerging data sources for the understanding of cities. *Applied Geography*, 49: 45-53. <https://doi.org/10.1016/j.apgeog.2013.09.012>
- ARRIBAS BEL, D. & FLEISCHMANN, M. (2022): *Spatial Signatures – Understanding (urban) spaces through form and function*. Habitat International 128: 102641. <https://doi.org/10.1016/j.habitatint.2022.102641>
- ARRIBAS BEL, D. & GREEN M. & ROWE, F. & SINGLETON, A. (2021): Open data products-A framework for creating valuable analysis ready data. *Journal of Geographical Systems*, 23 (4): 497-514. <https://doi.org/10.1007/s10109-021-00363-5>
- BATTY, M. (2013): *The New Science of Cities*. Cambridge, Massachusetts: MIT Press.
- BERGHAUSER PONT, M. & HAUPT, P. (2007): *The Spacemate: Density and the Typomorphology of the Urban Fabric*. IOS Press.
- BERGHAUSER PONT, M. & HAUPT, P. (2009): *Space, Density and Urban Form*. PhD, Technische Universiteit Delft.
- BERGHAUSER PONT, M. & MARCUS, L. (2015): *What can typology explain that configuration cannot?* SSS10 Proceedings of the 10th International Space Syntax Symposium.
- BERGHAUSER PONT, M. & STAVROULAKI, G. & MARCUS, L. (2019): Development of urban types based on network centrality, built density and their impact on pedestrian movement. *Environment and Planning B: Urban Analytics and City Science*. <https://doi.org/10.1177/2399808319852632>
- BERNABÉ, A. & CALMET I. & MUSY, M. & BOCHER, E. & ANDRIEU, H. (2013): *Classification automatique des tissus urbains par la méthode des nuées dynamiques*. En 31^{èmes} rencontres universitaires de l'AUGC.
- BOBKOVA, E. & BERGHAUSER-PONT, M. & MARCUS, L. (2019): Towards analytical typologies of plot systems: Quantitative profile of five European cities. *Environment and Planning B: Urban Analytics and City Science*. <https://doi.org/10.1177/2399808319880902>
- BOBKOVA, E. & MARCUS, L. & BERGHAUSER-PONT, M. (2017): *Multivariable measures of plot systems: Describing the potential link between urban diversity and spatial form based on the spatial capacity concept*. En Proceedings – 11th International Space Syntax Symposium, SSS 2017.
- BOEING, G. (2018): Measuring the complexity of urban form and design. *Urban Design International*, 23: 281-92. <https://doi.org/10.1057/s41289-018-0072-1>
- BOEING, G. (2019): Spatial information and the legibility of urban form: Big data in urban morphology. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2019.09.009>
- BONDARUK, B. (2020): *Discrete Global Grid Systems: Operational Capability of the Current State of the Art*.
- BONDARUK, B. & STEVEN A. R. & ROBERTSON C. (2020): Assessing the state of the art in Discrete Global Grid Systems: OGC criteria and present functionality. *Geomatica*, 74: 9-30. <https://doi.org/10.1139/geomat-2019-0015>
- BOURASSA, S. C. & HOESLI M. & MERLIN L. & RENNE J. (2021): Big data, accessibility and urban house prices. *Urban Studies*, 58 (15): 3176-95.
- BOURASSA, S. C. & HOESLI M. & PENG V. S. (2003): Do housing submarkets really matter? *Journal of Housing Economics*, 12 (1): 12-28. [https://doi.org/10.1016/S1051-1377\(03\)00003-2](https://doi.org/10.1016/S1051-1377(03)00003-2)
- BOWES, D. R. & IHLANFELDT K.R. (2001): Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics*, 50 (1): 1-25.
- CANIGGIA, G. & MAFFEI, G. L. (1979): *Composizione architettonica e tipologia edilizia. 1. Lettura dell'edilizia di Base. 2. Il Progetto nell'edilizia di Base*. Venezia: Marsilio Editori.
- CAO, J. A. (2015): *The Chinese real estate market: Development, regulation and investment*. Routledge.
- CARPIO PINEDO, J. & RAMÍREZ G. & SALAS-MONTES & LAMIQUIZ P.L. (2020): New Urban Forms, Diversity, and Computational Design: Exploring the Open Block. *Journal of Urban Planning and Development*, 146 (2): 4020002. [https://doi.org/10.1061/\(ASCE\)UP.1943-5444.0000555](https://doi.org/10.1061/(ASCE)UP.1943-5444.0000555)

- d'Acci, L. (2019): Quality of urban area, distance from city centre, and housing value. Case study on real estate values in Turin. *Cities*, 91: 71-92.
- DIBBLE, J. & PRELORENDJOS, A. & ROMICE O. & ZANELLA, M. & STRANO E. & PAGEL, M. & PORTA, S. (2019): On the origin of spaces: Morphometric foundations of urban form evolution. *Environment and Planning B: Urban Analytics and City Science*, 46 (4): 707-30. <https://doi.org/10.1177/2399808317725075>
- DIRECCIÓN GENERAL DEL CATASTRO (1993): *Normas Técnicas de Valoración*.
- Directiva Inspire 2007/2/EC, de marzo de 2007 (L 108/1).
- ECHAVES GARCÍA, A. & MARTÍNEZ DEL OLMO, A. (2021): Emancipación residencial y acceso de los jóvenes al alquiler en España: Un problema agravado y su diversidad territorial. *Ciudad y Territorio Estudios Territoriales*, 53 (M): 27-42. <https://doi.org/10.37230/CyTET.2021.M21.02>
- ESRI (2023): How Grouping Analysis Works. [08-11-2023]. ESRI. <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/how-grouping-analysis-works.htm>
- FLEISCHMANN, M. & ROMICE O. & PORTA S. (2020): Measuring urban form: Overcoming terminological inconsistencies for a quantitative and comprehensive morphologic analysis of cities. *Environment and Planning B: Urban Analytics and City Science*, 48 (8). <https://doi.org/10.1177/2399808320910444>
- FLEISCHMANN, M. & FELICCIOTTI, A. & ROMICE, O. & PORTA, S. (2021): Methodological foundation of a numerical taxonomy of urban form. *Environmental and Planning B: Urban Analytics and City Science*, 49 (4). <https://doi.org/10.1177/23998083211059835>
- FLEISCHMANN, M. & ARRIBAS-BEL, D. (2022): Geographical characterisation of British urban form and function using the spatial signatures framework. *Scientific Data*, 9 (546, 1). <https://doi.org/10.1038/s41597-022-01640-8>
- FOTOCASA (2021): *Número de anuncios publicados en fotocasa* [10/09/2021]. <https://www.fotocasa.es/>
- GABRIEL, S. A. & WOLCH, J. R. (1984): Spillover effects of human service facilities in a racially segmented housing market. *Journal of Urban Economics*, 16: 339-50. [https://doi.org/10.1016/0094-1190\(84\)90031-7](https://doi.org/10.1016/0094-1190(84)90031-7)
- GIL, J & BEIRÃO, J. N. & MONTENEGRO, N. & DUARTE, J. P. (2012): On the discovery of urban typologies: Data mining the many dimensions of urban form. *Urban Morphology*, 16 (1): 27-40.
- GOODMAN, A. C. (1981): Housing submarkets within urban areas: definitions and evidence. *Journal of Regional Science*, 21 (2): 175-85.
- GOODMAN, A. C. & THIBODEAU, T. G. (1998): Housing Market Segmentation. *Journal of Housing Economics*, 7 (2): 121-43. <https://doi.org/10.1006/jhec.1998.0229>
- GREYER, D. M. & MIESZKOWSKI, P. (1980): The effects of nonresidential land uses on the prices of adjacent housing: Some estimates of proximity effects. *Journal of Urban Economics*, 8 (1): 1-15.
- GRILICHES, Z. (1961): Hedonic price indexes for automobiles: An econometric of quality change. En *The price statistics of the federal government*, 173-96. NBER.
- HANCOCK, K. & JONES, C. & MUNRO, M. & SATSANGI, M. & MCGUCKIN, A. (1991): *Housing Costs and Subsidies in Glasgow: The Impact of Housing Subsidies in the Glasgow Travel-to-work Area*. Joseph Rowntree Foundation York Publishing Services, York.
- HAYWARD, R. (1987): *The use of housing tissues in urban design*. *Urban Design Quarterly*, December: 4-6.
- HEIKKILA, E. & GORDON P. & KIM, J. I., PEISER, R.B. & RICHARDSON, H. W. & DALE-JOHNSON, D. (1989): What Happened to the CBD-Distance Gradient?: Land Values in a Policentric City. *Environment and Planning A*, 21 (2): 221-32.
- HEYMAN, A. & LAW S., & BERGHAUSER-PONT M. (2018): How is Location Measured in Housing Valuation? A Systematic Review of Accessibility Specifications in Hedonic Price Models. *Urban Science*, 3 (1): 3. <https://doi.org/10.3390/urbansci3010003>
- HILL, R. J. (2013): Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of Economic Surveys*, 27 (5): 879-914. <https://doi.org/10.1111/j.1467-6419.2012.00731.x>
- HILL, R. J. & SCHOLZ, M. (2018): Can Geospatial Data Improve House Price Indexes? A Hedonic Imputation Approach with Splines. *Review of Income and Wealth*, 64 (4): 737-56. <https://doi.org/10.1111/roiw.12303>
- HILLIER, B., & J. HANSON (1988): *The social logic of space*. <https://doi.org/10.4324/9780429450174-9>
- IDEALISTA (2022): *Conjunto de datos desagregados de anuncios publicados en el portal para 2018*. <http://www.idealista.com/data>
- IDEALISTA (2023): *Datos del informe de precios del mercado inmobiliario de Idealista – diciembre 2022*. <http://www.idealista.com/data>
- Instituto Nacional de Estadística, INE (2022): *Datos del Índice de Precios de la Vivienda – diciembre 2022*. https://www.ine.es/prensa/ipv_tabla1.htm
- KEENAN, J. M. (2020): COVID, resilience, and the built environment. *Environment Systems and Decisions*. <https://doi.org/10.1007/s10669-020-09773-0>
- KIRMAN, A. P. & VRIEND, N.J. (2001): Evolving market structure: An ACE model of price dispersion and loyalty. *Journal of Economic Dynamics and Control*, 25 (3-4): 459-502.
- KOSTOF, S. (1991): *The City Shaped: Urban Patterns and Meanings through History*. New York, NY: Bulfinch Press.
- KROPP, K. (2017): *The Handbook of Urban Morphology*. Chichester: John Wiley & Sons. <https://doi.org/10.1002/9781118747711>
- LACH, S. (2002): Existence and persistence of price dispersion: an empirical analysis. *Review of economics and statistics*, 84 (3): 433-44.
- LAW, S. & STONOR T. & LINGAWI, S. (2013): Urban value: Measuring the impact of spatial layout design using space syntax. En *2013 International Space Syntax Symposium*.
- LEUNG, C. & YUI, K. & LEONG, Y. C. F. & WONG, S.K. (2006): Housing price dispersion: an empirical investigation. *The Journal of Real Estate Finance and Economics*, 32 (3): 357-85.
- LLOYD, S. P. (1982): Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*. <https://doi.org/10.1109/TIT.1982.1056489>
- LOBERTO, M. & LUCIANI A. & PANGALLO M. & al. (2018): *The potential of big housing data: an application to the Italian real-estate market*. Banca d'Italia, Eurosistema.
- LOPERFIDO, N. & TARPEY, T. (2018): Some remarks on the R2 for clustering. *Statistical Analysis and Data Mining* 11 (3): 135-48.
- LORD, E. & WILLEMS, M. & LAPOINTE, F.J. & MAKARENKOV, V. (2017): Using the stability of objects to determine the number of clusters in datasets. *Information*

- Sciences, 393: 29-46. <https://doi.org/10.1016/j.ins.2017.02.010>
- MACLENNAN, D. (1987): *The efficient market framework and real estate economics (mimeo)*. University of Glasgow, Glasgow.
- MACLENNAN, D & TU, Y (1996): Economic perspectives on the structure of local housing systems. *Housing Studies*, 11 (3): 387-406.
- MANLEY, D. (2014): Scale, aggregation, and the modifiable areal unit problem. *Handbook of Regional Science*, editado por M. M. Fischer y Nijkamp P., 1157-71. Heidelberg: Springer. https://doi.org/10.1007/978-3-642-23430-9_69
- MARTÍN CONSUEGRA, F. & DE FRUTOS, F. & HERNÁNDEZ AJA A. & OTEIZA, I. & ALONSO, C. & DE FRUTOS, B (2021): Utilización de datos catastrales para la planificación de la rehabilitación energética a escala urbana: aplicación a un barrio ineficiente y vulnerable de Madrid. *Ciudad y Territorio Estudios Territoriales*, 54 (211): 115-36. <https://doi.org/10.37230/CyTET.2022.211.7>
- MEHAFFY, M. & PORTA S. & ROFÉ, Y. & SALINGAROS, N. (2010): Urban nuclei and the geometry of streets: The 'emergent neighborhoods' model. *Urban Design International*, 15 (1): 22-46. <https://doi.org/10.1057/udi.2009.26>
- MERCADÉ, A. & MAGRINYÀ, J.F. & Y CERVERA, M. (2018): Descifrando la forma urbana: un análisis de patrones de agrupamiento basado en SIG. *GeoFocus Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 22: 3-19. <https://doi.org/10.21138/gf.612>
- Ministerio de Transportes, Movilidad y agenda Urbana, MITMA (2020): «Sistema Estatal de Índices de Referencia del Precio del Alquiler de Vivienda». <http://www.fomento.gob.es/be2/?nivel=2&orden=34000000>
- Ministerio de Transportes, Movilidad y agenda Urbana, MITMA (2023a): *Estadísticas de transacciones inmobiliarias*. <http://www.fomento.gob.es/be2/?nivel=2&orden=34000000>
- Ministerio de Transportes, Movilidad y agenda Urbana, MITMA (2023b): *Segmentación en clústeres tipológicos y caracterización geométrica del parque residencial de viviendas en España (URBAN3R)*. https://cdn.mitma.gob.es/portal-web-drupal/planes_estartegicos/1_2020_segmentacion_parque_residencial_clusteres.pdf
- Ministerio de Transportes, Movilidad y agenda Urbana, MITMA (2023c): *URBAN3R Plataforma de Datos Abiertos para impulsar la regeneración urbana en España*. <https://www.mitma.gob.es/arquitectura-vivienda-y-suelo/urbanismo-y-politica-de-suelo/urban3r>
- PÁEZ, A. (2021): Open spatial sciences: an introduction. *Journal of Geographical Systems*. Springer.
- PANERAI, P. & CASTEX, J. & DEPAULE, J.C. & SAMUELS, I. & SAMUELS, O.V. (2012): *Urban forms: The death and life of the urban block*. <https://doi.org/10.4324/9780080481548>
- REY BLANCO, D. & SELVA ROYO, J.R. (2023): *Housing submarket segmentation house price dataset. A typology-based case study in Madrid [Dataset]*, Mendeley Data. <https://doi.org/10.17632/ksypsm2zh3>
- REY BLANCO, D. & ARBUÉS, P. & LÓPEZ, F.A. & PÁEZ, A. (2023): Using machine learning to identify spatial market segments. A reproducible study of major Spanish markets. *Environment and Planning B: Urban Analytics and City Science*, Online First. <https://doi.org/10.1177/23998083231166952>
- REY BLANCO, D. & ZOFIO, J.L. & GONZÁLEZ-ARIAS, J. (2024): Improving hedonic housing price models by integrating optimal accessibility indices into regression and random forest analyses. *Expert Systems with Applications* 235 (121059). <https://doi.org/10.1016/j.eswa.2023.121059>
- RODE, P. & KEIM C. & ROBBAZZA, G. & VIEJO, P. & SCHOFIELD, J. (2014): Cities and energy: Urban morphology and residential heat-energy demand. *Environment and Planning B: Planning and Design*, 41 (1): 138-62. <https://doi.org/10.1068/b39065>
- RODRÍGUEZ LÓPEZ, J. (2022a): El mercado de vivienda resiste las primeras consecuencias de la guerra de Ucrania. *Ciudad y Territorio Estudios Territoriales*, 54 (213): 743-56. <https://doi.org/10.37230/CyTET.2022.213.13>
- RODRÍGUEZ LÓPEZ, J. (2022b): Mercado de vivienda y coyuntura económica general. *Ciudad y Territorio Estudios Territoriales*, 54 (214): 1027-38. <https://doi.org/10.37230/CyTET.2022.214.13>
- ROGERS, D. & POWER, E. (2020): *Housing policy and the COVID-19 pandemic: the importance of housing research during this health emergency*. <https://doi.org/10.1080/19491247.2020.1756599>
- ROSEN, S (1974): Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy* 82 (1): 34-55.
- SAMARDZHIEV, K. & FLEISCHMANN, M. & ARRIBAS-BEL, D. & CALAFIORE, A. & ROWE F. (2022): Functional signatures in Great Britain: A dataset. *Data in Brief* 43: 108335. <https://doi.org/10.1016/j.dib.2022.108335>
- STONOR, T. (2014): Essay: Space syntax: A SMART approach to urban planning, design and governance. *A+U Architecture and Urbanism*, n.º 530: 12-21.
- SUBDIRECCIÓN GENERAL DE POLÍTICAS URBANAS, VIVIENDA Y SUELO, MINISTERIO DE FOMENTO (2019): *Estudio (01) para la ERESEE 2020 "Estrategia a largo plazo para la Rehabilitación Energética en el Sector de la Edificación en España"*. https://cdn.mitma.gob.es/portal-web-drupal/planes_estartegicos/1_2020_segmentacion_parque_residencial_clusteres.pdf
- Uber Inc. (2018): H3: A hexagonal hierarchical geospatial indexing system. <https://uber.github.io/h3/#/>
- VIALARD, A. (2014): Typological atlases of block and block-face. En *ISUF 21st International Seminar of the Urban Form*.
- WATKINS, C. A. (2001): The definition and identification of housing submarkets. *Environment and Planning A*, 33: 2235-53. <https://doi.org/10.1068/a34162>
- WITTEN, I. H. & FRANK, E. & HALL, M.A. (2011): *Data Mining: Practical Machine Learning Tools and Techniques*, Third Edition.
- WU, Y. & WEI, Y.D. & LI H. (2020): Analyzing spatial heterogeneity of housing prices using large datasets. *Applied Spatial Analysis and Policy*, 13 (1): 223-56.

10. Listado de Acrónimos/Siglas

CV	Coefficiente de Variación
DGGS	Discrete Global Grid System (Sistema de Mallas Globales y Discretas)
INE	Instituto Nacional de Estadística
IQR	Interquartile Range (rango intercuartílico)
IPV	Índice de Precios de la Vivienda
MITMA	Ministerio de Transportes, Movilidad y Agenda Urbana.
OGC	Open Geospatial Consortium.