UNIVERSIDAD DE GRANADA

TESIS DOCTORAL

El efecto de la localización sobre el valor de la vivienda. Los casos de Santa Marta y Atenas.

IVAN DARIO TAMARIS TURIZO

2023

DIRECTORES:

Jorge Chica Olmo

Rafael Cano Guervós

Editor: Universidad de Granada. Tesis Doctorales
Autor: Iván Darío Tamaris Turizo
ISBN: 978-84-1195-364-1
URI: https://hdl.handle.net/10481/92833

El doctorando / The *doctoral candidate* **[Iván Dario Tamaris Turizo]** y los directores de la tesis / and the thesis supervisor/s: **[Jorge Chica Olmo y Rafael Cano Guervós]**

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

/

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor/s and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Lugar y fecha / Place and date:

Granada a 24 de noviembre de 2023

Director/es de la Tesis / *Thesis supervisor/s;*

Jorge Chica Olmo

Firma / Signed

Rafael Cano Guervós

Firma / Signed

Doctorando / *Doctoral candidate*:

Ivan Dario Tamaris Turizo

Firma / Signed

AGRADECIMIENTOS

A mis directores, Jorge Chica Olmo y Rafael Cano Guervós, por animarme, ofrecerme su confianza y guiarme para llevar a cabo este proyecto; aceptar este reto ha sido muy satisfactorio y el inicio a un mundo de conocimiento completamente nuevo para mí, donde ellos han sido un apoyo fundamental para la realización de estos estudios.

A mis padres y en especial a mi madre Berta Turizo quienes siempre nos ha inculcado mí y a mis hermanos la importancia del estudio.

A mi esposa Margaret Vila, quien me ha apoyado en este proyecto, ella ha sido soporte para cumplir mis metas, y me ha animado a ser un ejemplo para nuestros hijos Diego y Santiago; en general, agradecido con mi familia.

También a mis grandes amigos José de la Peña y Arnold de la Peña con quienes comenzamos este reto académico por medio de una beca ERASMUS para realizar el Master en la UGR y que posteriormente, gracias al ánimo de mis tutores, me ha llevado a seguir con este proyecto de vida.

A CEMIX-UGR y Banco Santander por financiar parte de este trabajo y a Ángel Blázquez Diéguez por su mentoría.

ÍNDICE

RESUMEN

Resumen

El trabajo que a continuación se presenta bajo la modalidad de tesis, en el marco de las normas actuales que regulan los requisitos para su presentación, se ha realizado por medio de la agrupación de tres artículos científicos que cumplen con los debidos criterios de indexación y que a su vez avalan el presente trabajo. Estos artículos se encuentran indexados en Journal Citation Report y se presentan cronológicamente de la siguiente manera:

- Chica-Olmo, Jorge; Cano-Guervós Rafael; Tamaris-Turizo Ivan. (2019). Determination of buffer zone for negative externalities: Effect on housing prices. *The Geographical Journal, 185*(2), 222-236.

- Chica-Olmo, J., Cano-Guervós, R., Moschovaki, M., Tamaris-Turizo, I. (2021). What is the effect of location on rental housing prices in Athens? *Revista Brasileira de Gestão de Negócios, 23*(3), 439-453.

- Tamaris-Turizo, I., Chica-Olmo, J., Cano-Guervós, R. (2022). The use of geostatistics to estimate missing data in a spatial econometric model of housing prices. *Revista Ingeniería de Construcción. 37*(3), 404-415.

El objeto principal de esta tesis es realizar un análisis geoespacial tanto del precio de venta de la vivienda de segunda mano, como su precio para arrendamiento en el sector urbano. En los mencionados trabajos se sustenta estadísticamente que la ubicación del inmueble es un determinante fundamental a la hora de estimar su precio. Para ello, se parte de la base metodológica relacionada con los precios hedónicos, para luego incorporar herramientas de la

Geoestadística y la Econometría espacial, con el fin de brindar mayor sustento y solidez a los resultados encontrados.

Esta investigación pretende realizar un estudio para determinar los factores asociados al precio de la vivienda, teniendo en cuenta sus características constructivas y localizativas, las cuales a través de diferentes métodos estadísticos evidenciaron la relevancia de la ubicación espacial de los inmuebles. Para ello, se utilizaron instrumentos de Sistemas de Información Geográfica (SIG), junto con las herramientas citadas anteriormente. Así, se destaca la utilización del método de Krigeaje para realizar interpolación espacial, y de esta forma resolver el problema de los datos faltantes, así como para obtener mapas de isovalores.

Para analizar el comportamiento de los precios de la vivienda para venta o arrendamiento, la literatura ampliamente se apoya en la teoría de los precios hedónicos, en la que el precio se determina en función de los atributos físicos del inmueble (Rosen, 1974). Clásicamente, esto se lleva a cabo mediante la aplicación del método de Mínimos Cuadrados Ordinarios (MCO), bajo el supuesto de la independencia de las observaciones, es decir, que no hay autocorrelación espacial.

Dentro del ámbito de la valoración inmobiliaria, las variables espaciales (relacionadas con la ubicación) entran a jugar un papel importante en el análisis de estos precios; por ende, una de las implicaciones que presentan estos modelos, es que existe dependencia o autocorrelación espacial entre las observaciones. Gracias al desarrollo de los métodos como la Econometría Espacial, la Geoestadística y los SIG en diversos campos, como la valoración de inmuebles, se ha demostrado a través de diversos estudios empíricos en el ámbito inmobiliario, las ventajas que proporciona el uso de estas metodologías, que permiten tener en cuenta dicha dependencia espacial.

Esta dependencia se respalda desde el punto de vista teórico con lo que se conoce como la Ley de Tobler (1970), la cual establece que "todo está relacionado con todo, pero las cosas cercanas están más relacionadas que las lejanas". En nuestro caso esto implica que existe una dependencia espacial en los precios, esto es, una dependencia respecto de la ubicación del inmueble. Esto implica que la relación de precios entre las viviendas disminuye a medida que aumenta la distancia entre ellas.

En el primer trabajo, publicado en junio de 2019 en la revista The Geographical Journal, titulado "Determination of buffer zone for negative externalities: Effect on housing prices", se desarrolló un modelo de precios hedónicos que consideró distintas variables constructivas y la ubicación de las viviendas. Así mismo, se tuvo en cuenta la proximidad con la línea ferroviaria de un tren de carga, con el fin de estimar una distancia de afectación causada por las externalidades negativas que genera el paso de éste. Para determinar dicha distancia se utilizaron métodos de Econometría Espacial y SIG. En este trabajo se encontró que las viviendas ubicadas en una zona alrededor de 320 metros a lo largo de la línea del tren presentan precios significativamente más bajos que el resto de las viviendas.

El segundo trabajo, publicado en 2021 en la Revista Brasilera de Gestión de Negocios (RBGN), titulado "What is the effect of location on rental housing prices in Athens?" busca determinar la relación que existe entre el precio de alquiler, los factores constructivos y la ubicación del inmueble. En el trabajo se utiliza el método clásico de Minimos Cuadrados Ordinarios (MCO), así como el método de Regression Kriging (RK), que es en una combinación del modelo de regresión y la Geoestadística. En el desarrollo de este estudio, se lograron obtener resultados estadísticamente significativos que respaldan las conclusiones alcanzadas. Estos resultados se basan en un análisis de los precios de arrendamiento en relación con la ubicación de los inmuebles. Para visualizar y comprender mejor la relación entre los precios de arrendamiento y la ubicación de los bienes, se utilizó el mapa de isovalores. A través de este mapa, se pudo observar de manera gráfica la existencia de una fuerte relación entre los precios de arrendamiento y la ubicación de los inmuebles.

El tercer trabajo, publicado en 2022 en la revista "Ingeniería de Construcción" de la Pontificia Universidad Católica de Chile, titulado "The use of geostatistics to estimate missing data in a spatial econometric model of housing prices", consistió en resolver el problema de datos faltantes para la variable antigüedad, debido a que para algunas de las viviendas no estaba disponible esta información. Para resolver este problema se utilizó interpolación espacial (Krigeaje). Esto permitió incrementar la muestra y mejorar la bondad del modelo. Además, se utilizó el método de regresión ortogonal para resolver el problema de multicolinealidad grave. Finalmente se estimaron modelos espaciales para determinar el precio de la vivienda.

La estructura desarrollada en esta tesis consta de los siguientes capítulos: se inicia con una introducción general, relacionada con los artículos publicados. A continuación, se realiza la correspondiente justificación teórica; posteriormente, la metodología y objetivos planteados, los tres artículos publicados y finalmente las conclusiones generales, con sus correspondientes limitaciones y líneas futuras de investigación.

Bajo la modalidad de tesis por compendio, se presentan a continuación los tres artículos que soportan la presente tesis doctoral.

Tabla 1. Indicadores de calidad de los artículos publicados

Autores: Chica-Olmo Jorge, Cano-Guervos Rafael, Tamaris-Turizo Ivan.

Título: Determination of buffer zone for negative externalities: Effect on housing prices

Referencia (DOI): 10.1111/geoj.12289

Revista: The Geographical Journal.

Editorial: Wiley Blackwell

Factor de Impacto: 5.2 (4-Year)

Posición:

Cite score rank 2020 (Clarivate Analytics)

30/85 GEOGRAPHY (Q1)

Revisión por al menos 2 expertos independientes

Autores: Chica-Olmo, Jorge; Cano-Guervos, Rafael; Moschovaki, María; Tamaris-Turizo, Ivan.

Título: What is the effect of location on rental housing prices in Athens?

Referencia (DOI): 10.7819/rbgn.v23i3.4114

Revista: Revista Brasileira de Gestão de Negócios (RBGN)

Editorial: Fundacao Escola de Comercio Alvares Penteado

Factor de Impacto: 1.3 (2017-2020)

Posición:

Journal Citation Indicator (JCI-2020)

143/1.199 BUSINESS AND INTERNATIONAL MANAGEMENT (Q3)

Revisión por al menos 2 expertos independientes

Autores: Tamaris-Turizo, Ivan; Chica-Olmo, Jorge; Cano-Guervos, Rafael

Título: The use of geostatistics to estimate missing data in a spatial econometric model of housing prices

Referencia (DOI): 10.7764/RIC.00044.21

Revista: Ingeniería de Construcción

Editorial: Pontificia Universidad Católica de Chile

Factor de Impacto: 0,227

Posición:

Journal Citation Indicator (JCI-2020)

138/185 BUILDING AND CONSTRUCTION (Q3)

Revisión por al menos 2 expertos independientes

Listado de figuras y tablas

CAPÍTULO 1. Introducción general

1.1 Justificación teórica

La vivienda es un objeto de estudio relevante para las familias, empresas y organismos oficiales, entre otros agentes. Tener una vivienda es fundamental y de gran importancia como patrimonio familiar, puesto que brinda bienestar, protección entre otras cualidades. Para lograr acceder ella, las familias requieren destinar una proporción considerable de sus ingresos. En muchos casos conseguir a una casa propia implica todo un reto financiero, y en la mayoría de las veces se debe acudir a créditos de largo plazo para poder conseguirla (Igan, & Loungani, 2012). No obstante, las restricciones de liquidez pueden ser un obstáculo para que las familias accedan a este tipo de crédito (Tejada & Vidaurre, 2007).

Así mismo, la adquisición de viviendas no sólo es relevante para las familias, sino que también representa una fuente de financiación para el sector público, al estar sujeto a tributación; también es de interés para el sector de seguros y el sector bancario.

Dada la importancia de este tipo de activos, se considera relevante realizar estudios para la estimación de su precio. De manera que la información aportada en este campo es una base importante para que potenciales compradores o inversionistas tomen sus decisiones al momento de adquirir o alquilar una propiedad de una manera más acertada. El enfoque principal de este estudio se centra en llevar a cabo estimaciones del precio de la vivienda de segunda mano, tanto para transacciones de venta como de alquiler.

Esta evaluación también se considera relevante teniendo en cuenta el interés de individuos o inversionistas que buscan obtener beneficios por medio del arrendamiento de propiedades, manifestando su inclinación por esta clase de activos. Como consecuencia, la demanda de estos inmuebles termina favoreciendo el sector de la construcción, que tradicionalmente ha sido un elemento clave en la economía y que a su vez genera un impacto importante para el crecimiento económico (Ofori, 2012). El sector de la construcción genera una amplia demanda de bienes y servicios, como materiales de ferretería, servicios profesionales y no profesionales en diversas áreas, etc., además de otros tipos de gastos indirectos que se irrigan en la sociedad como resultado del desarrollo de un proyecto inmobiliario. Este sector también absorbe una cantidad importante de mano de obra, que ayuda a disminuir la tasa de desempleo. Por todo ello, el sector de la construcción es considerado un dinamizador fundamental para la economía de un país (Palomino Silva et al., 2017).

Con base en lo anterior se considera pertinente realizar estudios para la estimación de precios de venta de vivienda de segunda mano o para arrendamiento. Gran parte de la literatura se apoya en la aplicación de los modelos hedónicos, que resulta del precio esperado según las características constructivas que ofrece la vivienda (Goodman, 1978). De esta manera, se cuantifican los precios implícitos relacionados con las características del inmueble (Rosen, 1974). Según la literatura este tipo de modelos hedónicos se considera la base tanto para la estimación de precios en casos de venta de viviendas de segunda mano, como en la estimación de su precio de arrendamiento.

Existen circunstancias particulares en las que la cercanía a puntos de interés o lugares específicos ejerce influencia en el precio del inmueble, pudiendo generar efectos tanto positivos como negativos, dependiendo de las ventajas o desventajas que estos brinden. Estos efectos producidos son denominados externalidades, donde según Nicholson (1998) en su definición seminal se presentan cuando “las actividades de un agente afectan a la actividad de otro agente de manera que no se reflejan en las transacciones de mercado”. Lo anterior se refiere a la existencia de impactos indirectos o externos que no se reflejan de manera inmediata en las transacciones comerciales.

A continuación, se mencionarán algunos aspectos descritos en la literatura, mencionando en primer lugar aquellos que impactan positivamente sobre el inmueble. Así, el precio de la vivienda y de su arrendamiento para viviendas de tipo residencial considera el suministro adecuado de servicios públicos y/o amenidades ambientales que se encuentren en su entorno (Vásquez, W, & Beaudin, L, 2020). Siendo estas últimas de gran importancia porque tienen un impacto significativo en el valor de la propiedad, lo cual se refleja en un aumento de su precio, debido a las bondades que ofrecen las amenidades ambientales contiguas a la vivienda. El resultado es igualmente válido cuando se trata de viviendas utilizadas para fines vacacionales, pues el acceso a las diversas amenidades de recreación impacta sobre el precio de arrendamiento justificando una mayor alza del precio de renta (Nelson, 2010).

Sadayuki (2018) encontró una relación geográfica del precio de arrendamiento de las viviendas con las amenidades situadas en lugares circundantes, tal como paradas de transporte público (subterráneos y trenes). Así, las viviendas más próximas a estos sitios contemplan precios de arrendamientos más altos, y en la medida que incrementa la distancia con relación a estos puntos, el precio de arrendamiento decrece.

También resulta interesante analizar la relación entre los establecimientos educativos (de primaria y bachillerato) de alta calidad con las viviendas de tipo multifamiliar. Según Gabe et al., (2022), Kuroda (2018), los arrendatarios están dispuestos a pagar una renta mayor en viviendas cercanas a colegios con mayor calidad académica, lo que podría revelar el interés de las familias por demandar estos colegios. En este sentido, determinó que la calidad de la escuela tiene un efecto positivo sobre las viviendas donde vive cierto grupo familiar, no siendo este efecto positivo cuando se trata de alquiler por parte de personas solteras.

Además de estos factores localizativos, una fuerte demanda de vivienda con fines turísticos también influye en el precio, debido a que esta situación termina presionando al alza el precio de arrendamiento (Blanco-Romero et al., 2018). También Nelson (2010) considera los efectos positivos que se generan por la cercanía de amenidades de recreación.

En cambio, las externalidades negativas inciden en el precio tanto de la vivienda como del alquiler, provocando una disminución significativa. Zambrano-Monserrate & Ruano, (2021) encontraron que en el mercado de alquiler de viviendas aquellas que se encontraban más cercanas a cuerpos de agua contaminados generan un efecto negativo en su precio. Esto implica que cuando se presenta esta situación, el precio de alquiler aumenta en la medida que lo hace la distancia con respecto a estos estuarios contaminados.

Otra situación en la que se presentan externalidades negativas, según Agudelo et al., (2015), es la cercanía a estadios deportivo de futbol (se cita el caso de Atanacio Girardot, en la ciudad de Medellín - Colombia), debido a los problemas de orden público que algunas veces se presentan entre los aficionados de los equipos de futbol.

Adicionalmente, la literatura científica ha evidenciado la alta correlación que existe ente el precio y la localización, considerando esta premisa esencial en la mayoría de los trabajos aplicados en esta línea. La presencia de autocorrelación espacial es la principal hipótesis considerada al estudiar el precio de alquiler y precio de la vivienda usada desde un punto espacial. Esta autocorrelación espacial se debe a que los propietarios de viviendas próximas en el espacio valoran con precios parecidos sus viviendas.

Para la estimación del precio y del alquiler de la vivienda se pueden adoptar diversos enfoques metodológicos de investigación, tal como la aplicación de modelos de regresión múltiple y/o análisis espacial. El Modelo de Regresión Lineal Múltiple (MRLM) ha sido uno de los métodos más utilizados en la literatura para la determinación de los factores que explican dicho precio. Los modelos hedónicos permiten determinar el precio asociado a cada uno de estos factores (Rosen, 1974), por lo tanto, el MRLM es un punto de partida adecuado para alcanzar este objetivo.

Con el propósito de determinar un modelo que explique la relación entre el precio de la vivienda y las variables que influyen en él, es fundamental considerar como factor explicativo la ubicación o localización de la misma. Por ello es conveniente utilizar métodos que permitan tener en cuenta este factor, como los métodos geoestadísticos. En particular el método de Krigeaje o kriging, que permite considerar la presencia de autocorrelación espacial en la variable analizada o en las perturbaciones del modelo, proporcionando predictores más precisos (Matheron, 1970; Basu & Thibodeau, 1998). También se puede considerar la utilización de cokriging en aquellos casos en que no sólo la variable de interés (precio) presenta autocorrelación espacial, sino que también la presentan las variables auxiliares (Chica-Olmo, 2007).

Además, la utilización de los SIG son herramientas que ayudan a representar y analizar los datos georreferenciados, como es el caso del precio de la vivienda. La combinación de estas herramientas junto con los métodos geoestadísticos permite analizar desde un punto de vista espacial los precios de las viviendas (Chica-Olmo et al., 2013).

En diversos trabajos se evidencia que tanto el precio de la vivienda como el precio de alquiler se encuentra determinado por sus características constructivas y localizativas. Entre las características constructivas se pueden citar las dependencias de las que se compone una vivienda, las habitaciones, baños, garaje, etc. Mientras que las localizativas hacen referencia a la accesibilidad a distintos lugares de interés, como centros de educación, puntos de transporte, comercio, entre otros.

Diversas variables juegan un rol determinante en el precio de la vivienda y estas pueden incrementar o reducir el precio. La antigüedad, por ejemplo, juega un papel importante para este tipo de estimación, de tal manera que se ha encontrado que las viviendas más nuevas tendrán mayor precio, y en la medida que sean más antiguas su precio será menor. De otra manera, determinados factores localizativos tienen efectos positivos sobre el precio, como la proximidad de la vivienda a diferentes amenidades (Chen & Jim, 2010), a las estaciones de bus (Yang et al., 2019), a centros educativos (Wen et al., 2014).

CAPÍTULO 2. Objetivos, hipótesis y metodología

2.1 Objetivos e hipótesis

A partir de una evaluación sobre la literatura expuesta en el capítulo 1, se han identificado necesidades de investigación que requieren la formulación de una serie de objetivos, con el propósito de impulsar el avance del conocimiento en relación con los determinantes del precio de la vivienda y la estimación de su precio de arrendamiento.

En particular, el objetivo principal de la siguiente tesis doctoral es realizar una modelización hedónica del precio de la vivienda de segunda mano o su precio de alquiler, haciendo especial énfasis en los efectos de la localización sobre dicho precio. Más concretamente, un objetivo será cuantificar la influencia de la localización de las viviendas en el sector urbano de Santa Marta (Colombia) y en Atenas (Grecia), y para ello se requiere analizar la autocorrelación espacial de los precios.

Objetivos específicos:

- ✓ Objetivo 1: Determinar el impacto de la ubicación sobre el precio de arrendamiento de la vivienda

Con relación a este objetivo, el estudio realizado sobre el precio de alquiler de la vivienda en la ciudad de Atenas (Chica-Olmo, J., Cano-Guervós, R., Moschovaki, M.-D., & Tamaris-Turizo, I. (2021) se apoya en la aplicación de modelos de precios hedónicos, los cuales establecen una relación entre el precio de alquiler y las características de la vivienda. Sin embargo, esto no es suficiente, porque bajo esta estimación no se considera la ubicación del inmueble, por lo que se han implementados métodos geoestadísticos, como el RK (Regresión Kriging) (Chica-Olmo, 1995; Dubin, 1992; Hengl et al., 2007), para robustecer la estimación del modelo. De esta manera, se han obtenido conclusiones que respaldan la relación entre el precio de arrendamiento y la ubicación del inmueble, proporcionando una mayor comprensión de los factores que influyen en los precios de alquiler de vivienda.

- ✓ Objetivo 2: Estimar datos faltantes de la variable antigüedad

Este objetivo, planteado en el trabajo de Tamaris-Turizo, I., Chica-Olmo, J., Cano-Guervós, R. (2022), busca resolver el problema de la falta de información en algunos puntos no muestreados, en particular, de la variable "antigüedad" de las viviendas. La finalidad es encontrar una solución que permita completar los datos faltantes sin perder información y, al mismo tiempo, mejorar el tamaño de la muestra, lo cual contribuirá a aumentar la significancia del modelo estimado. De esta manera, se busca obtener estimaciones confiables de la variable "antigüedad" en los puntos de muestra donde no se dispone de estos valores, asegurando una mayor integridad y utilidad de los datos recolectados, lo que resultará en una mejora de la significancia del modelo.

- ✓ Objetivo 3. Obtener el mapa de isovalores para representar gráficamente la distribución espacial del precio de la vivienda

Este objetivo desarrollado en el trabajo Chica-Olmo, J., Cano-Guervós, R., & Tamaris-Turizo, I. (2019), busca profundizar en el análisis de la distribución espacial sobre los precios de la vivienda a lo largo del territorio, incorporando en ella la modelización de la tendencia en el precio que se observa desde la costa al resto de las ubicaciones de la ciudad. Para lograrlo, se utilizarán los SIG, los cuales ofrecen una herramienta muy ilustrativa y de gran utilidad en este contexto. En particular, se empleará el mapa de isovalores, obtenido mediante el método de Kriging. Dicho mapa es una representación gráfica que muestra de manera efectiva la estructura espacial de los precios de la vivienda. Este mapa permite visualizar cómo los precios varían en función de la ubicación geográfica. Al analizar la información de la costa hacia la periferia, se espera obtener información valiosa sobre la evolución de los precios a medida que nos alejamos del área costera y nos adentramos en las áreas periféricas de la ciudad. Este análisis permitirá identificar posibles patrones o gradientes de precios, así como comprender la dinámica que influye en la formación de valores inmobiliarios en diferentes zonas.

- ✓ Objetivo 4: Determinar el radio de influencia y cuantificar las externalidades negativas generadas por el paso de un tren de carga.

Para lograr este objetivo, en el trabajo Chica-Olmo, J., Cano-Guervós, R., & Tamaris-Turizo, I. (2019) se han utilizado distintos modelos de precios hedónicos, técnicas de econometría espacial y Geoestadística, que permiten examinar la relación entre las características de la vivienda y el impacto del tren de carga. Además, se han empleado los SIG con el fin de delimitar la zona de afectación que generan externalidades negativas causadas por el tren de carga.

Por lo tanto, los objetivos anteriores permiten considerar que es pertinente realizar estudios en los que se incorporen variables localizativas en la modelización del precio de la vivienda.

Con relación a los objetivos mencionados, se plantean las siguientes hipótesis:

H1: Se plantea si el precio de venta de la vivienda se encuentra distribuido aleatoriamente en el espacio o, por el contrario, se encuentra autocorrelacionado espacialmente.

H2: ¿La incorporación de información relativa a datos ausentes mejora la bondad del modelo y aumentan las variables significativas en el mismo?

H3: ¿Los precios de las viviendas en la ciudad de Santa Marta presentan una tendencia espacial creciente desde la periferia a la costa?

H4: ¿Los precios de las viviendas están relacionados con algunos lugares de interés?

H5: ¿La proximidad de las viviendas a la vía del tren de mercancías tiene un efecto negativo sobre el precio de las mismas?

H6: ¿Cuál es el radio de influencia de las vías del tren de carga dentro del cual se ven afectados los precios de las viviendas?

2.2 Metodología

Con el fin de lograr los objetivos planteados en esta investigación, fue necesario utilizar un enfoque metodológico que combina técnicas de la Econometría clásica, Geoestadística y Econometría espacial, adaptando cada una de ellas de acuerdo con los requerimientos de cada objetivo establecido.

La Econometría clásica proporciona una base metodológica que ha permitido elaborar modelos de regresión para explicar y estimar el precio de la vivienda a partir de las características constructivas y localizativas de la misma.

Por otro lado, la Geoestadística ha desempeñado un papel fundamental en cuanto a la cuantificación y análisis de las relaciones y la variabilidad espaciales de los datos. Por medio de esta técnica se ha podido comprender la estructura espacial, identificar patrones y tendencias y realizar interpolaciones espaciales para la estimación de la variable antigüedad en lugares no muestreados y para la construcción de mapas de isovalores del precio y del alquiler de la vivienda.

Y a manera de robustecer esta investigación, la incorporación de la Econometría espacial ha proporcionado herramientas y métodos específicos para datos espaciales, teniendo en cuenta la dependencia espacial y la influencia de la ubicación geográfica en las relaciones estudiadas. Esta técnica permitió modelizar y analizar la interacción espacial, evaluar la propagación de efectos y realizar inferencias espaciales más precisas.

En el sector de los bienes raíces gran parte de la literatura relacionada con la estimación de precio de viviendas se sustenta en la modelización de precios hedónicos, siendo el método de regresión tradicionalmente usado para obtener modelos econométricos (Chica-Olmo, 2007). Estos modelos se aplican de manera similar cuando se trata de determinar el precio de alquiler de las viviendas con el fin de determinar los factores que influyen en ese precio (Efthymiou & Antoniou, 2013), de tal manera que los precios dependerán de las características estructurales del inmueble. En concreto, la Econometría clásica mediante el método de MCO (Mínimos Cuadrados Ordinarios) se contempla como punto de partida para tratar de establecer una relación del precio en función de las variables constructivas, tal como el número de habitaciones, baños y otras características similares, así como en función de variables localizativas, tales como la distancia al centro de la ciudad y otros lugares de interés. De este modo se pretende determinar el peso que tienen estos factores que influyen sobre el precio de venta o de alquiler.

Por lo tanto, la metodología planteada anteriormente se aplica en los trabajos recogidos en los capítulos 3, 4 y 5. La expresión del modelo lineal general se expresa de la siguiente manera:

$$y = \alpha + \beta X + \varepsilon$$

Donde:

y: variable explicada (precio de la vivienda o precio de alquiler)

α: intercepto o constante

β: coeficientes de las variables explicativas

X: matriz de variables explicativas

ε: perturbaciones

No obstante, este modelo no considera las interacciones espaciales que podrían existir entre las observaciones (unidades residenciales), es decir, que no tiene en cuenta la correlación o dependencia espacial que podría darse en el precio de las viviendas. No obstante, esta modelización sí permite incorporar variables localizativas (como las distancias a lugares de interés) para cuantificar su efecto sobre el precio de venta de la vivienda o su precio de alquiler.

El trabajo con datos georreferenciados puede dar lugar a la presencia de autocorrelación espacial (Anselin, 1998), es decir, que los valores (precios de venta o arrendamiento) sean similares entre los vecinos o zonas vecinas. En consecuencia, se requiere la aplicación de técnicas de la Geoestadística y de Econometría espacial, para tener en cuenta la existencia de autocorrelación espacial entre las observaciones. La Geoestadística ofrece una alternativa para determinar estas relaciones, mediante la Regresión Kriging (RK), que consiste en una combinación de regresión con krigeaje (Chica-Olmo, 1995; Dubin, 1992; Hengl et al., 2007). Tal modelo se expresa de la siguiente manera:

$$\hat{y}(s_o) = x(s_o)'\hat{\beta} + \lambda' e(s_i)$$

Donde:

$\hat{y}(s_0)$: precio estimado de la vivienda en la localización s_0

$x(s_o)$: variables explicativas

$\hat{\beta}$: Coeficientes estimados de las variables explicativas

λ: es el vector de pesos Kriging

$e(s_i)$: es el vector de residuos

Los pesos Kriging se obtienen teniendo en cuenta la estructura de autocorrelación espacial de los residuos, mediante el variograma (véase la ecuación en los capítulos 3 y 5).

El modelo presentado anteriormente permite obtener los mapas de isovalores, el cual se incorpora como herramienta en los SIG. Estos mapas muestran la distribución espacial de los precios de las viviendas, y permiten identificar zonas donde las viviendas presentan precios similares de acuerdo con su ubicación. Al mismo tiempo esta técnica contribuyó a revelar la distribución espacial de la antigüedad de las viviendas a lo largo de la ciudad de Santa Marta, y de esta manera se pudo apreciar las zonas donde se encuentran las viviendas con mayor y menor antigüedad.

Por otro lado, la Econometría espacial también ha desarrollado modelos en los que se considera la autocorrelación o dependencia espacial, de tal manera que se determina una matriz de pesos, llamada W, en la cual se puede establecer el tipo de vecindad mediante diversos criterios.

De acuerdo con lo anteriormente dicho, hay muchos estudios que contemplan la aplicación de SIG y Econometría espacial, en los que se tiene en cuenta la localización para explicar el precio de la vivienda. Sin embargo, la aplicación de métodos Geoestadísticos para la estimación de modelos hedónicos aplicados al precio de la vivienda no es tan común (Chica-Olmo, et al., 2019). Con relación a la Econometría espacial, se ha considerado la utilización de los modelos Spatial Autorregresive Model (SAR) y el Spatial Error Model (SEM). En el primero se asocia un proceso regresivo con la variable explicada, mientras que el segundo se tiene en cuenta el retardo de las perturbaciones:

$$SAR: y = \rho Wy + X\beta + \varepsilon$$

$$SEM: y = X\beta + u$$

$$u = \lambda Wu + \varepsilon$$

Donde:

y: vector (n) de la variable dependiente

X: matrix (nxk) con k variables explicativas

β: vector de parámetros

W: matrix de pesos espaciales,

Wy: vector y con retardo espacial

u: vector de perturbaciones

Wu: vector de perturbaciones con retardo espacial

λ: parámetro asociado a las perturbaciones retardadas

ε: vector perturbaciones distribuido normalmente.

Una vez seleccionado el modelo más adecuado, su dependencia se determina por medio de algunos estadísticos. Uno de los más utilizado en la literatura es el test I de Moran, cuya significación estadística permite detectar la presencia de autocorrelación espacial.

CAPÍTULO 3. What is the effect of location on rental housing prices in Athens?

Jorge Chica-Olmo[1]
Rafael Cano-Guervos[1]
María-Despoina Moschovaki[2]
Ivan Tamaris-Turizo[3]

[1]*University of Granada, Department of Quantitative Methods for Economics and Business, Granada, Spain.*
[2]*Economic analyst, NGA Human Resources, Granada, Spain.*
[3]*International School of Administration and Marketing, Sergio Arboleda University, Santa Marta, Colombia*.

Abstract

Purpose – The aim of this study is to quantify the effect of location on rental housing prices in the city of Athens.

Theoretical framework – The right to adequate housing is a fundamental human right defended by democratic societies. Therefore, it is of interest to examine housing tenures for both owned and rented accommodation.

Design/methodology/approach – Geostatistical methods (regression-kriging) were used to obtain the results, which are represented on an isovalue map of rental housing prices displaying the minor and major effects of location by zone.

Findings – This study highlights the impact of location on rental housing prices by showing how the rent of a standard dwelling in the city of Athens varies depending on its location.

Research Practical & Social implications – The main social implications of this work is it helps investors determine where to direct investments and it assists public authorities in deciding where to focus urban management policies, in order to control the undesirable effects of an excessive rise in rents caused by tourism.

Originality/value – The main originality of this paper lies in its isovalue map of rental housing prices for standard dwellings, which can also be interpreted as a locational isovalue map.

Keywords – rental housing prices, geostatistics, isovalue map, effects of location, Athens.

3.1 Introduction

The right to adequate housing is a fundamental human right defended by democratic societies, and has always formed part of the class struggle (Danieli et al., 2018). Rex (1968) considered three types of housing tenure: multiple-house owners, homeowners, and renters. These types of tenure gave rise to the concept of housing class developed by Weber, who argued that the class struggle is not only a struggle for the means of production, but also for access to housing (Chen et al., 2018). In addition, “space” can be considered as being where the class struggle takes place given that the upper social classes tend to occupy the best locations, displacing the lower classes towards marginal areas (Goodall, 2013). In fact, according to Tse (2002), as housing prices increase, families tend to move to places with poorer accessibility where prices are generally lower. This spatial segregation gives rise to a concentration of the upper social classes in areas with dwellings of superior quality and location assets, which are therefore more highly priced (Ozanne & Thibodeau, 1983).

The problem of housing tenures in large cities is currently of great concern to many citizens, particularly young people seeking to become independent. The price of rental properties for residential use is being influenced by the strong demand for tourist rental accommodation, which has given rise to a potential “bubble” in rental prices in some European cities (Blanco-Romero et al., 2018) and what is known as the “turistification” phenomenon.

The rental housing market is an economic activity that consists of renting property acquired through purchase, donation, inheritance, etc. from one individual to another. Investors in this market make long-term financial investments by purchasing properties with the aim of renting them for residential use (buy-to-rent). There are generally two types of rental agreements for dwellings targeted at residential use depending on the period of time stipulated in the agreement: short-term or long-term agreements. Given that investment in housing is typically long term, homeowners, whether they are large or small investors, generally prefer long-term rentals (Wulff & Maher, 1998). As highlighted by several authors, long-term rentals have traditionally been the object of research given the interest of policymakers and politicians. However, as indicated above, there is currently a boom in another type of apartment rental: tourism rentals, which are in increasing supply due to the potential profits to be gained (Nasreen & Ruming, 2019).

From the perspective of rental housing management, there are three types of managers: public companies, private companies, and private owners. As Lee (2007) indicated, private owners are generally more common in Western capitalist societies, leading to an asymmetric pattern in public and private housing investment.

Nonetheless, regardless of the type of management or agreement (short or long term), rental housing is one of the main factors that determines housing prices (Gallin, 2008). In fact, one of the most commonly used methods for valuing housing is capitalization, which determines the price of a dwelling based on its rental price (Pagourtzi et al., 2003). However, this may be subject to an endogenous effect, since rental price is also determined by the price of a dwelling. In some countries, for example, it is a common practice among banks and real estate agents to set a rental price equivalent to a percentage of the selling price. In other words, rent is the dividend in the real estate stock market (Leamer, 2002). Furthermore, the price-to-rent ratio is an indicator of the real estate market, which could have the potential for predicting house prices and determining the factors that explain price evolution over time (Gallin, 2008; Taipalus, 2006).

It may also be of interest to analyze the variation in rental prices in space by taking into account the location of rental dwellings. For this purpose, it is convenient to perform a spatial analysis of rental housing prices using different tools such as geographic information systems (GIS), spatial econometrics, and geostatistics. These tools can assist in determining the effect of location on rental housing prices, which could help investors and public bodies to improve their decision-making processes.

Given the geographical scope of this study, it is important to highlight that, until the 2008 crisis, the building industry was among the most influential sectors in the Greek economy. It was boosted by state development, the government's fiscal and social policies, and the propensity of Greek citizens to own real estate (Papadimitriou et al., 2013). Although housing prices in Greece have fallen by 38.5% since the crisis (Ramos, 2015), the cost of rental accommodation now represents 40% of disposable income, while the EU average is 28% (Eurostat, 2018).

As indicated above, rental housing is a subject of social interest. Although this topic has been widely examined using a temporal approach, we believe that the scarcity of studies that use a spatial approach justifies this research. The main objective of this study is to determine the impact of a property's location on its rental price. To this end, we studied a sample of rental housing prices in the city of Athens in the second quarter of 2014.

In what follows, we first present a review of the literature on the importance of location to rental housing prices and describe the methods used. We then discuss the data and results and, finally, draw some conclusions.

3.2 Literature review

As indicated above, housing can be considered as an investment product whose profitability can be measured by its ability to generate a higher or lower income. Consequently, rental housing prices and their explanatory factors are topics of interest. Hedonic modeling can be used to quantify the implicit prices of factors that determine the overall price of a good (Lancaster, 1966; Rosen, 1974). Hedonic models have been used in a large number of studies primarily to explain housing prices. However, the hedonic housing regression model has received some criticism (Chau & Chin, 2003), mostly due to the limitations of the classical estimation method used: ordinary least squares (OLS). The use of alternative methods has made it possible to overcome some of the limitations of OLS, such as the weaker explanatory and predictive capacity of OLS in the presence of spatial dependence (Anselin, 1988). Different solutions have been proposed to improve this explanatory and predictive capacity using a spatial model approach.

Within this type of modeling, spatial econometrics have provided the most widely used solutions through the specification of spatial autoregressive (SAR) models, spatial error models (SEM), and geographically weighted regression (GWR) models (see Dubin et al., 1999; Krause & Bitter, 2012; Osland, 2010; Wen et al., 2017). Geostatistical methods have also been successfully used (see Bourassa et al., 2007; Cellmer, 2014; Chica-Olmo, 1995; Dubin, 1992). These methods have the advantage of providing isovalue maps for the analyzed variable. As regards the modeling of rental housing prices, a few studies have used the classical methods (Hoch & Waddell, 1993; Nishi et al., 2019; Wheaton, 1977), quantile regression models (Cui et al., 2018), spatial econometric models (Efthymiou & Antoniou, 2013), and GWR (Iliopoulou & Stratakis, 2018). Studies using geostatistical methods are virtually non-existent, hence one of the main novelties of this study. The common denominator in these methodologies is that they consider the spatial component as a key element for the modeling of rental housing.

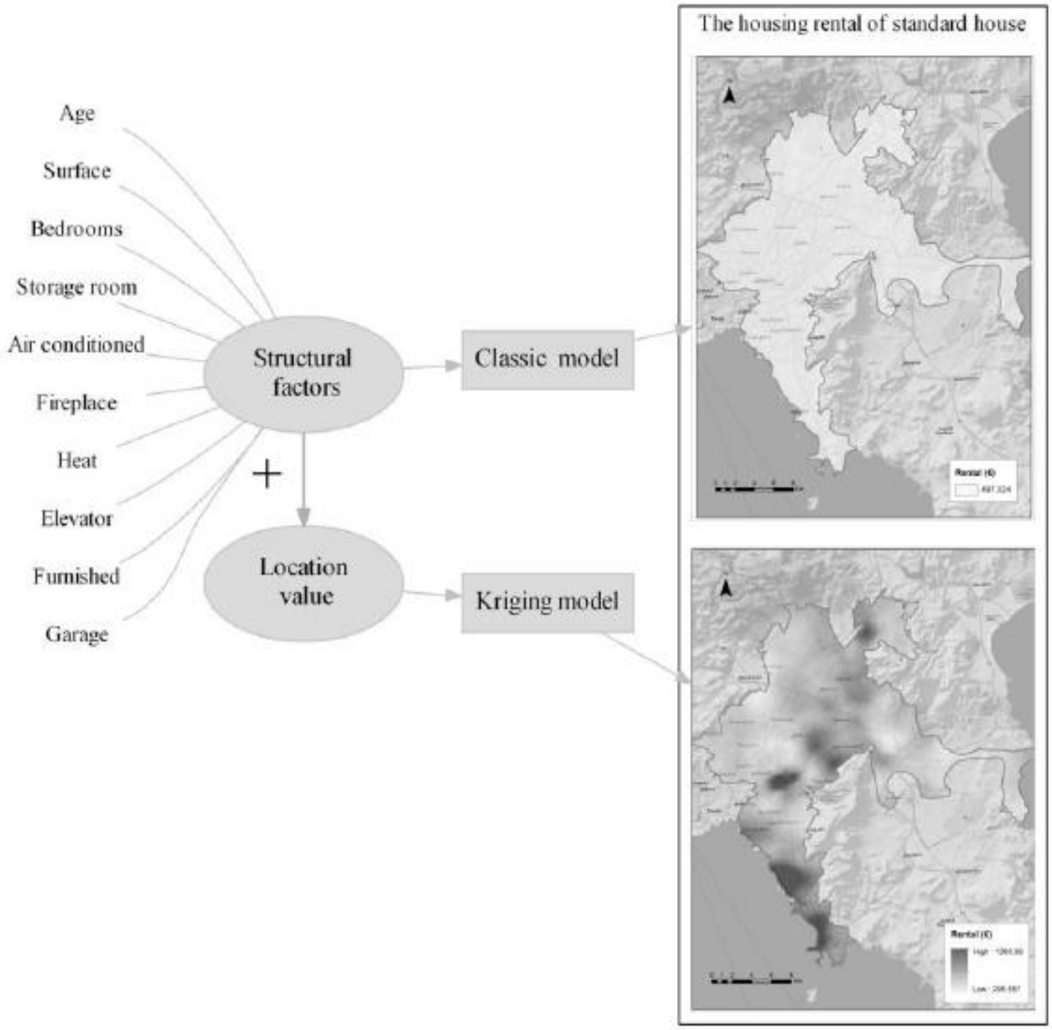
From a theoretical perspective, the spatial component must be considered when explaining rental housing prices. In the case of monocentric cities, the classic urban land bid rent theory establishes that real estate rents decrease as the distance to the city center increases (Von

Thünen, 1966). This theory also applies to polycentric cities (Harris & Ullman, 1945). These approaches take into account friction costs associated with accessibility, which is measured as the distance from a dwelling to the central business district or the cost incurred or time spent traveling between the two. Some authors argue that there is a relationship between utility and rent maximization, which would ensure that location and rent patterns are identical (Wheaton, 1977). In fact, according to Alonso (1964), "location" is a hedonic good whose value should be explained.

The factors that determine the price of housing have been traditionally classified into structure, accessibility, and neighborhood attributes (Can, 1992; Chica-Olmo, 1995; Dubin, 1998). Of these, the structural characteristics of dwellings are relatively easy to measure and are generally included in publicly-available data. These characteristics include floor area, number of rooms, age, and the general physical or structural characteristics of the dwelling.

Figure 1 shows the structural characteristics considered in this study. However, accessibility and neighborhood attributes are closely related to the location of each dwelling in a specific area. For example, transport, nearby services of different kinds, accessibility to one's workplace and different places of interest, and the quality of the neighborhood and environment are just some of the characteristics that give each residential unit a unique set of attributes depending on the location (Efthymiou & Antoniou, 2013; Tse, 2002). Location characteristics are not always easy to measure although they form part of the price, and as such the price of a dwelling will be affected by its location (Chica-Olmo et al., 2013). What these characteristics have in common is that they depend on location. As a result, some authors, such as Dubin (1992), have stated that the three most important factors that determine housing prices are location, location, and location. It is evident, then, that accessibility and neighborhood characteristics (i.e., location components) should affect rental housing prices, although, as said, they are not always easy to measure.

Figure 1. Factors that influence rental housing prices and models



For example, location attributes related to accessibility, such as proximity to the central business district, workplaces, main transport routes, public schools, and shopping centers, among others (Chen & Jim, 2010; Chica-Olmo et al., 2019; Efthymiou & Antoniou, 2013; Trojanek & Gluszak, 2018), require the unit of measurement to be determined (distance, time, or cost). At other times the difficulty lies in determining the area of influence pertaining to the previously mentioned factors or other factors, such as ethnicity, crime rate, and environmental factors, such as noise and air pollution (Beimer & Maennig, 2017; Fabusuyi, 2018; Le Boennec & Salladarré, 2017; Moye & Thomas, 2018; Swoboda et al., 2015). Location factors are more difficult to measure than structural characteristics (Basu & Thibodeau, 1998). In addition, when assessing the effects of location components, it is difficult to determine neighborhood boundaries and what measures to use to determine those boundaries (Diao et al., 2017; Diao et al., 2016). As a result, there are areas where boundaries are quite obvious, such as municipal boundaries, school district boundaries, and other boundaries based on geographical elements such as highways, roads, rivers, and parks (Fan et al., 2016). However, there are other areas where neighborhood boundaries are not as evident. Despite these difficulties, it is important to incorporate the impact of location when estimating hedonic rental housing prices (Basu & Thibodeau, 1998) given that location is considered to influence both the selling price and the rental price of a dwelling (Efthymiou & Antoniou, 2013).

The omission of location variables in the regression model leads to bias in the OLS. Spatial autocorrelation can be observed in the disturbances when location variables are omitted. Nishi et al. (2019) quantified the bias caused by the omission of factors that determine the rental price of dwellings in Tokyo, Japan. However, spatial prediction methods, such as geostatistical methods (kriging), can reduce bias by taking into account autocorrelation in disturbances (Dubin, 1988). Autocorrelation is caused by the first law of geography, according to which “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970, p. 236). In the case of housing, autocorrelation occurs for several reasons. Given that neighborhoods generally develop at the same time, properties have very similar structural characteristics, such as size or design. In turn, dwellings belonging to the same area share the same advantages and disadvantages of location (Basu & Thibodeau, 1998).

The effect of spatial contagion caused by the transfer of information between neighboring homeowners in a space regarding the predominant prices of the area must also be taken into account. All of these factors influence both housing and rental prices (Hoch & Waddell, 1993), sometimes positively and sometimes negatively. In short, the effect of location should be considered in a hedonic model.

If rental housing prices are estimated using a traditional hedonic regression model (OLS) that does not consider the effect of location on prices but only the structural characteristics of a standard dwelling, the estimates would be the same for the entire geographical space. In contrast, the kriging model accounts for the effect of location, and hence the rent for a standard dwelling will vary depending on its location (see Figure 1).

The main interest of hedonic models is that the implicit price of each of the characteristics that determine rental housing prices is quantified in the model coefficient. That is, the adjusted regression model estimates the influence of each of the housing characteristics on rental housing prices. This work proposes a method that can be widely applied, since it only requires information that is generally available on many real estate websites (rental housing prices, certain structural characteristics of the dwellings and their locations). The second reason for using this method is that it allows property values, which are sometimes overly subjective, to be estimated in an objective manner, especially with regard to locational factors. The third reason is that the method not only enables the rental price of a particular property to be calculated, but it can also be used for mass valuations of homes, thus making it an effective tool for real estate and construction companies, real estate investment funds, financial and mortgage entities, as well as public administrations for the purposes of taxation or urban management.

3.3 Methods and Data

3.3.1 Methods

Geostatistics is a science that facilitates the analysis of spatial data using spatial quantitative techniques. Geostatistical models are based on the estimation of a variances-covariances matrix of disturbances to subsequently estimate the model parameters (Cressie, 1991).

Kriging is a spatial prediction method used in geostatistics to minimize the quadratic error and obtain weights that depend on the distances between individual dwellings and the dwelling being assessed. The autocorrelation structure is detected by means of a variogram (Matheron, 1970). Kriging generally attributes less weight to dwellings farthest from the one being assessed and more weight to the nearest ones. Therefore, kriging can be considered a weighted average technique that takes into account the presence of spatial autocorrelation (Felus, 2001).

Based on the widely accepted hypothesis in the literature that the value of urban properties is fundamentally conditioned by structural and locational characteristics, a method is proposed for estimating urban rental housing prices taking into account the influence of both types of characteristics. Specifically, the method used is geostastical kriging combined with econometric regression. Because this is a widely-known method, an in-depth explanation is not provided in this paper.

One variant of the kriging method is regression-kriging (RK), which is a combination of econometric and geostatistical methods (Chica-Olmo, 1995; Dubin, 1992; Hengl et al., 2007). Some recent studies have applied this methodology in the field of property valuation, particularly for house price valuations (Bajat et al., 2018; Chica-Olmo et al., 2019).

To estimate the rental prices of houses located at point s0, whose price is not known, the following best linear unbiased predictor can be used (Christensen, 1987):

$$\hat{y}(S_0) = x(S_0)'\hat{\beta} + \lambda' e(S_i) \quad (1)$$

where $x(S_0)$ is a vector which represents the values of the independent variables (apartment age, surface area, bathrooms, etc.) in S_0; $\hat{\beta}$ is the vector of parameters estimated by generalized least squares (GLS); λ is the vector of kriging weights and $e(S_i)$ is a vector of residuals in the sample points, S_i. These weights (λ) are obtained taking into account the spatial autocorrelation structure given by the variogram function *(y(h))* (Matheron, 1970). In our study, the residuals represent the effect of the location variables not included in the regression model. The spatial autocorrelation structure of the residuals is determined by the following empirical variogram:

$$\widehat{\Upsilon}_e(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [e(S_i + h) - e(S_i)]^2 \quad (2)$$

where (S_i + h) and (S_i) are the location of the apartments in the sample and N (h) is the number of h distant location-pairs. To apply the kriging method, it is necessary to adjust the empirical variogram model. The exponential model is one of the most widely used models in the literature (Webster & Oliver, 2007):

$$\text{Exponential model: } \gamma(h) = \begin{cases} C_0 + C\left[1 - exp\left(-\frac{h}{a}\right)\right], & |h| > 0 \\ 0, & |h| = 0 \end{cases} \quad (3)$$

where C_0 is the nugget effect, a is the range, C is the partial sill, and the effective range is equal to 3a. The range measures the distance at which the location variables cease to have an influence on rental housing prices.

3.3.2 Data

Athens is the capital of Greece and currently the country's largest city. It is the main center of Greek economic, cultural, and political life. The metropolitan area of Athens extends over the plains of the Attica peninsula, and is bordered by the Saronic Gulf to the south, Mount Aigaleo to the west, Mount Parnitha to the northwest, Mount Pentelicus to the northeast, and Mount Hymettus to the east. The original city of Athens was in the center of the plain but following the urban expansion in the 20th century the city has merged with surrounding populations, giving rise to the metropolitan area, which includes 54 municipalities.

In order to study apartment rental prices in the city of Athens, a sample of 757 residential apartments distributed throughout the city was collected in the second quarter of 2014. Given that there is no official register of rented housing (Magginas & Pateli, 2009), the data for this study were obtained from various real estate company websites (www.tospitimou.gr, www.spitogatos.gr, www. spiti24.gr, www.acropolis- realestate.gr).

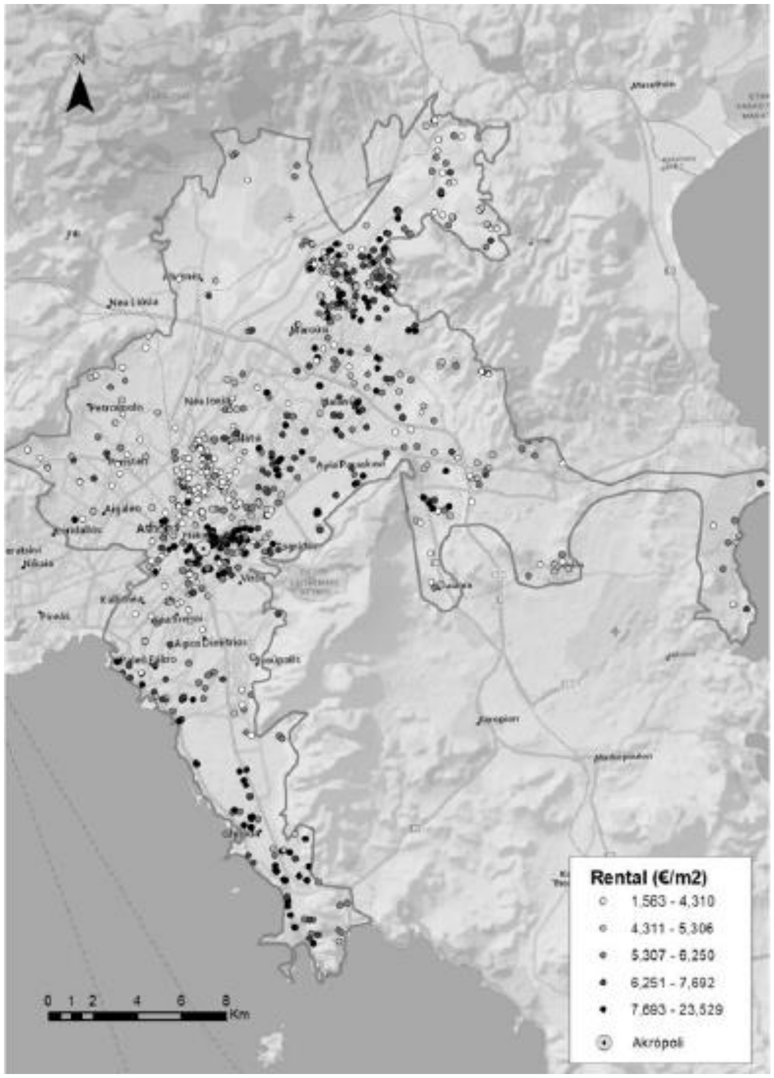
The dependent variable in the model is rental housing price measured in euros (Rental housing). The explanatory variables are primarily structural characteristics. As a result, we have included the age of the building in years (Age); total built surface area in square meters (Surface); number of bedrooms (Bedrooms); number of additional toilets apart from the main bathroom (WC); whether or not the dwelling has a storage room (Storage room); whether or not it has air conditioning (Air conditioned); whether or not it has views (View); whether or not it has a fireplace (Fireplace); whether or not it has heating (Heat); whether or not it has an elevator (Elevator); whether or not it is furnished (Furnished), and whether or not it has a garage (Garage). The basic descriptive statistics of all the variables are shown in Table 2.

Table 2. Descriptive statistics of the variables

	Mean	Mode	SD	Min	Max
Dependent variable					
Rental housing prices	817.713	800	673.819	120	6,000
Explanatory variables					
Age	23.460	44	17.227	0	114
Surface	127.974	100	82.566	25	600
Bedrooms	2.336	2	1.081	0	6
WC	0.526	1	0.569	0	3
Storage room	0.472	0	0.500	0	1
Air conditioned	0.361	0	0.481	0	1
View	0.421	0	0.494	0	1
Fireplace	0.462	0	0.499	0	1
Heat	0.966	1	0.182	0	1
Elevator	0.627	1	0.487	0	2
Furnished	0.292	0	0.455	0	1
Garage	0.472	0	0.500	0	1

Figure 2 shows the spatial distribution of the housing in the sample and the rental price per square meter in Athens by quintiles. The results show that the highest prices are concentrated in the Acropolis area, which occupies part of the city center (corresponding to the major tourist area); in Glyfada, which is in the south of the city (coastal area); in the northeast, and in some areas of the north where there are numerous luxury homes. In contrast, the lowest prices are found in the central area north of the Acropolis, close to several universities, as well as in the eastern and western areas furthest from the city center.

Figure 2. Location of the housing in the sample



3.4 Results and Discussion

In this study, we used a semi-logarithmic hedonic model widely employed in the literature on property values (Nishi et al., 2019). This type of specification facilitates normal model disturbances and interprets the model coefficients (multiplied by 100) of the continuous variables as the percentage impact of this variable on rental housing prices. For the dummy variables, the expression $[\exp(\beta) - 1] * 100$ is used to interpret the percentage impact (Halvorsen & Palmquist, 1980).

As indicated, in addition to modeling rental housing prices in Athens using the classical OLS method, we used RK to run a model with the GLS estimator. Table 3 shows the estimates of the model coefficients using OLS and RK. Severe multicollinearity was not detected in the model given that the highest variance inflation factor (VIF) is for the Age variable (7.4326) and does not exceed the value of 10. However, the model disturbances exhibit spatial autocorrelation given that the Moran's I statistic (I = 0.4242) is significant at the 5% level (p = 0.000). To consider the spatial autocorrelation in the model disturbances, we used an exponential variogram model whose parameters are shown in Table 3. The relative residual nugget effect (nugget/sill) is 0.33, thus indicating clear spatial dependence, which was detected due to the significant value of the Moran's I statistic. The presence of spatial dependence in the model disturbances could be due to the existence of variables that are not included in the model, although they exhibit spatial autocorrelation due to the location variables, such as spatial contagion and location characteristics. The experimental variogram in Figure 3 shows that as the distance between dwellings increases, the variability between the residuals also increases.

Three models (spherical, exponential, and Gaussian) were used to select the type of spatial autocorrelation structure, all of which yielded very similar results in terms of the R-squared of the RK model (Gaussian = 0.8419, exponential = 0.8434, and spherical = 0.8439).

Ultimately, the exponential model was chosen as it is the most frequently used model in the literature (Webster & Oliver, 2007). The figure also shows the adjusted variogram model, whose effective range is equal to 2,635.818 meters (3*878.606). This distance indicates the area of influence of local variables on rental housing prices in Athens.

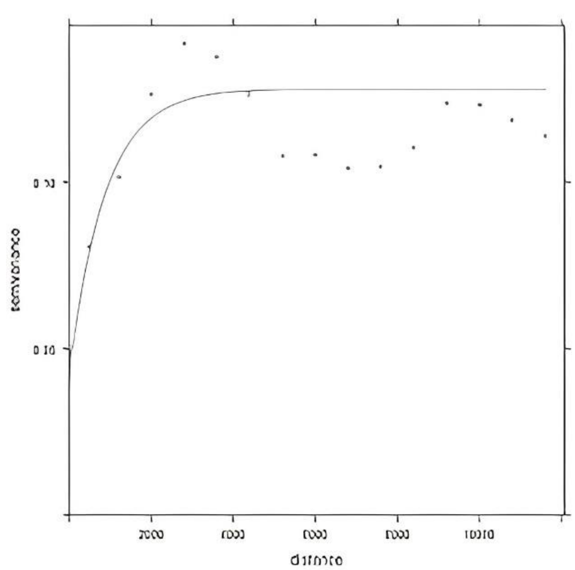
As can be seen in Table 3, the variables that are significant at the 5% and 10% levels have the expected signs in both models. As regards the goodness-of-fit of the model, it can be observed that the RK model is more suitable than the OLS one, since the R-squared has increased and the root-mean-square prediction error (RMSE) has decreased.

Table 3. Model estimated using OLS and RK. Dependent variable: Rental housing price

Explanatory variables	OLS	RK
Const.	5.214 (0.000)	5.4425 (0.000)
Age	-0.0055 (0.006)	-0.0084 (0.000)
Age-squared	0.0001 (0.000)	0.0001 (0.000)
Surface	0.0038 (0.000)	0.0038 (0.000)
Bedrooms	0.0596 (0.009)	0.0719 (0.000)
WC	0.1348 (0.000)	0.0768 (0.001)
Storage room	0.0811 (0.008)	0.0314 (0.020)
Air conditioned	0.0789 (0.005)	0.0298 (0.019)
View	0.0963 (0.000)	0.0679 (0.003)
Fireplace	0.2052 (0.000)	0.1107 (0.000)
Heat	0.1758 (0.015)	0.0984 (0.080)
Elevator	0.0761 (0.005)	0.0734 (0.002)
Furnished	0.1493 (0.000)	0.1279 (0.000)
Garage	0.1867 (0.000)	0.1775 (0.000)
Variogram		Exponential
Nugget	--	0.0432
Partial sill		0.0848
Range		878.606
Model fit		
R-squared	0.7504	0.8434
RMSE	0.3490	0.2768

Note: N = 757; p-values in parenthesis.

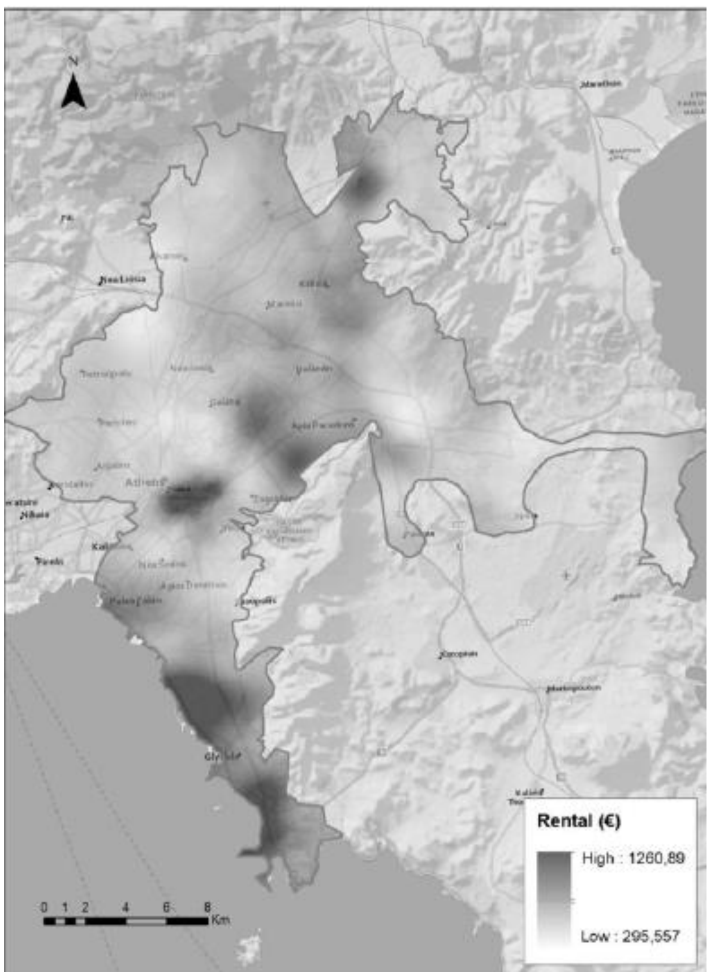
Figure 3. Experimental variogram and adjusted exponential model



From the point of view of the housing market, it is clearly of interest to determine housing prices at any given location and thus obtain isovalue maps (Clapp, 2004; Yue et al., 2010). To do so, it is first necessary to estimate the price of housing at any point on the map. However, given that it is impossible to know the structural characteristics for all points on the map, we defined a standard dwelling. A standard dwelling was obtained by assigning the numerical value of the sample mean if the variable was continuous or the mode if the variable was binary.

Figure 4 shows the estimated values for a standard dwelling in Athens. As can be observed, the highest prices are found in the center of the city, around the Acropolis.

Figure 4. Rental price of standard dwellings



There are several reasons for this. The main one is that the central business district and the most important monuments (Syntagma Square, the Cathedral, the Hellenic Parliament, most of the ministries, and almost all the embassies) are located in this area. This is also the central point where most of the shops, bars, restaurants, designer boutiques, museums, and other amenities and leisure facilities are located. In addition, there are a large number of hotels, banks, post offices, as well as subway, bus, and tram lines that provide easy access to the area.

However, the center is not the only area where high rental prices can be found. The prices in the southern coastal zone (Glyfada) are similar to in the central area. This is due to the proximity of the housing to the sea and the Port of Piraeus, which is the chief port in Greece and one of the most important in Europe. The presence of numerous hotels, the proximity to the city center, and the good transport links enable residents from these areas to enjoy the beaches while they can continue to work close to the main hub. In addition, there are numerous nautical companies, which are paramount to the country's economy.

Another high-priced area is located in the north (Kifissia), which is characterized by relatively new and exclusive housing owned by people with high purchasing power. This area is far enough from the center to escape the noise and hustle and bustle of the city and is surrounded by many green spaces and forests. The transport links to the center are good and there is quick access to the motorway out of Athens. Lastly, the area between Galatsi and Agia Paraskevi is also characterized by its high housing prices. This is due to the numerous squares, parks, avenues, and green areas surrounding both districts. The extension of the subway into these areas has also increased rental prices. In contrast, one of the areas with the lowest prices in the city due to socio-economic problems is located between Aigaleo, Galatsi, and Peristeri. Lastly, the eastern zone also has low prices because it is far from the urban center and has poor transport links.

The results show the areas where the rents for residential properties are higher due to their cultural and environmental attractions, among other location characteristics. These same attractions are also highly valued in tourism accommodations. As a result, some of the available housing can be exploited for purely tourism purposes, which could lead to the touristification phenomenon, as is being observed in other geographical areas (Blanco-Romero et al., 2018; Yrigoy, 2019). This phenomenon consists of the displacement of an area's habitual residents by tourists as a result of the high demand for rental accommodation for tourism purposes.

Therefore, tools to determine the impact of location on rental housing prices could assist relevant authorities in their decision making. For example, the Athens City Council may be interested in determining the impact of location on rental prices in order to control the undesirable effects of tourist accommodation rental prices.

The methodology used in this study can also help in detecting the housing rent gap. The rent gap is calculated as the disparity between residential rentals and tourism rentals, which leads to more intensive use of housing (Ley, 1986; Smith, 1987). Moreover, models that help to determine and spatially represent rental prices can be useful for managing social policies that enable citizens to access an essential asset such as housing.

3.5 Conclusions

This study highlights the impact of location on rental housing prices by showing how the rent of a standard dwelling in the city of Athens varies depending on its location. To do so, we hypothesized that rental housing prices are not randomly spatially distributed, but rather that neighboring dwellings are affected by location factors in a similar way.

In addition, we assumed that neighboring homeowners will tend to offer similar rental prices due to the spatial contagion effect. Both location factors and spatial contagion lead to spatial autocorrelation in rental housing prices. To achieve the results obtained in this study, a geostatistical method was used (regression-kriging) that takes into account the structural characteristics of a dwelling, as well as the spatial autocorrelation caused by undetected location factors and spatial contagion.

The methodology enabled us to obtain an isovalue map of rental housing prices for standard dwellings, which can also be interpreted as a locational isovalue map given that the structural characteristics of standard dwellings are constant for the entire plane. The price of standard dwellings is higher in the most attractive areas (the Acropolis, Glyfada, Kifissia, Galatsi, etc.) since these are historic areas with higher environmental quality (proximity to the beach, green areas, etc.) and better transport links.

It should also be noted that these results may serve to anticipate the emergence of the touristification phenomenon given that these areas are attractive for both residential and tourism use.

We would like to highlight the usefulness of the methodology, which could help investors to determine where to target investments and assist public authorities in deciding where to focus urban management policies in order to control the undesirable effects of an excessive rise in rents caused by tourism.

Finally, one advantage of the methodology used in this work is that it can be applied to any city in the world, as data on rental housing prices, structural characteristics, and spatial location are easy to obtain through searches in real estate portals. The main differences between some cities and others would mainly be related to locational factors, which are difficult to reproduce in space, since they are specific to each city. However, another advantage of the proposed methodology is that it is not strictly necessary to specify these factors in the model.

3.6 References

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models* (Vol. 4). Springer Netherlands. https://doi.org/10.1007/978-94-015-7799-1

Bajat, B., Kilibarda, M., Pejović, M., & Petrović, M. S. (2018). Spatial Hedonic Modeling of Housing Prices Using Auxiliary Maps. In J.-C. Thill (Ed.), *Spatial Analysis and Location Modeling in Urban and Regional Systems* (pp. 97–122). Springer. https://doi.org/10.1007/978-3-642-37896-6_5

Basu, S., & Thibodeau, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics*, *17*(1), 61–85. https://doi.org/10.1023/A:1007703229507

Beimer, W., & Maennig, W. (2017). Noise effects and real estate prices: A simultaneous analysis of different noise sources. *Transportation Research Part D: Transport and Environment*, *54*, 282–286. https://doi.org/10.1016/j.trd.2017.05.010

Blanco-Romero, A., Blázquez-Salom, M., & Cànoves, G. (2018). Barcelona, Housing Rent Bubble in a Tourist City. Social Responses and Local Policies. *Sustainability*, *10*(6), 2043. https://doi.org/10.3390/su10062043

Bourassa, S. C., Cantoni, E., & Hoesli, M. (2007). Spatial Dependence, Housing Submarkets, and House Price Prediction. *The Journal of Real Estate Finance and Economics*, *35*(2), 143–160. https://doi.org/10.1007/s11146-007-9036-8

Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, *22*(3), 453–474. https://doi.org/10.1016/0166-0462(92)90039-4

Cellmer, R. (2014). The Possibilities and Limitations of Geostatistical Methods in Real Estate Market Analyses. *Real Estate Management and Valuation*, *22*(3), 54–62. https://doi.org/10.2478/remav-2014-0027

Chau, K. W., & Chin, T. L. (2003). A Critical Review of Literature on the Hedonic Price Model. *International Journal for Housing Science and Its Applications*, *27*(2), 145–165.

Chen, J., Wu, Y., Guo, F., & Wang, H. (2018). Domestic property and housing class in contemporary urban China. *Journal of Housing and the Built Environment*, *33*(1), 91–109. https://doi.org/10.1007/s10901-017-9545-6

Chen, W. Y., & Jim, C. Y. (2010). Amenities and disamenities: A hedonic analysis of the heterogeneous urban landscape in Shenzhen (China). *The Geographical Journal*, *176*(3), 227–240. https://doi.org/10.1111/j.1475-4959.2010.00358.x

Chica Olmo, J. (1995). Spatial Estimation of Housing Prices and Locational Rents. *Urban Studies*, *32*(8), 1331–1344. https://doi.org/10.1080/00420989550012492

Chica-Olmo, J., Cano-Guervos, R., & Chica-Olmo, M. (2013). A Coregionalized Model to Predict Housing Prices. *Urban Geography*, *34*(3), 395–412. https://doi.org/10.1080/02723638.2013.778662

Chica-Olmo, J., Cano-Guervos, R., & Chica-Rivas, M. (2019). Estimation of Housing Price Variations Using Spatio-Temporal Data. *Sustainability*, *11*(6), 1551. https://doi.org/10.3390/su11061551

Chica-Olmo, J., Cano-Guervos, R., & Tamaris-Turizo, I. (2019). Determination of buffer zone for negative externalities: Effect on housing prices. *The Geographical Journal*, *185*(2), 222–236. https://doi.org/10.1111/geoj.12289

Christensen, R. (1987). *Plane answers to complex questions: The theory of linear models*. Springer-Verlag.

Clapp, J. M. (2004). A Semiparametric Method for Estimating Local House Price Indices. *Real Estate Economics*, *32*(1), 127–160. https://doi.org/10.1111/j.1080-8620.2004.00086.x

Cressie, N. (1991). *Statistics for Spatial Data*. John Wiley & Sons.

Cui, N., Gu, H., Shen, T., & Feng, C. (2018). The Impact of Micro-Level Influencing Factors on Home Value: A Housing Price-Rent Comparison. *Sustainability*, *10*(12), 4343. https://doi.org/10.3390/su10124343

Danieli, Y., Stamatopoulou, E., & Dias, C. (2018). *The Universal Declaration of Human Rights: Fifty Years and Beyond*. Routledge. https://doi.org/10.4324/9781315223469

Diao, M., Leonard, D., & Sing, T. F. (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics*, *67*, 64–77. https://doi.org/10.1016/j.regsciurbeco.2017.08.006

Diao, M., Qin, Y., & Sing, T. F. (2016). Negative Externalities of Rail Noise and Housing Values: Evidence from the Cessation of Railway Operations in Singapore. *Real Estate Economics*, *44*(4), 878–917. https://doi.org/10.1111/1540-6229.12123

Dubin, R. A. (1988). Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms. *The Review of Economics and Statistics*, *70*(3), 466–474. https://doi.org/10.2307/1926785

Dubin, R. A. (1992). Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics*, *22*(3), 433–452. https://doi.org/10.1016/0166-0462(92)90038-3

Dubin, R. A. (1998). Spatial Autocorrelation: A Primer. *Journal of Housing Economics*, *7*(4), 304–327. https://doi.org/10.1006/jhec.1998.0236

Dubin, R., Pace, R. K., & Thibodeau, T. G. (1999). Spatial Autoregression Techniques for Real Estate Data. *Journal of Real Estate Literature*, *7*(1), 79–96. https://doi.org/10.1023/A:1008690521599

Efthymiou, D., & Antoniou, C. (2013). How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. *Transportation Research Part A: Policy and Practice*, *52*, 1–22. https://doi.org/10.1016/j.tra.2013.04.002

Eurostat. (2018). *Housing statistics*. Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Housing_statistics/es

Fabusuyi, T. (2018). Is crime a real estate problem? A case study of the neighborhood of East Liberty, Pittsburgh, Pennsylvania. *European Journal of Operational Research*, *268*(3), 1050–1061. https://doi.org/10.1016/j.ejor.2017.12.003

Fan, Q., Hansz, J. A., & Yang, X. (2016). The Pricing Effects of Open Space Amenities. *The Journal of Real Estate Finance and Economics*, *52*(3), 244–271. https://doi.org/10.1007/s11146-015-9508-1

Felus, Y. A. (2001). *New methods for spatial statistics in geographic information systems* [Doctoral dissertation, The Ohio State University].

https://etd.ohiolink.edu/acprod/odb_etd/r/etd/search/10?p10_accession_num=osu1486400446373917&clear=10&session=114911163308190

Gallin, J. (2008). The Long-Run Relationship Between House Prices and Rents. *Real Estate Economics*, *36*(4), 635–658. https://doi.org/10.1111/j.1540-6229.2008.00225.x

Goodall, B. (2013). *The Economics of Urban Areas* (1st ed.). Pergamon.

Halvorsen, R., & Palmquist, R. (1980). The Interpretation of Dummy Variables in Semilogarithmic Equations. *The American Economic Review*, *70*(3), 474–475.

Harris, C. D., & Ullman, E. L. (1945). The Nature of Cities. *The Annals of the American Academy of Political and Social Science*, *242*(1), 7–17. https://doi.org/10.1177/000271624524200103

Hengl, T., Heuvelink, G. B. M., & Rossiter, D. G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences*, *33*(10), 1301–1315. https://doi.org/10.1016/j.cageo.2007.05.001

Hoch, I., & Waddell, P. (1993). Apartment Rents: Another Challenge to the Monocentric Model. *Geographical Analysis*, *25*(1), 20–34. https://doi.org/10.1111/j.1538-4632.1993.tb00277.x

Iliopoulou, P., & Stratakis, P. (2018). Spatial analysis of housing prices in the Athens Region, Greece. *RELAND: International Journal of Real Estate & Land Planning*, *1*, 304–313. https://doi.org/10.26262/reland.v1i0.6486

Korpi, W. (2018). *The Democratic Class Struggle*. Routledge. https://doi.org/10.4324/9780429441714

Krause, A. L., & Bitter, C. (2012). Spatial econometrics, land values and sustainability: Trends in real estate valuation research. *Cities*, *29*(Suppl. 2), S19–S25. https://doi.org/10.1016/j.cities.2012.06.006

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, *74*(2), 132–157. https://doi.org/10.1086/259131

Le Boennec, R., & Salladarré, F. (2017). The impact of air pollution and noise on the real estate market. The case of the 2013 European Green Capital: Nantes, France. *Ecological Economics*, *138*, 82–89. https://doi.org/10.1016/j.ecolecon.2017.03.030

Leamer, E. E. (2002). Bubble Trouble? Your Home Has a P/E Ratio Too. *UCLA Anderson Forecast*.

Lee, C.-I. (2007). Does provision of public rental housing crowd out private housing investment? A panel VAR approach. *Journal of Housing Economics*, *16*(1), 1–20. https://doi.org/10.1016/j.jhe.2006.09.004

Ley, D. (1986). Alternative Explanations for Inner-City Gentrification: A Canadian Assessment. *Annals of the Association of American Geographers*, *76*(4), 521–535. https://doi.org/10.1111/j.1467-8306.1986.tb00134.x

Magginas, N., & Pateli, E. (2009). *Greek residential real estate market: Recovery in sight by mid-2010*. National Bank of Greece, Economic and Market Analysis.

Matheron, G. (1970). *La Théorie des variables régionalisées, et ses applications*. Ecole Nationale Supérieure des Mines de Paris.

Moye, R., & Thomas, M. (2018). “Race and Housing Values: What Happens When Whites Don’t All Move Out?” *City & Community*, *17*(1), 109–133. https://doi.org/10.1111/cico.12289

Nasreen, Z., & Ruming, K. (2019). Room Sharing in Sydney: A Complex Mix of Affordability, Overcrowding and Profit Maximisation. *Urban Policy and Research*, *37*(2), 151–169. https://doi.org/10.1080/08111146.2018.1556632

Nishi, H., Asami, Y., & Shimizu, C. (2019). Housing features and rent: Estimating the microstructures of rental housing. *International Journal of Housing Markets and Analysis*, *12*(2), 210–225. https://doi.org/10.1108/IJHMA-09-2018-0067

Osland, L. (2010). An Application of Spatial Econometrics in Relation to Hedonic House Price Modeling. *Journal of Real Estate Research*, *32*(3), 289–320. https://doi.org/10.1080/10835547.2010.12091282

Ozanne, L., & Thibodeau, T. (1983). Explaining metropolitan housing price differences. *Journal of Urban Economics*, *13*(1), 51–66. https://doi.org/10.1016/0094-1190(83)90045-1

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: A review of valuation methods. *Journal of Property Investment & Finance*, *21*(4), 383–401. https://doi.org/10.1108/14635780310483656

Papadimitriou, D. B., Nikiforos, M., & Zezza, G. (2013). *The Greek Economic Crisis and the Experience of Austerity: A Strategic Analysis*. The Levy Economics Institute of Bard College. https://www.levyinstitute.org/publications/the-greek-economic-crisis-and-the-experience-of-austerity

Ramos, P. (2015, July 10). *Grecia, un escaparate de viviendas a precios democratizados*. EL MUNDO. https://www.elmundo.es/economia/2015/07/10/559fa01c268e3efc178b457c.html

Rex, J. A. (1968). The Sociology of a Zone of Transition. In R. E. Pahl (Ed.), *Readings in Urban Sociology* (1st ed., pp. 211–231). Elsevier. https://shop.elsevier.com/books/readings-in-urban-sociology/pahl/978-0-08-013293-8

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Smith, N. (1987). Gentrification and the Rent Gap. *Annals of the Association of American Geographers*, *77*(3), 462–465. https://doi.org/10.1111/j.1467-8306.1987.tb00171.x

Swoboda, A., Nega, T., & Timm, M. (2015). Hedonic Analysis Over Time and Space: The Case of House Prices and Traffic Noise. *Journal of Regional Science*, *55*(4), 644–670. https://doi.org/10.1111/jors.12187

Taipalus, K. (2006). A global house price pubble? Evaluation based on a new rent-price approach. *Bank of Finland Research Discussion Papers*, Paper 29. https://doi.org/10.2139/ssrn.1018329

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234–240. https://doi.org/10.2307/143141

Trojanek, R., & Gluszak, M. (2018). Spatial and time effect of subway on property prices. *Journal of Housing and the Built Environment*, *33*(2), 359–384. https://doi.org/10.1007/s10901-017-9569-y

Tse, R. Y. C. (2002). Estimating Neighbourhood Effects in House Prices: Towards a New Hedonic Model Approach. *Urban Studies*, *39*(7), 1165–1180. https://doi.org/10.1080/00420980220135545

von Thünen, J. H. (1966). *Isolated State: An English Edition of Der Isolierte Staat*. Pergamon Press.

Webster, R., & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*. John Wiley & Sons.

Wen, H., Jin, Y., & Zhang, L. (2017). Spatial heterogeneity in implicit housing prices: Evidence from Hangzhou, China. *International Journal of Strategic Property Management*, *21*(1), 15–28. https://doi.org/10.3846/1648715X.2016.1247021

Wheaton, W. C. (1977). A bid rent approach to housing demand. *Journal of Urban Economics*, *4*(2), 200–217. https://doi.org/10.1016/0094-1190(77)90023-7

Wulff, M. N. N., & Maher, C. (1998). Long-term Renters in the Australian Housing Market. *Housing Studies*, *13*(1), 83–98. https://doi.org/10.1080/02673039883506

Yrigoy, I. (2019). Rent gap reloaded: Airbnb and the shift from residential to touristic rental housing in the Palma Old Quarter in Mallorca, Spain. *Urban Studies*, *56*(13), 2709–2726. https://doi.org/10.1177/0042098018803261

Yue, W., Liu, Y., & Fan, P. (2010). Polycentric Urban Development: The Case of Hangzhou. *Environment and Planning A: Economy and Space*, *42*(3), 563–577. https://doi.org/10.1068/a42116

CAPÍTULO 4. The use of geostatistics to estimate missing data in a spatial econometric model of housing prices

Ivan Tamaris-Turizo[1]
Jorge Chica-Olmo[2]
Rafael Cano-Guervos[2]

[1]*Universidad Pedagógica y Tecnológica de Colombia UPTC, COLOMBIA*
[2]*Universidad de Granada, Granada, ESPAÑA*

Abstract

Housing prices have been the subject of many studies, and some of them have tried to determine the influencing structural and location factors through hedonic econometric models. One of the main factors considered in the literature on real estate appraisals is the location of the dwellings. For this reason, this study combines the spatial methodologies of geostatistics and spatial econometrics. On the one hand, this work uses geostatistics to estimate missing data to account for the lack of information in the sampled real estate websites. On the other hand, the explanatory factors of prices are determined through spatial econometrics. The combination of both methods facilitates estimating housing prices in Santa Marta (Colombia), solving the problem of missing data. In the modeling, the problems of spatial heteroscedasticity and multicollinearity are corrected. This combination of methods could be of great interest to companyies and public agencies related to real estate activity, which is sustained by the information available on these real estate websites.

Keywords: Hedonic mode; geostatistics; spatial econometrics; housing prices; missing data; heteroscedasticity

4.1 Introduction

When it comes to estimating the price of urban real estate, the theory of hedonic prices (which values property according to the profit generated by its features) is one of the most widely used, as it allows for determining the implicit prices of the features that determine the price of such property. These implicit prices are obtained through the application of the hedonic regression model (Rosen, 1974), in which the coefficients of the explanatory features affecting the price of the property measure the profit they report. However, on many occasions, the application of the regression model causes problems related to multicollinearity (Chinloy, 1977; Mark, 1980; Tiwari & Parikh, 1998) and heteroscedasticity (Rehák & Káčer, 2019; Trojanek & Gluszak, 2018), leading to efficiency issues for the estimated coefficients of the model.

In one of his first works, Goodman (1978) used the hedonic price method to estimate housing prices based on the expression $P = (C)$, where P represents the price of each dwelling and C is the set of components that are considered to contribute to explain the price. The selling price of the dwelling will depend on the structural quality of these components. Therefore, under equilibrium conditions, the marginal price of each of the components will be equal to their marginal valuation (Rosen, 1974; Nelson, 1978). The econometric method is useful for estimating such marginal effects and the price of urban housing (Can, 1992).

The application of hedonic models to determine prices in the housing market has continued to be important, as it is considered an essential tool and a starting point for the analysis of the urban housing market (Can, 1990; Goodman & Thibodeau, 2003; Li et al., 2013; Zhang et al., 2019).

One of the most widely used methods for estimating econometric models is the Ordinary Least Squares (OLS), but this method does not consider spatial effects (Anselin, 1992). In the context of OLS, the model coefficients are estimated under the assumption of no spatial autocorrelation in the disturbances (Zhu et al., 2011). But this assumption is not usually met in the case of hedonic housing price models since prices are not independent of their spatial location (Case et al., 2004; Wen et al., 2017), leading to inefficient estimates (Anselin, 1988).

Subsequent research developed techniques that aim to overcome the limitations of the OLS method in the presence of spatial autocorrelation by improving the estimate of prices, such as geostatistics (Dubin, 1992; Chica-Olmo, 1995; Basu & Thibodeau, 1998; Yoo and Kyriakidis, 2009; Gröbel & Thomschke, 2018) and spatial econometrics (Can, 1992; Pace et al., 1998; Von Graevenitz & Panduro, 2015; Wen et al., 2017).

The features that explain housing prices can be classified into construction and location (Chica-Olmo, 2007; Li et al., 2012; Li et al., 2019; Poeta et al., 2019). The construction features most frequently used in hedonic models are housing area, number of bedrooms, number of bathrooms, kitchen, and garage (with positive effects on house price), as well as age (with a negative effect on housing price) (Goodman & Thibodeau, 2003; Clapp, 2004; Conroy & Milosch, 2011; Sah et al., 2015; Li et al., 2016; Xiao et al., 2017).

In addition, the locational features are determined by the location of the dwelling in relation to its surroundings (Li et al., 2013; Simlai, 2014; Sah et al., 2015; Chica-Olmo et al., 2019). For this reason, it is presumed that houses close to each other tend to have similar prices. According to Tobler's (1970) geographic law, "everything is related to everything else, but near things are more related than distant things." This explains why housing prices show a high degree of spatial dependence since location factors that strongly influence price, such as accessibility or proximity to externalities, have a similar influence on nearby dwellings. Specifically, location can generate a positive effect if the property is close to an amenity or attraction, but it can also have a negative effect when the property is affected by a negative externality in the neighborhood (Can, 1992, Zhang et al., 2014). For example, in some cases, factors such as noise, safety or vibrations have an adverse effect on home appraisals (Bowes & Ihlanfeldt, 2001), or other factors such as atmospheric pollution, electromagnetic radiation, noise pollution, taxes, among other factors, also have a negative impact on housing prices (Chay & Greenstone, 2005; Geng et al., 2015). This spatial dependence on housing prices will also lead to spatial autocorrelation in the disturbances of the econometric model (Can, 1998; Basu & Thibodeau, 1998).

Regarding Latin America, some studies on hedonic prices have been carried out. They have included construction and location variables in their models. Thus, including location variables has improved the results of the models. For example, in the city of Medellín (Colombia), the results of the model improved after including the spatial component (Urrea et al., 2019). Some studies also showed that structural features had more influence on housing prices than locational features, as shown in the city of Juarez (Mexico) (Fierro et al., 2009) or in the Metropolitan Region of Chile, where age and surface area were the most relevant variables (Sagner, 2011).

However, other characteristics related to the quality of life and income level have also been used to estimate hedonic price models, as in the city of Buenos Aires (Argentina), where the results were consistent and showed a strong correlation between price, income level and neighborhood characteristics (Cruces et al., 2008). Similarly, in the city of Santa Marta (Colombia), a strong relationship was found between income level and the demand for this real estate (Garza & Ovalle, 2019).

It is important to consider that public services are not adequately provided in several areas of Latin America. This deficit in services generates a differential that affects housing prices in these areas. Thus, in Machala (Ecuador), a study to assess the price of housing rental concluded that the most relevant variables, those with the greatest impact, were the quality and availability of public services, together with the distance to the city's central park (Zambrano-Monserrate, 2016). In Brazil, it was observed that areas with an adequate water supply and garbage collection services significantly increase housing rental prices (Morais & Cruz, 2015).

In the Latin American region, there have also been studies that considered the ease of access to public transportation through the proximity of the dwelling to an urban transportation station. These studies show that proximity has an impact but in different ways. According to Rodríguez & Targa (2004) y Muñoz-Raskin (2010), proximity to transport stations positively influences housing values in the case of Bogota (Colombia), while in Buenos Aires (Argentina), the impact of proximity to transport stations on property prices is not so significant (D'Elia et al., 2020). However, in Santa Marta (Colombia), it has been observed that proximity to freight railways has a negative effect on housing prices (Chica-Olmo et al., 2019).

A problem when applying regression models to models involving different information sources is the absence of information on some explanatory variables. Sometimes the sample information is not consistently provided by the different sources. This problem must be addressed by using some statistically valid method. Anselin (1992) suggests solving this type of difficulty by applying the Kriging interpolation technique to predict the values at the locations where no observations are available. Therefore, this is a tool of special importance when seeking to improve the estimates in case data are unavailable for all the variables at the sampling points.

In this work, three hypotheses are proposed:

H1. Incorporating missing data information improves the quality and significance of the model variables.

H2. Housing prices in the city of Santa Marta show an increasing spatial trend from the outskirts to the coast.

H3. Housing prices do not necessarily have to be related to some points of interest.

This paper is structured as follows: after the literature review related to hedonic prices and the different methods used for their determination, there is a brief explanation of the spatial methodologies used in this work, such as geostatistics and spatial econometrics. Then, the main results obtained by applying the aforementioned methods are presented. Finally, the conclusions of the study are included.

4.2 Metodology

4.2.1 Geostatistics

Classically, hedonic models have been estimated using Ordinary Least Squares (OLS). However, this technique may present problems since it does not consider the presence of spatial dependence, which means that the housing price is not independent of its location in space. In addition, it may present structural instability, which implies that the parameters take different values if the property is located in a certain area or not (Baronio et al., 2018).

Fortunately, the development of Geographic Information Systems (GIS), spatial econometrics and geostatistics has enabled to consider the presence of spatial dependence.

Through GIS, a cartographic model is generated that shows the relationship between the housing price and its location, i.e., the spatial distribution of housing prices can be observed. According to Can (1998), GIS is a powerful tool for obtaining a geographic visualization of dwellings to facilitate spatial analysis and data management.

GISs also contribute to spatial analysis using geostatistical methods, which are extremely useful tools to be applied in real estate. Also, through the application of the geostatistical Kriging interpolation tool, it is possible to extend the estimates to areas where there is a lack of sample information.

Kriging has proven to be a useful tool not only in real estate or mining, where it originated but also in other fields of study. For example, in the area of health, it was possible to identify the geographic disparities of breast cancer in the USA, using Kriging to solve the problem of the lack of homogeneous information caused by a change in the data collection methodology, with almost 70% of missing data (Chien et al., 2013). This geostatistical tool has also been used in meteorology to solve problems of incomplete information. According to Kilibarda et al. (2014), it is required to use interpolation methods when forecasting daily temperatures from spatio-temporal data. When pixel loss of more than 50% occurs, Kriging regression models are used to obtain better maps resulting from a global geostatistical model.

Geostatistics has also been used in real estate appraisals with missing data through the Cokriging method. This method was used to estimate the prices of commercial premises in Toledo (Spain), with auxiliary information provided by the prices of surrounding dwellings, which significantly improved the quality of price predictions (Montero-Lorenzo et al., 2009). This methodology has also been applied in the city of Granada when using heterotopic data samples to estimate housing prices (Chica-Olmo, 2007).

Besides geostatistics, other research solved the problem of missing data using different methods. Lesage & Pace (2004,) in their study on hedonic prices to determine housing values, used a Monte Carlo experiment to improve the sample, managing to reduce the prediction errors by more than 75%. Also, Kelejian & Prucha (2010) applied a model with spatial lags to overcome this problem. They worked with the complete sample and obtained better estimators that minimize the disturbance terms associated with the spatial correlation. In addition, Li & Li (2018) addressed this problem by using a Cubic Spline Interpolation Model to analyze the negative relationship between the bad odor from landfills and housing prices in Hong Kong.

4.2.2 Spatial Econometrics

Furthermore, in order to carry out the spatial econometric modeling, an Exploratory Spatial Data Analysis (ESDA) can be performed to contrast the possible presence of spatial autocorrelation (Anselin et al., 2006). Finally, SAR/SEM spatial regression models are estimated, and the Lagrangian Multiplier (LM) is used (Anselin, 1988) to choose the model with the best significance.

In spatial econometrics, when considering the presence of spatial autocorrelation, it is necessary to specify the form and degree of the neighborhood between georeferenced data. For this purpose, the W-weight matrix is used, which helps to establish multidirectional spatial relationships, i.e., the interdependence between regions (Can, 1990; Corrado & Fingleton, 2011; Anselin & Rey, 2012).

To carry out the house pricing estimate in this study, we will use spatial econometrics and geostatistics as techniques that rely on spatial effects related to location. On the one hand, the lack of information in some sample observations has been overcome by applying the Kriging spatial interpolator, which helped to enlarge the sample. On the other hand, spatial econometrics has been applied to perform tests to validate assumptions and estimate the effects of price drivers.

When data is searched throughout space, the value will depend on the location, which translates into structural instability in space because the observations are not homogeneous. This aspect implies heterogeneity in the data, so the previous condition implies that heteroscedasticity problems are generated (Anselin, 1988). A proposal to solve this problem is based on the application of spatial models developed by Cliff & Ord (1973, 1981,) widely used in the field of spatial modeling (Kelejian & Prucha, 1998, 1999, 2010; Arraiz et al., 2010), which is based on the following expression:

$$y = \rho W y + X\beta + u$$

$$u = \lambda W u + \varepsilon$$

In this expression, y is a n x 1 vector of observations of the explained variable, X is the n x k matrix of observations of k explanatory variables, W is the n x n matrix of known spatial weights, β is a k x 1 vector of regression parameters, λ and ρ are scalar autoregressive parameters, u is a n x 1 vector of regression disturbances which may present spatial autocorrelation and ε are normally and independently distributed disturbances.

If λ and ρ are different from zero, the model is called SARAR. When ρ is different from 0 and λ is equal to zero, it is called SAR (Spatial autoregressive), and when ρ is equal to 0 and λ is different from zero, it is called SEM (Spatial Error Model). Of these three models, SAR and SEM are the most widely used in practice. Lagrange tests (LM) are used (Anselin et al., 1996) to determine the most appropriate model.

4.3 Data and area of study

This study estimates housing prices in the city of Santa Marta (Colombia), which is the capital of the Magdalena Department. The city has a population of 479.000 inhabitants (according to the

2018 census of the National Administrative Department of Statistics (from the Spanish Departamento Administrativo Nacional de Estadística, DANE). This city has an area of 2,393 km2, of which 36.6 km2 belong to the urban area. The urban area is known for being monocentric, having its main commercial and business center next to the coastal area on the shores of Santa Marta Bay, which belongs to the Caribbean Sea. According to the DANE census, 75% of the dwellings are single-family homes.

The information related to the sample data of the properties for sale in Santa Marta was obtained from different real estate agency websites. Specifically, the real estate websites consulted were Metrocuadrado, Safe, Coldwell Banker and Habitat Kasa. Data collection focused on the urban perimeter of the city of Santa Marta, in particular, on second-hand single- family dwellings for sale. The prices offered for these properties were obtained from these websites. The construction or structural features observed were the price of the dwelling (RV), expressed in millions of Colombian pesos; number of bedrooms (HABITAC); number of bathrooms (BANOS); constructed area in square meters (M2); age, in years (ANTIG); the dichotomous variables GARAGE: 1 if the dwelling has a garage and 0 if it does not; and PISCINA: 1 if the dwelling has a swimming pool and 0 if it does not. This cross-sectional information, obtained from real estate websites during the first half of 2016, represents approximately 24% of the dwellings offered in that period, as approximately 900 single-family properties were for sale.

In addition, the following location variables were considered: distance to the nearest university (DUNIV), distance in km to the nearest supermarket (DSUPERMDO), distance in km to the nearest shopping center (DCTROCCIAL), distance in km to the center (DISTC) and the UTM coordinates of each dwelling in deviations from the mean (COORDXM and COORDYM). These coordinates allow considering the increasing spatial trend of prices (Chica-Olmo, 1995 and

2007).

The information is cross-sectional and corresponds to dwellings located in the urban perimeter of the city of Santa Marta (between the 1st and 66th lanes and between the 4th and 50th streets approximately). It is important to mention that the information provided by the different real estate websites is heterogeneous. In particular, not all the websites provided the age of the property. In short, within the sample of 211 dwellings, only complete data was available for 142 dwellings. Figure 6 shows the location of the sample dwellings and the price. This figure shows that the highest prices are found close to the center of the city, which is located near the seaport. There is also a spatial drift determined by a plane that goes from the coast (with high values) to the city outskirts (with lower values). It is also observed that there are points of interest such as universities, supermarkets or shopping centers, both in the downtown area, where values are high and, in the outskirts, where values are low.

Figure 5 shows the outline that summarizes the development of the work. First, a base model (Mod 0) was estimated in which regression was performed with the data from the original sample, with the distinction that the age variable was not found in all the dwellings. After that, Kriging was applied to estimate the missing data for age, and a new model (Mod 1) was obtained. The orthogonality method was then used to solve the multicollinearity problem, obtaining a new model (Mod 2). Finally, based on Anselin et al. (1996), two spatial SEM models are estimated, one where heteroscedasticity is not controlled (Mod 3) and the other where this issue is controlled (Mod 4.) Therefore, the assumptions for the validity of the model are fulfilled.

Figure 5. Outline of the research process

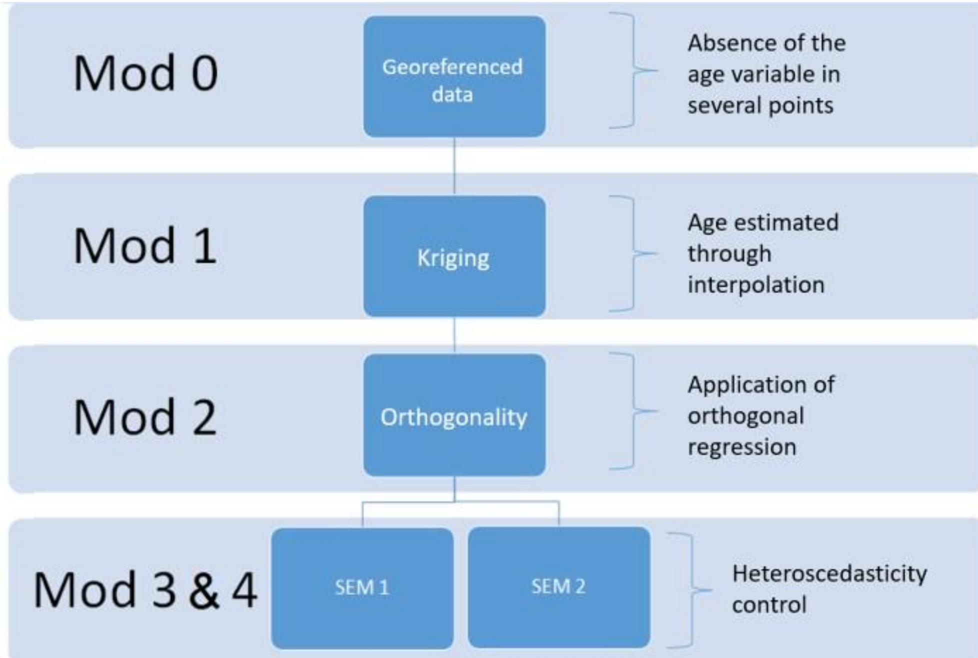
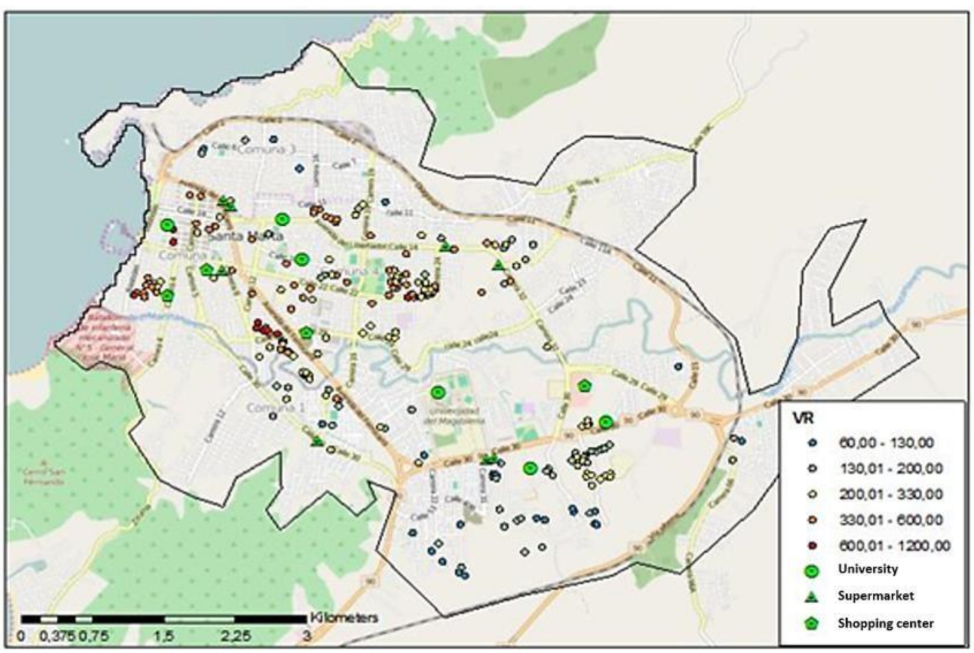


Figure 6. Spatial distribution of housing prices and some sites of interest in the city of Santa Marta



To overcome the problem of the lack of information on age, the Kriging method was used, which consists of an estimate through interpolation of the weighted average value that requires experimental and structural information (Chica-Olmo, 1994). That is to say, a statistical prediction of a value for the unsampled sites is obtained from the available data (Felus, 2001). Therefore, this method will estimate the missing data for the age variable. Figure 7 shows the isovalue map of age obtained through Kriging.

Figure 7. Spatial distribution of the age variable of dwellings in Santa Marta and their estimates obtained by kriging

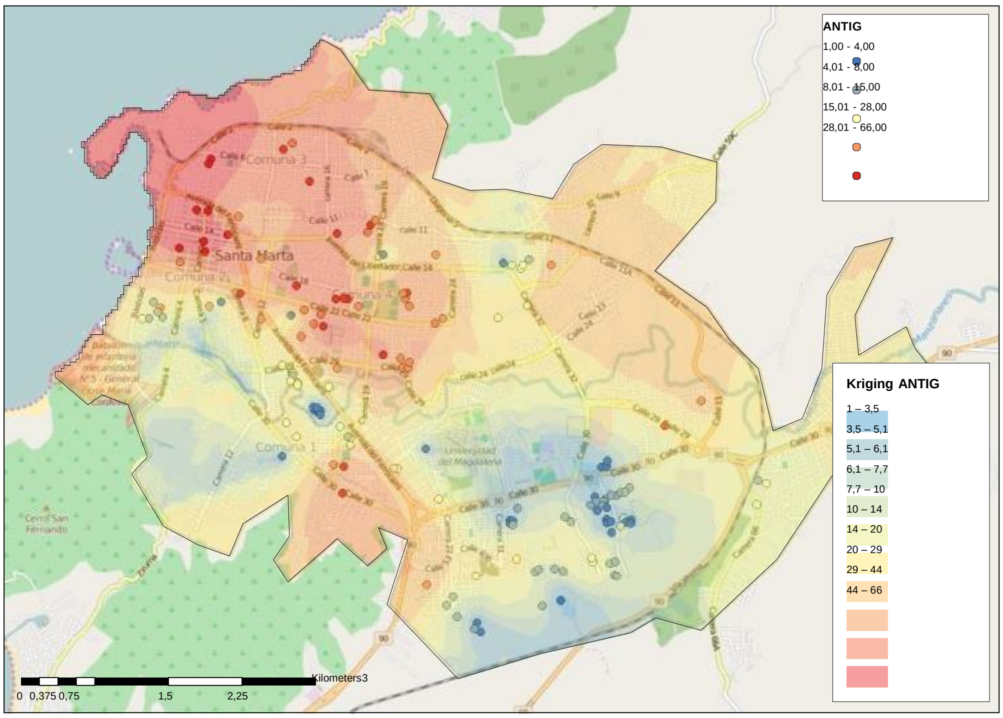


Table 4 shows the descriptive statistics of the construction and locational features. It should be noted that the variable ANTIG represents the age of sampled dwellings and the estimate by Kriging of the dwellings without available values. From the results of this table, it can be summarized that 87% of the properties have a garage, 15% have access to a swimming pool, and, approximately, on average, have 3 bedrooms, 3 bathrooms, 156 m2 and are 16 years old. It is noteworthy that the average number of bathrooms is 3, equal to the average number of bedrooms. It is also striking that there is one property with 6 bathrooms, as it is a luxurious property of about 615 m2 in an exclusive neighborhood of the city.

Table 4. Descriptive statistics

Variable	Mean	St. Dev.	Var. Coef.	Min	Max
VR	287.737	167.544	0.609	60	1200
GARAJE	0.872	0.335	0.410	0	1
HABITAC	3.242	0.739	0.277	2	6
BANOS	2.938	0.952	0.350	1	6
M2	156.545	93.116	0.672	50	615
PISCINA	0.152	0.360	2.536	0	1
DUNIV	0.858	0.399	0.465	0.160	1.923
DSUPERMDO	0.789	0.386	0.489	0.063	2.601
DCTROCCIAL	1.024	0.543	0.530	0.110	2.465
COORDXM	79.475	1,428.291	17.972	-2,707.738	3,672.262
COORDYM	-55.599	1,160.799	-20.878	-2,557.409	2,296.591
ANTIG	16.380	12.389	0.756	1	66

4.4 Results

This section presents the main results obtained from the methodology proposed in this study. First, we show the results of the model estimated by OLS, whose endogenous variable is the logarithm of housing prices and whose explanatory variables are the construction and locational features of the dwellings (Mod0-OLS, Table 5). To estimate this model, the 142 dwellings with all the information available were used. This model shows that all the construction variables are significant except ANTIG and PISCINA. As for the location variables, no variable related to the distance to points of interest (distance to the university, supermarkets or shopping centers) is significant, although the variables that include the distance to the city center and the spatial trend of prices, collected by the geographic coordinates are significant.

To solve the problem of missing data corresponding to the ANTIG variable, the Kriging method has been applied, and the missing values of this variable (69 dwellings) have been replaced by the estimates obtained using this method. Thus, a sample of 211 data was available. From these new data, Mod1-MCO was estimated (Table 5). In this model, the significance of the ANTIG variable and the quality of the model improved. Regarding the verification of the hypotheses of the regression model, serious multicollinearity problems were detected with the DISTC variable, as it presents an IVF (variance inflator factor) of 159, which is much higher than the maximum recommended value of 10 (Kennedy, 1992), with 116,111 being the number of condition. To solve this multicollinearity problem, García et al. (2020) propose the use of orthogonal regression, estimating the following auxiliary model:

$$DISTC_i = \alpha_1 + \alpha_2 COORDXM_i + \alpha_3 COORDYM_i + u_i$$

This methodology to solve the multicollinearity problem has been applied in models 2, 3 and 4. The residuals of the auxiliary model collect the information provided by DISTC that COORDXM and CORDYM do not explain. Model 2 (Mod2- Ort, Table 5) includes these residuals as an explanatory variable, which has been called DISTCRES, representing the effect of the distance to the city center, which is not determined by the geographic coordinates of the dwellings. This model no longer presents serious multicollinearity problems since the highest IVF of this model is 3.124, which corresponds to the DCTROCCIAL variable, with a condition number of 23.565. For this reason, the orthogonalized variable DISTCRES will be maintained in subsequent models. Furthermore, this model complies with the normality assumption since the Jarque-Bera test statistic (which asymptotically follows a chi-square distribution with 2 degrees of freedom) takes the value of 1.093, corresponding to a p-value of 0.579. However, this model presents problems of spatial autocorrelation in the disturbances (Moran's I of the errors = 0.246, p-value < 0.000), and of heteroscedasticity (Breusch-Pagan test = 57.938, p-value < 0.000).

We followed the criteria proposed by Anselin et al. (1996) to determine whether a SAR or SEM model is more appropriate. The LMlag, RLMlag, LM-Lerr and RLMerr statistics were obtained for the analysis of the presence of spatial autocorrelation. The results of the LM tests (Table 6) show the spatial dependence in terms of the error, so the SEM specification was chosen. The results of the SEM model estimates are shown in Table 5 (Mod3-SEM1). However, the issue of heteroscedasticity persists in this model (Breusch-Pagan test = 59.351, p-value < 0.000). In this situation, Kelejian & Purcha (1998) propose applying two-stage least squares to control heteroskedasticity, alternating the generalized moments method and the instrumental variables method.

Table 5 shows the estimates of the SEM model controlling for the heteroscedasticity issue. The results of this model indicate that all the structural variables are significant at 95%, except for the PISCINA variable. It is also observed that the location variables of distances to different points of interest (such as universities, supermarkets or shopping centers) are not significant. This is an expected result, considering that these types of points of interest are distributed throughout the city, so they are located near expensive housing (located in the center of the city) and near more affordable housing (located in areas farther away from the center) (see Figure 6). On the other hand, variables that account for the effect of large-scale price variation, such as the coordinates of the dwellings or distance to the city center, are significant. This allows us to contrast the effect noted above, that there is a monotonic decreasing effect in housing prices from the coast to the outskirts. It is also observed that the coefficient λ is significantly different from zero, which indicates that housing prices depend not only on the variables specified in this model but also on other location variables not included.

Table 5. Mod0-OLS: model 0 before kriging. Mod1-OLS: model 1 calculated by OLS after kriging. Mod2-Ort: orthogonal regression model. Mod3-SEM1: SEM model. Mod4-SEM2: SEM model controlling for heteroscedasticity

	Mod0-OLS	**Mod1- OLS**	**Mod2-Ort**	**Mod3-SEM1**	**Mod4-SEM2**
Constant	6.4195***	6.259***	4.565***	4.561***	4.598***
Construction. Variable					
GARAJE	0.1745**	0.214***	0.214***	0.210 ***	0.221***
HABITAC	0.0876**	0.076***	0.076***	0.080***	0.067**
BANOS	0.1386***	0.125***	0.125***	0.115***	0.122***
M2	0.0020***	0.002***	0.002***	0.002***	0.002***
PISCINA	0.0562	0.049	0.049	0.064	0.067
ANTIG	−0.0037	-0.004 **	-0.004**	-0.004*	-0.005**
Location Variable					
DUNIV	0.0000	-0.062	-0.062	-0.070	-0.074
DSUPERMDO	0.0000	0.004	0.004	0.007	0.009
DCTROCCIAL	−0.0001	-0.098*	-0.098*	-0.054	-0.064
COORDXM	0.00045***	0.00045***	-0.00004*	-0.00005**	-0.00004*
COORDYM	−0.00014*	-0.00014*	0.00010***	0.00010***	0.00012***
DISTC	−0.0006***	-0.00058***			
DISTCRES			-0.00058***	-0.00072***	-0.00064***
λ				0.4250***	0.2150**
R Square	0.760	0.819	0.819	0.838	0.840
N	142	211	211	211	211

Note: "" not significant, "" significant at 10%, "**" significant at 5%, "***" significant at 1%*

Table 6. Lm statistics

Statistics	
LMerr	24.634 (0.000)
RLMerr	6.774 (0.009)
LMlag	20.182 (0.000)
RLMlag	2.322 (0.128)

Note: p-values between -brackets

4.5 Conclusions

It is important to mention that although the starting point was a sample with incomplete information due to the lack of information on the age of 69 dwellings, it was possible to extend the sample and improve the quality of the model and the significance of the age variable (H1) thanks to the spatial estimate of the age of the dwelling, using the Kriging method. In addition, this work has included location variables. Among these, the distance to the center and the spatial trend were highly significant (H2).

It was also found that some location factors, such as distance to certain points of interest, were not significant in this case (H3), probably because these points of interest are located either in areas with higher prices or with lower prices.

In addition, the problem of heteroscedasticity was controlled by applying the generalized moments method and instrumental variables. All structural variables were also significant at 95%, except PISCINA. It was also found that the Mod4-SEM2 model obtained better results than the conventional hedonic price model. This model showed advantages because it considers the interaction between the characteristics of the dwelling and its location. Therefore, according to the results, when the data obey a heterogeneous spatial structure, spatial econometric models are preferable, which have been fundamental in the study of housing price estimates within the hedonic price approach.

The results obtained in this study may be relevant for the real estate or construction sector because these results provide information on the most influential attributes of housing prices. These include, for example, construction factors such as surface area, the existence of a garage, number of bathrooms, age, and number of rooms or location factors such as distance to the city center or the decreasing spatial trend of prices from the coast to the outskirts of the city. The study also shows that incorporating missing information improves the significance of some variables in the model, for example, the age of the dwellings. The results obtained can serve as a reference for the city appraisers when assigning a price to a property. These results can also be useful for developing social real estate projects related to the District Mayor's Office.

As future lines of work, we propose the geostatistical modeling of housing prices considering the above variables and complementing it with the addition of variables with neighborhood characteristics, such as proximity to parks, recreational areas and schools to determine whether they would be significant and to what extent they would improve the level of significance of the model. Therefore, it can be considered that housing prices not only depend on the variables specified in this model but also on other location variables that are not included in the model, such as places of interest in the city, which could have an impact on the price.

4.6 References

Anselin, L. (1988). *Spatial Econometrics: Methods and Models* (Vol. 4). Springer Netherlands. https://doi.org/10.1007/978-94-015-7799-1

Anselin, L. (1992). Space and applied econometrics: Introduction. *Regional Science and Urban Economics*, *22*(3), 307–316. https://doi.org/10.1016/0166-0462(92)90031-U

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, *26*(1), 77–104. https://doi.org/10.1016/0166-0462(95)02111-6

Anselin, L., & Rey, S. J. (2012). Spatial econometrics in an age of CyberGIScience. *International Journal of Geographical Information Science*, *26*(12), 2211–2226. https://doi.org/10.1080/13658816.2012.664276

Anselin, L., Sridharan, S., & Gholston, S. (2007). Using Exploratory Spatial Data Analysis to Leverage Social Indicator Databases: The Discovery of Interesting Patterns. *Social Indicators Research*, *82*(2), 287–309. https://doi.org/10.1007/s11205-006-9034-x

Arraiz, I., Drukker, D. M., Kelejian, H. H., & Prucha, I. R. (2010). A Spatial Cliff-Ord-Type Model with Heteroskedastic Innovations: Small and Large Sample Results. *Journal of Regional Science*, *50*(2), 592–614. https://doi.org/10.1111/j.1467-9787.2009.00618.x

Baronio, A., Vianco, A., & Rabanal, C. (2018). *UNA INTRODUCCIÓN A LA ECONOMETRÍA ESPACIAL*. https://www.researchgate.net/publication/328812646_UNA_INTRODUCCION_A_LA_ECONOMETRIA_ESPACIAL

Basu, S., & Thibodeau, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics*, *17*(1), 61–85. https://doi.org/10.1023/A:1007703229507

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. https://doi.org/10.1006/juec.2001.2214

Can, A. (1990). The Measurement of Neighborhood Dynamics in Urban House Prices. *Economic Geography*, *66*(3), 254–272. https://doi.org/10.2307/143400

Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, *22*(3), 453–474. https://doi.org/10.1016/0166-0462(92)90039-4

Can, A. (1998). GIS and Spatial Analysis of Housing and Mortgage Markets. *Journal of Housing Research*, *9*(1), 61–86. https://doi.org/10.1080/10835547.1998.12091927

Case, B., Clapp, J., Dubin, R., & Rodriguez, M. (2004). Modeling Spatial and Temporal House Price Patterns: A Comparison of Four Models. *The Journal of Real Estate Finance and Economics*, *29*(2), 167–191. https://doi.org/10.1023/B:REAL.0000035309.60607.53

Chay, K. Y., & Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, *113*(2), 376–424. https://doi.org/10.1086/427462

Chica Olmo, J. (1994). *Teoría de las variables regionalizadas. Aplicación en economía espacial y valoración inmobiliaria*. Universidad de Granada. https://editorial.ugr.es/libro/teoria-de-las-variables-regionalizadas_138704/

Chica Olmo, J. (1995). Spatial Estimation of Housing Prices and Locational Rents. *Urban Studies*, *32*(8), 1331–1344. https://doi.org/10.1080/00420989550012492

Chica-Olmo, J. (2007). Prediction of Housing Location Price by a Multivariate Spatial Method: Cokriging. *Journal of Real Estate Research*, *29*(1), 91–114. https://doi.org/10.1080/10835547.2007.12091188

Chica-Olmo, J., Cano-Guervos, R., & Tamaris-Turizo, I. (2019). Determination of buffer zone for negative externalities: Effect on housing prices. *The Geographical Journal*, *185*(2), 222–236. https://doi.org/10.1111/geoj.12289

Chien, L.-C., Yu, H.-L., & Schootman, M. (2013). Efficient mapping and geographic disparities in breast cancer mortality at the county-level by race and age in the U.S. *Spatial and Spatio-Temporal Epidemiology*, *5*, 27–37. https://doi.org/10.1016/j.sste.2013.03.002

Chinloy, P. T. (1977). Hedonic price and depreciation indexes for residential housing: A longitudinal approach. *Journal of Urban Economics*, *4*(4), 469–482. https://doi.org/10.1016/0094-1190(77)90007-9

Clapp, J. M. (2004). A Semiparametric Method for Estimating Local House Price Indices. *Real Estate Economics*, *32*(1), 127–160. https://doi.org/10.1111/j.1080-8620.2004.00086.x

Cliff, A. D., & Ord, J. K. (1981). *Spatial Processes: Models & Applications*. Pion.

Conroy, S. J., & Milosch, J. L. (2011). An Estimation of the Coastal Premium for Residential Housing Prices in San Diego County. *The Journal of Real Estate Finance and Economics*, *42*(2), 211–228. https://doi.org/10.1007/s11146-009-9195-x

Corrado, L., & Fingleton, B. (2012). Where Is the Economics in Spatial Econometrics? *Journal of Regional Science*, *52*(2), 210–239. https://doi.org/10.1111/j.1467-9787.2011.00726.x

Cruces, G., Ham, A., & Tetaz, M. (2008). Quality of Life in Buenos Aires Neighborhoods: Hedonic Price Regressions and the Life Satisfaction Approach. *Inter-American Development Bank*, Working Paper No. R-559. https://doi.org/10.2139/ssrn.1815924

D'Elia, V. V., Grand, M. C., & León, S. (2020). Bus rapid transit and property values in Buenos Aires: Combined spatial hedonic pricing and propensity score techniques. *Research in Transportation Economics*, *80*, 100814. https://doi.org/10.1016/j.retrec.2020.100814

Dubin, R. A. (1992). Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics*, *22*(3), 433–452. https://doi.org/10.1016/0166-0462(92)90038-3

Felus, Y. A. (2001). *New methods for spatial statistics in geographic information systems* [Doctoral dissertation, The Ohio State University]. https://etd.ohiolink.edu/acprod/odb_etd/r/etd/search/10?p10_accession_num=osu1486400446373917&clear=10&session=114911163308190

Fierro, K. P., Fullerton, T. M., & Donjuan-Callejo, K. E. (2009). Housing Attribute Preferences in a Northern Mexico Metropolitan Economy. *Atlantic Economic Journal*, *37*(2), 159–172. https://doi.org/10.1007/s11293-009-9174-x

García, C. B., Salmerón, R., García, C., & García, J. (2020). Residualization: Justification, properties and application. *Journal of Applied Statistics*, *47*(11), 1990–2010. https://doi.org/10.1080/02664763.2019.1701638

Garza, N., & Ovalle, M. C. (2019). Tourism and housing prices in Santa Marta, Colombia: Spatial determinants and interactions. *Habitat International*, *87*, 36–43. https://doi.org/10.1016/j.habitatint.2019.04.001

Geng, B., Bao, H., & Liang, Y. (2015). A study of the effect of a high-speed rail station on spatial variations in housing price based on the hedonic model. *Habitat International*, *49*, 333–339. https://doi.org/10.1016/j.habitatint.2015.06.005

Getis, A. (1995). Cliff, A.D. and Ord, J.K. 1973: Spatial autocorrelation. London: Pion. *Progress in Human Geography*, *19*(2), 245–249. https://doi.org/10.1177/030913259501900205

Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, *5*(4), 471–484. https://doi.org/10.1016/0094-1190(78)90004-9

Goodman, A. C., & Thibodeau, T. G. (2003). Housing market segmentation and hedonic prediction accuracy. *Journal of Housing Economics*, *12*(3), 181–201. https://doi.org/10.1016/S1051-1377(03)00031-7

Gröbel, S., & Thomschke, L. (2018). Hedonic pricing and the spatial structure of housing data – an application to Berlin. *Journal of Property Research*, *35*(3), 185–208. https://doi.org/10.1080/09599916.2018.1510428

Kelejian, H. H., & Prucha, I. R. (1998). A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *The Journal of Real Estate Finance and Economics*, *17*(1), 99–121. https://doi.org/10.1023/A:1007707430416

Kelejian, H. H., & Prucha, I. R. (1999). A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review*, *40*(2), 509–533. https://doi.org/10.1111/1468-2354.00027

Kelejian, H. H., & Prucha, I. R. (2010). Spatial models with spatially lagged dependent variables and incomplete data. *Journal of Geographical Systems*, *12*(3), 241–257. https://doi.org/10.1007/s10109-010-0109-5

Kennedy, P. E. (1992). *A Guide To Econometrics*. MIT Press.

Kilibarda, M., Hengl, T., Heuvelink, G. B. M., Gräler, B., Pebesma, E., Perčec Tadić, M., & Bajat, B. (2014). Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *Journal of Geophysical Research: Atmospheres*, *119*(5), 2294–2313. https://doi.org/10.1002/2013JD020803

LeSage, J. P., & Pace, R. K. (2004). Models for Spatially Dependent Missing Data. *The Journal of Real Estate Finance and Economics*, *29*(2), 233–254. https://doi.org/10.1023/B:REAL.0000035312.82241.e4

Li, H., Wei, Y. D., Wu, Y., & Tian, G. (2019). Analyzing housing prices in Shanghai with open data: Amenity, accessibility and urban structure. *Cities*, *91*, 165–179. https://doi.org/10.1016/j.cities.2018.11.016

Li, H., Wei, Y. D., Yu, Z., & Tian, G. (2016). Amenity, accessibility and housing values in metropolitan USA: A study of Salt Lake County, Utah. *Cities*, *59*, 113–125. https://doi.org/10.1016/j.cities.2016.07.001

Li, R. Y. M., & Li, H. C. Y. (2018). Have Housing Prices Gone with the Smelly Wind? Big Data Analysis on Landfill in Hong Kong. *Sustainability*, *10*(2), 341. https://doi.org/10.3390/su10020341

Li, Y., He, L., & Jiao, J. (2012). A spatial analysis of housing prices in Chinese coastal cities, a case study of the city of Xiamen, China. *The 6th International Association for China*

Planning Conference (IACP) Proceedings, 144–148. https://doi.org/10.1109/IACP.2012.6342973

Li, Y., He, L., Xu, W., Wang, H., & He, Z. (2013). Using GIS and Hedonic in the modelling of spatial variation of housing price in Xiamen city. *International Review for Spatial Planning and Sustainable Development*, *1*(4), 29–42. https://doi.org/10.14246/irspsd.1.4_29

Mark, J. H. (1980). A Preference Approach to Measuring the Impact of Environmental Externalities. *Land Economics*, *56*(1), 103–116. https://doi.org/10.2307/3145833

Montero-Lorenzo, J.-M., Larraz-Iribas, B., & Páez, A. (2009). Estimating commercial property prices: An application of cokriging with housing prices as ancillary information. *Journal of Geographical Systems*, *11*(4), 407–425. https://doi.org/10.1007/s10109-009-0095-7

Morais, M. da P., & Cruz, B. de O. (2015). Demand for Housing and Urban Services in Brazil: A Hedonic Approach. *Repository of the Brazilian Institute for Applied Economic Research*. https://repositorio.ipea.gov.br/handle/11058/4968

Muñoz-Raskin, R. (2010). Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia. *Transport Policy*, *17*(2), 72–84. https://doi.org/10.1016/j.tranpol.2009.11.002

Nelson, J. P. (1978). Residential choice, hedonic prices, and the demand for urban air quality. *Journal of Urban Economics*, *5*(3), 357–369. https://doi.org/10.1016/0094-1190(78)90016-5

Pace, R. K., & LeSage, J. P. (2004). Spatial Statistics and Real Estate. *The Journal of Real Estate Finance and Economics*, *29*(2), 147–148. https://doi.org/10.1023/B:REAL.0000035307.99686.fb

Poeta, S., Gerhardt, T., & Stumpf Gonzalez, M. (2019). Hedonic price analysis of single-family housing. *Revista Ingeniería de Construcción*, *34*(2), 215–220. https://doi.org/10.4067/S0718-50732019000200215

Rehák, Š., & Káčer, M. (2019). Estimating price gradient in Bratislava with different distance measurements. *Journal of European Real Estate Research*, *12*(2), 190–206. https://doi.org/10.1108/JERER-10-2018-0045

Rodríguez, D. A., & Targa, F. (2004). Value of accessibility to Bogotá's bus rapid transit system. *Transport Reviews*, *24*(5), 587–610. https://doi.org/10.1080/0144164042000195081

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Sagner T., A. (2011). Determinantes del precio de viviendas en la región metropolitana de Chile. *El trimestre económico*, *78*(312), 813–839.

Sah, V., Conroy, S. J., & Narwold, A. (2016). Estimating School Proximity Effects on Housing Prices: The Importance of Robust Spatial Controls in Hedonic Estimations. *The Journal of Real Estate Finance and Economics*, *53*(1), 50–76. https://doi.org/10.1007/s11146-015-9520-5

Simlai, P. (2014). Estimation of variance of housing prices using spatial conditional heteroskedasticity (SARCH) model with an application to Boston housing price data. *The Quarterly Review of Economics and Finance*, *54*(1), 17–30. https://doi.org/10.1016/j.qref.2013.07.001

Tiwari, P., & Parikh, J. (1998). Affordability, Housing Demand and Housing Policy in Urban India. *Urban Studies*, *35*(11), 2111–2129. https://doi.org/10.1080/0042098984033

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234–240. https://doi.org/10.2307/143141

Trojanek, R., & Gluszak, M. (2018). Spatial and time effect of subway on property prices. *Journal of Housing and the Built Environment*, *33*(2), 359–384. https://doi.org/10.1007/s10901-017-9569-y

Urrea Zora, M. A., & Cárdenas Mazo, N. (2019). *Un modelo espacial de precios hedónicos y calidad de vecindario: Una aplicación para Medellín, Colombia* [BachelorThesis, Universidad EAFIT]. http://repository.eafit.edu.co/handle/10784/14941

von Graevenitz, K., & Panduro, T. E. (2015). An Alternative to the Standard Spatial Econometric Approaches in Hedonic House Price Models. *Land Economics*, *91*(2), 386–409. https://doi.org/10.3368/le.91.2.386

Wen, H., Jin, Y., & Zhang, L. (2017). Spatial heterogeneity in implicit housing prices: Evidence from Hangzhou, China. *International Journal of Strategic Property Management*, *21*(1), 15–28. https://doi.org/10.3846/1648715X.2016.1247021

Xiao, Y., Chen, X., Li, Q., Yu, X., Chen, J., & Guo, J. (2017). Exploring Determinants of Housing Prices in Beijing: An Enhanced Hedonic Regression with Open Access POI Data. *ISPRS International Journal of Geo-Information*, *6*(11), 358. https://doi.org/10.3390/ijgi6110358

Yoo, E.-H., & Kyriakidis, P. C. (2009). Area-to-point Kriging in spatial hedonic pricing models. *Journal of Geographical Systems*, *11*(4), 381–406. https://doi.org/10.1007/s10109-009-0090-z

Zambrano-Monserrate, M. A. (2016). Formación de los precios de alquiler de viviendas en Machala (Ecuador): Análisis mediante el método de precios hedónicos. *Cuadernos de Economía*, *39*(109), 12–22. https://doi.org/10.1016/j.cesjef.2015.10.002

Zhang, L., Zhou, J., Hui, E. C. M., & Wen, H. (2019). The effects of a shopping mall on housing prices: A case study in Hangzhou. *International Journal of Strategic Property Management*, *23*(1), 65–80. https://doi.org/10.3846/ijspm.2019.6360

Zhang, Y., Zheng, S., & Sun, C. (2014). What's the Effect of Urban Villages on Commercial Housing Price? An Analysis Based on Second-Hand Housing Transactions in Beijing. In J. Wang, Z. Ding, L. Zou, & J. Zuo (Eds.), *Proceedings of the 17th International Symposium on Advancement of Construction Management and Real Estate* (pp. 173–183). Springer. https://doi.org/10.1007/978-3-642-35548-6_19

Zhu, B., Füss, R., & Rottke, N. B. (2011). The Predictive Power of Anisotropic Spatial Correlation Modeling in Housing Prices. *The Journal of Real Estate Finance and Economics*, *42*(4), 542–565. https://doi.org/10.1007/s11146-009-9209-8

CAPÍTULO 5. Determination of buffer zone for negative externalities: Effect on housing prices

---

Jorge Chica-Olmo[1]
Rafael Cano-Guervos[1]
Ivan Tamaris-Turizo[2]

[1]*Department of Quantitative Methods for Economics and Business, University of Granada, Granada, Spain.*
[2]*Budget Management Service, University of Magdalena, Magdalena, Colombia*

Abstract

Numerous works have addressed the problem of the hedonic modelling of housing prices, in which structural and locational variables of dwellings are considered. Among the latter, it is common to include variables related to transport accessibility. However, few studies have considered the effects of proximity to railway lines, and even less so freight railroads. This study determines the size of the width of influence (buffer zone) of negative externalities associated with the freight train line in the city of Santa Marta, Colombia. Moreover, the effect of this width on housing prices is studied. A hedonic model was estimated using spatial econometrics and geostatistical methods. Geostatistical techniques were used to obtain an isovalues map for the price of a standard dwelling. The study is of interest to the fields of real estate, territorial planning and transport systems, among others.

KEYWORDS: buffer zone, freight trains, geostatistics, housing prices, Santa Marta (Colombia), spatial econometrics

5.1 Introduction

Hedonic price models are currently the main method used to determine the price of housing and to quantify the implicit price of each of the characteristics that influence housing prices (Wen et al., 2017). This method is used to make real estate valuations and determine the characteristics that affect property prices (Rosen, 1974). The classical statistical method of ordinary least squares (OLS) does not consider the presence of spatial autocorrelation (Anselin, 1988). There are two approaches in spatial hedonic modelling: spatial econometrics (Can, 1992) and the geostatistical approach (Chica-Olmo, 1995; Dubin, 1992).

According to Can (1992), the characteristics that determine the price of housing can be classified into two groups: structural and locational. The structural characteristics include the age of the dwelling, the number of bathrooms and other features, while the locational characteristics include accessibility to the central business district (CBD) and neighbourhood characteristics.

Accessibility refers to the proximity of services demanded by citizens, such as schools, hospitals, commercial establishments, leisure facilities, green areas, urban transport, CBD and others. Some studies have considered the effect on housing price of accessibility to different amenities and measure accessibility in terms of distance (Chen & Jim, 2010). Fan et al. (2016), for example, studied distances to different amenities to estimate the marginal willingness to pay for an additional metre proximate to open space amenities. However, the accessibility between two locations can be measured in different ways, including time, cost, mode of transportation and physical distance (Jang & Kang, 2015), although it is usually measured in terms of distance or time. In this regard, Apparicio et al. (2003) compared physical distance and time and concluded that, at the city level, both are highly correlated. Physical distance has also been used in several studies as a proxy for spatial proximity (Jang & Kang, 2015).

Several studies have analyzed locational or neighbourhood factors. A classic classification can be found in the work of Chau & Chin (2003) and Li et al. (2016). Some of these factors have negative effects, such as air pollution (Li et al., 2016), air and noise pollution (Chasco & Gallo, 2013; Le Boennec & Salladarré, 2017), water pollution (Chen & Li, 2018), airport noise (Cohen & Coughlin, 2008; Pope, 2008), traffic noise (Wilhelmsson, 2000) and crime rates (Tita et al., 2006). Others, however, have positive effects, such as schools (Kane et al., 2006), shopping centres (Jang & Kang, 2015), beaches (Conroy & Milosch, 2011), lakes (Wen et al., 2014), parks (Wu et al., 2017), sports facilities (Feng & Humphreys, 2008) and proximity to air transport infrastructure (McMillen, 2004), a highway and a bus stop (Fan et al., 2016).

One of the least studied locational factors in the literature on housing prices is proximity to the rail transport system. This externality can have positive or negative effects. Therefore, one aspect of interest is to determine the sign of these effects and the area of influence or buffer zone of this externality.

Geographical information systems (GIS) employ different tools to analyse spatial data. One of the most useful tools is the creation of buffer zones around objects such as points, lines or polygons (Longley et al., 2001). These buffer zones determine a space that represents the area of influence of the objects. To obtain the buffer zones, it is necessary to determine initially the distance or radius of influence, which is usually arbitrary (Perchoux et al., 2016). As in this study, Diao et al. (2016) and Xu & Zhang (2016) used a buffer-ring-based approach to identify the treatment zone of a rail line. Our study aims to determine objectively the radius of influence in order to maximise the explanatory capacity of the model and quantify the negative externalities caused by the proximity of dwellings to a freight railroad.

The hedonic model was estimated from both a spatial econometrics and geostatistics perspective. This is a novel contribution of the study, since this twofold approach has largely been neglected in the literature on the subject (Case et al., 2004; Militino et al., 2004). In addition, multivariate and univariate geostatistical methods were used, respectively, to obtain the hedonic model and estimate the missing data for the explanatory variable age of dwelling, whose estimates are used in the hedonic model.

This work is structured as follows. A literature review on rail externalities that influence housing prices is provided in the next section. Following this, a brief overview of the methodology used in the study is presented. The main results of applying these methods to housing prices in the city of Santa Marta, Colombia, are then discussed. Finally, some conclusions are drawn.

5.2 Background

Studies on externalities associated with proximity to rail transport infrastructure have obtained ambiguous results. Kahn (2007), for example, found that the construction of a new railway station results in the positive gentrification of communities. In a similar vein, public investment in the Washington DC Metro caused property values to rise around station site areas (Grass, 1992). In the literature review of Wardrip (2011), it was concluded that although access to public transport can lead to higher housing prices, the construction of a new transit line in a real estate market with high demand may not cause the same increase. Gadziński & Radzimski (2016) found that although 20% of respondents are willing to pay a higher price for homes located near to train stations, their spatial econometrics model shows a weak relation between proximity to stations and apartment prices. Using the difference in differences approach, Diao et al. (2017) found that the opening of new lines increased housing values within 600 m of the stations. Moreover, they used a local polynomial regression-based approach to identify the treatment zone of subway stations in Singapore. Other studies have accounted for spatial autocorrelation in assessing the price premium of rail infrastructure from different approaches (Diao, 2015; Xu & Zhang, 2016).

However, Kilpatrick et al. (2007) detected negative externalities on the price of homes close to railway lines, but without direct access to stations. Wardrip (2011) highlighted the negative effects of the noise, traffic and pollution associated with transit systems on houses near transit lines. Dai et al. (2016) compared the negative effects of proximity of housing to a non-transfer station and a transfer station, concluding that these effects depend on whether the dwellings are located in suburbs or in the city. Perceptions of crime and excessive nuisance related to rail transit stations also negatively affect residential property values (Forouhar, 2016). Other authors have also considered the influence of environmental aesthetics on housing prices (Cetintahra & Cubukcu, 2015).

Therefore, there does not exist a categorical relationship between housing prices and proximity to railway infrastructure. This ambiguity can be observed in work such as that by Banister & Thurstain-Goodwin (2011) and Forouhar (2016) in this regard, Dubé et al. (2013) noted that of 19 studies that have examined this relationship, a positive association was observed in 13 cases, a negative association in seven and a neutral association in four.

Although the effect of accessibility to rail transport on house prices has been addressed in the literature, most studies focus on accessibility to railway stations. These studies find a more or less positive effect (Bowes & Ihlanfeldt, 2001; Debrezion et al., 2007; Gadziński & Radzimski, 2016; Guerra et al., 2012; Weinberger, 2001), but some negative effects due to noise nuisance have also been reported (Debrezion et al., 2011). However, these last authors emphasise the scarcity of literature on railway line accessibility. The study quantifies the negative effect of proximity to the railway line, but the distance of the influence of this negative externality is not determined. Diao et al. (2016) tested the existence of negative externalities associated with train noise for houses located less than 400 m from the line (threshold) and found that the removal of train noise externalities increases housing prices by 13.7%. In contrast, Strand & Vågnes (2001) observed that property prices in Norway increase by about 10% within a distance of 100 m from railway lines.

This literature review points to the scarcity of studies that address the effect of proximity to railway lines on housing prices, particularly freight railroads and the quantification of their area of influence, in contrast to the relative abundance of studies on proximity to railway stations, thus justifying the interest of this paper. We are interested in analysing the negative externalities of freight trains on housing prices in Santa Marta, Colombia, and in determining the area of influence of the railway lines (the buffer zone).

5.3 Study area and methods

Santa Marta is one of the oldest cities in South America and is located on the Caribbean coast. The municipality spans an area of 2,390 km2, of which only 36.6 km2 correspond to the urban area of the city considered in this study. According to the National Administrative Department of Statistics (DANE), the capital had a population of approximately 491,000 inhabitants in 2016. The city is divided into nine districts with some 300 neighbourhoods, in which 75.7% of the dwellings are single-family homes according to the last general census conducted by DANE in 2005. There has been a surge in real estate development in Santa Marta in recent years, although the city has mainly expanded on the coast (Ramos et al., 2015).

One of the main urban characteristics of Santa Marta is that the city was built based on a monocentric model with a CBD located next to the bay, which also has many tourism resources. The climate is warm, with an average relative humidity of 77% and temperatures ranging from 23 to 32°C throughout the year. The city has become an important international tourist destination due to its climate, resources and privileged location on the route to areas of natural beauty, such as the Tayrona National Park and the Sierra Nevada de Santa Marta. Currently, it is a magnet for investors who want to take advantage of the city's potential for the development of major projects.

5.4 Data

To carry out this study, primary information on second-hand single-family homes for sale in the first quarter of 2016 was used. In this quarter, according to the Register of Real Property, the total number of properties sold (single-family homes, apartments, offices, etc.) was about 1,400. Information on second-hand single-family homes was obtained from the websites of major real estate agencies operating in Santa Marta, such as Coldwell Banker, Metrocuadrado and Safe Bienes y Raíces. In the three-month period considered in these real estate portals, an average of approximately 900 single-family homes was on the market, some of which were offered on several websites simultaneously. Therefore, the sample represents approximately 30% of the homes for sale in the first quarter of 2016. It should be noted that availability of information on real estate sales in Santa Marta, and in general in Colombia, is heterogeneous since not all of these sources provide the same type of information. Safe Bienes y Raíces, for example, did not provide the age of the home. The final sample consisted of 270 second-hand homes, but complete data were only available for 201. There was missing information about the age of 69 houses, as it was not provided by any of the real estate companies. To solve this problem, it was necessary to perform a spatial interpolation using the kriging method with the variable age for the 201 houses for which the value of the variable was known.

The kriging method allowed estimation of this variable for the 69 houses for which this value was not available. Figure 8 shows the location of the 201 houses, their age and the choropleth map of age estimated by the kriging method.

Figure 8. Location of the 201 houses with complete data, age of the houses (quintiles) and choropleth map of estimated age (kriging age).

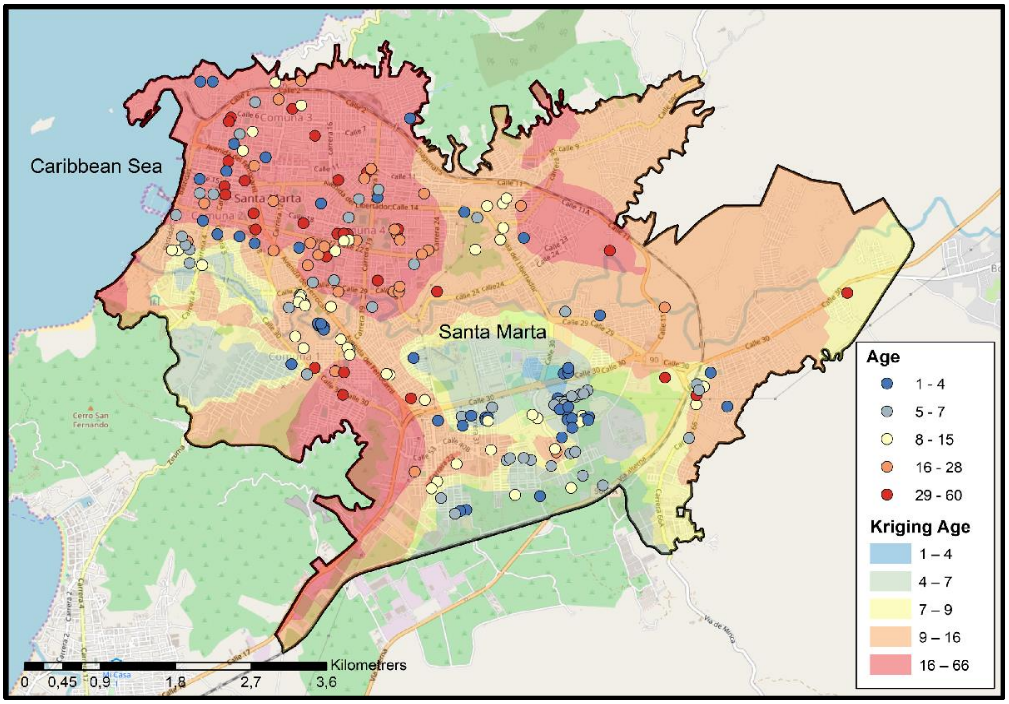
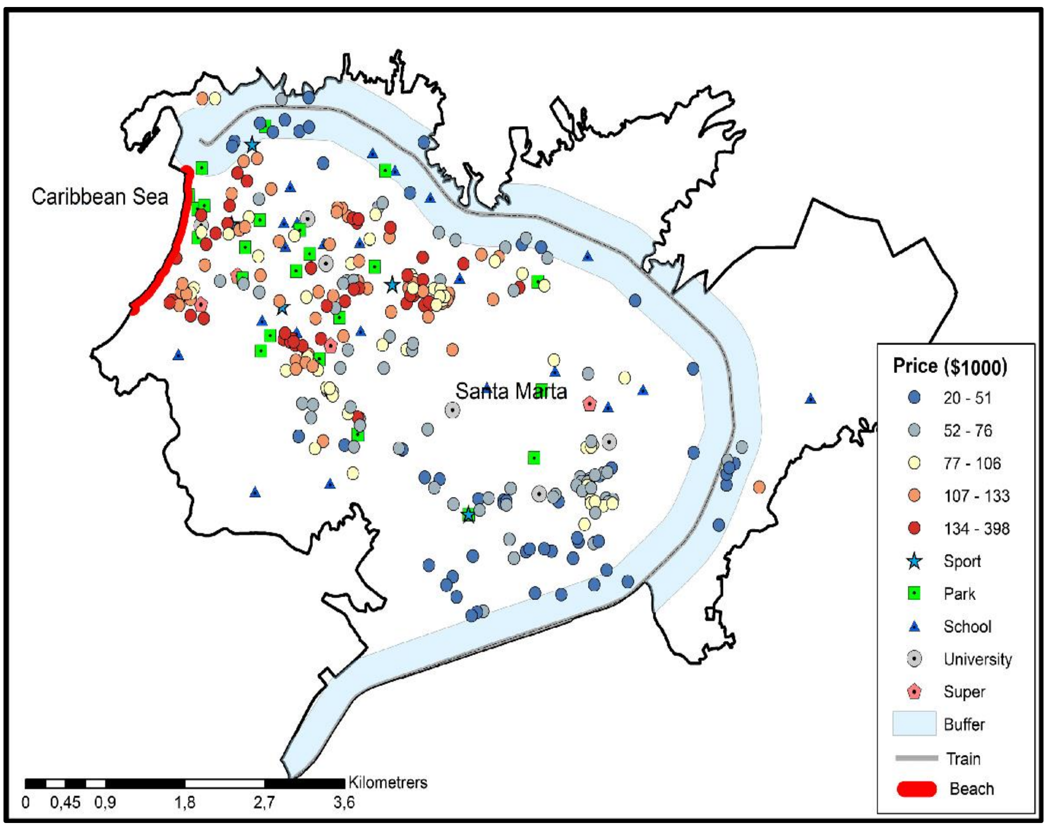


Figure 9 shows the location of the 270 houses that comprise the final sample and their price in US dollars. As can be observed, homes are for sale throughout the entire city. In addition, this figure shows the location of each of the amenities considered in this work, which are typically included in the hedonic housing price regression model (sports facilities, parks, schools, university, supermarket, beach and train line).

The Atlantic Railway Line, which was inaugurated in 1887, borders Santa Marta to the north and east (see Figure 9). This line transports goods to and from the city's port and the main cities in the interior of the country and terminates in the municipality of La Dorada. Initially, the Atlantic Railway was a cargo and passenger line, but currently operates as a freight line only (mainly coal, 40 million tons per year), although there are plans to diversify the containerised cargo to transport other types of product (coffee, fruit, industrial goods, household appliances, paper, plastics, fertilisers, cement, etc.). The freight train is made up of more than a hundred cars and operates day and night, crossing several populated areas without a set schedule. The train is not equipped with safety mechanisms such as lights, crossing signals or bells to warn residents of its arrival. This has resulted in a considerable number of fatalities and accidents, to such a point that it is considered the national railway that has caused the greatest number of deaths (Noticias-Uno, 2013). In addition, these dangers are compounded by the proximity of dwellings, many of which are located just a few metres from the line, causing noise, pollution, cracks in the walls of the residences and other nuisances.

Figure 9. Price in us dollars of the houses in the sample (quintiles). Urban amenities: sports facilities (sport), parks and squares (park), schools (school), universities (university), commercial establishments (super), beach area (beach), train line (train) and buffer zone (buffer).



Although the main goal of this paper is to examine the negative effects of the proximity of the freight train line on house prices (variable Price, in US dollars), structural and locational variables that are typically considered in the literature have been included in the specification of the hedonic model (Can, 1992; Fan et al., 2016; Wu et al., 2017). The structural variables are: Garage (number of garages in the house), Rooms (number of rooms), Baths (number of bathrooms), Surface (surface area of the house, in square metres), Swimmingpool (takes the value of 1 if the house has a swimming pool and 0 otherwise) and Age (age of the house in years). The locational variables are: Dist-University (distance of the house to the nearest college, in kilometres), Dist-School (distance of the house to the nearest school, in kilometres), Dist-Train (distance to the nearest railway line, in kilometres), Buffer-Train (takes the value of 1 if the house is located within the distance of 320 m or less from the railway line and 0 otherwise), Dist-Sport (distance to the nearest sports facility, in kilometres), Dist- Beach (distance to the beach, in kilometres), Dist-Park (distance to the nearest park, in kilometres) and Dist-Super (distance to the nearest shopping centre, in kilometres). All the locational variables are measured using a Euclidean distance between amenity and house. In the case of Dist-Train and Dist-Beach, a linear perpendicular distance is measured. Table 7 shows the descriptive statistics of the variables included in the hedonic models. It should be noted that 10% of the houses are located at a distance of less than or equal to 320 m from the railway line (the numerical value of which is obtained as explained in the Results section), and that the average price of the dwellings is $44,400.

Table 7. Descriptive statistics of variables

Variables	**Mean**	***SD***	**Min**	**Max**
Price	97,157.26	57,678.60	19,920.00	398,400.00
Structural variables:				
Garage	0.874	0.394	0	3
Rooms	3.315	0.800	2	6
Baths	2.948	0.986	1	6
Surface	162.374	103.825	50	670
Swimmingpool	0.144	0.352	0	1
Age	15.504	12.700	1	60

Variables	Mean	*SD*	Min	Max
Locational variables:				
Dist-University	0.872	0.428	0.063	3.260
Dist-School	0.652	0.455	0.026	2.158
Dist-Train	1.183	0.666	0.047	2.695
Buffer-Train	0.100	0.301	0	1
Dist-Sport	0.896	0.670	0.045	4.991
Dist-Beach	2.308	1.876	0.000	6.347
Dist-Park	0.579	0.455	0.058	3.410
Dist-Super	0.650	0.408	0.063	3.390

It is important to highlight that although accessibility to a city's historic district (Fan et al., 2016) or CBD is usually considered in the literature, these two areas coincide in Santa Marta as they are both located next to the beach. For this reason, we have preferred to use accessibility to the beach because it is perfectly demarcated by the shoreline (see red line in Figure 9).

5.5 Methods

As indicated above, there exist two methods to study spatial effects on housing prices: geostatistics and spatial econometrics (Anselin, 1992; Osland, 2010; Pace & LeSage, 1998). The geostatistical approach considers the autocorrelation structure based on the variance–covariance matrix of perturbations and models the matrix using the variogram function, which considers the spatial autocorrelation between the model perturbations and how the correlation decreases as the distance between data increases (Cressie, 1991). In contrast, spatial econometrics uses a W weights matrix, which considers the spatial relationship between the data based on different criteria. The W weights matrix is the cornerstone of the approach since the results obtained are sensitive to the criteria selected to build the matrix (Anselin & Rey, 2012). Hence, it is important to specify the W matrix adequately.

5.5.1 Spatial econometrics

In the present paper, a spatial econometrics model was used to explain housing prices based on information provided by real estate agencies. Classical estimation methods, such as OLS, assume independence between observations (Pace & LeSage, 2009). However, spatial econometric models consider the spatial dependence between observations, which is revealed by the presence of spatial autocorrelation (Anselin, 1988) and can be detected by Moran's I test (Moran, 1950). There are different models to consider geographical location in a spatial econometric model (Anselin, 1988) but traditionally two are used:

$$SAR: y = \rho Wy + X\beta + \varepsilon \qquad (1)$$

$$SEM: y = X\beta + u \qquad (2)$$

$$u = \lambda Wu + \varepsilon$$

where y is a vector (n) of the values of the dependent variable (housing prices); X is a matrix (n×k) with the k explanatory variables of the model that is related to the vector of the parameter β; W is a spatial weights matrix (n×n), wij, which reflects the neighbourhood spatial structure between observations i and j; W_y is a vector representing the spatial lagged dependent variable (endogenous interaction relationships) that is related to the parameter ρ, which represents a spatial spillover effect; W_u represents the spatial lagged perturbations with λ as its associated parameter; and ε denotes normal vector perturbations i.i.d. (independent and identically distributed)

Classically, the elements of the W matrix (wij) take the value of 1 if the geographical data i and j are neighbours and 0 otherwise. In the case of certain data, such as housing, different criteria can be applied to consider the neighbours (Corrado & Fingleton, 2012). For example, neighbours can be considered those that are within a radius or cut-off, or a certain number of nearest neighbours (k nearest; Osland, 2010). Another way of specifying the elements of the W matrix is that the weights of the matrix are equal to the inverse of the distance between houses, which allows taking into account that the relationship between the price of two houses decreases as the distance between them increases.

5.5.2 Geostatistics

Although geostatistical methods were originally used in earth sciences (Cressie, 1990; Matheron, 1970), this approach has been applied in other sciences, including geography. Several methods are used in geostatistics, with the most common one being kriging, which is a univariate method (Cressie, 1991). In this study, we use a multivariate method known as regression kriging (RK), which mixes the regression method with the kriging method (Chica-Olmo, 1995; Dubin, 1992; Hengl et al., 2007). To estimate housing prices at any point in the plane, s0, where the price of housing is not known, we will use the best linear unbiased predictor (BLUP; Christensen, 1987):

$$\hat{y}(S_0) = x(S_0)'\hat{\beta} + \lambda' e(S_i) \qquad (3)$$

where x(s0) is a vector of the known values of the explanatory variables in s0; $\hat{\beta}$is a vector of coefficients estimated by generalised least squares (GLS); λ is a vector of kriging weights and e(si) is the vector of the model residuals in the sample locations, si, whose spatial autocorrelation structure can be characterised by the variogram. An unbiased estimator of the variogram is (Matheron, 1965):

$$\hat{\Upsilon}_e(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [e(S_i + h) - e(S_i)]^2 \qquad (4)$$

where (si+h) and (si) are locations, and N(h) is the number of h distant point pairs. It is necessary to fit a model to the empirical variogram to perform estimations with the kriging method. We used the spherical or Matheron model:

$$\gamma(h) = \begin{cases} C_0 + C\left[\frac{3}{2}\frac{h}{a} - \frac{1}{2}\left(\frac{|h|}{a}\right)^3\right], & |h| > 0 \\ 0 \qquad\qquad\qquad\qquad\qquad\quad , & |h| = 0 \end{cases}$$

The parameters of the fitted model are: the nugget effect (C0), the range (a) and the sill (C0+C), where C is the partial sill. C0 is a measure of spatial continuity, the range is defined by the value of h in which the value of the variogram stabilizes and the sill is the value that the variogram model attains at the range. The spherical model is one of the most frequently used models in geostatistics (Webster & Oliver, 2007). The range of the variogram represents the distance that defines the area of influence of the variable. Spatial dependence can be investigated using the relative nugget effect, that is, the ratio of the nugget effect and the sill. A ratio of less than 25% indicates strong spatial autocorrelation, a ratio of 25%–75% represents moderate spatial autocorrelation, and a ratio greater than 75% indicates weak spatial autocorrelation (Rouhani, 1996).

5.6 Results

In this work we have specified a semi-logarithmic hedonic model which is widely used in the real estate literature (Debrezion et al., 2011). This type of specification normalises the model perturbations and the coefficients of the model (β) can be directly interpreted as semi-elasticities. That is, in the case of a continuous variable β*100 represents the percentage impact of this variable on housing price, ceteris paribus. The dummy variable coefficients cannot be directly interpreted (Halvorsen & Palmquist, 1980). In this case it is necessary to use the expression:

$$[\exp(\beta) - 1] * 100 \qquad (6)$$

In the semi-logarithmic hedonic model specified in this paper, different variables that capture accessibility to amenities from the sample of houses have been considered (university, school, sport, beach, park and supermarket). As indicated, these amenities have a positive effect on housing prices, such that the closer the house is to the amenity, the higher the price of the house or, equivalently, as the distance between the house and the amenity increases, the price of the house decreases continuously. This leads to the presence of negative coefficients for these amenities in the hedonic model. However, in the case of proximity to the freight train line, a negative externality occurs, which can be measured continuously by the distance between the houses and the railway line or by binary variables, considering either a threshold or cut-off distance (Diao et al., 2016) or more than one threshold (piecewise distances; Debrezion et al., 2011). In our case, we first specified Model 1 (Mod1-OLS), which includes the structural and amenity variables. Mod1-OLS includes a continuous measure of distance of houses to the railway line (variable Dist-Train). Table 8 (Mod1-OLS) shows the OLS estimates of the model.

All variables have the expected signs and are significant at the 5% level, with the exception of Dist-Train, Dist- Park and Dist-Super. The variable Dist-Train may not be significant because the negative externalities of the freight railway line (noise, pollution, vibrations, hazards, visual impact, barriers, etc.) on the price of housing decrease significantly from a certain distance.

Therefore, we have considered that this negative effect may have a limited range of influence or threshold. This range of influence demarcates the buffer zone within which the effects of negative externalities have a negative influence on housing prices. Therefore, we have specified Model 2 (Mod2-OLS), which is the same as Model 1 (Mod1-OLS) but replaces the variable Dist-

Train with the binary variable Buffer-Train, which takes the value of 1 if the house is located within a certain radius of influence of the train line and zero otherwise.

To detect this range of influence, OLS was used to estimate Model 2 with different distances ranging from 100 to 700 m in 20–20 m intervals. Figure 10 shows the coefficient of determination for each of these models and the p value corresponding to the variable Buffer-Train. As can be seen, the highest value of the R2 and the lowest p value are reached for the same distance, which is 320 m. Therefore, once this width of influence has been established, it is possible to represent the buffer zone of the train line (see Figure 9). In Table 8, Model 2 (Mod2-OLS) is shown with the variable Buffer-Train set with a threshold of 320 m. The coefficient is negative and significant, thus indicating that the houses within the buffer zone have a lower value than the rest of the houses. It should be noted that serious multicollinearity was not detected in either Model 1 (Mod1-OLS) or Model 2 (Mod2-OLS), since the largest variance inflation factor (VIF) does not exceed the value of 10 in any case (3.963 and 3.767, respectively).

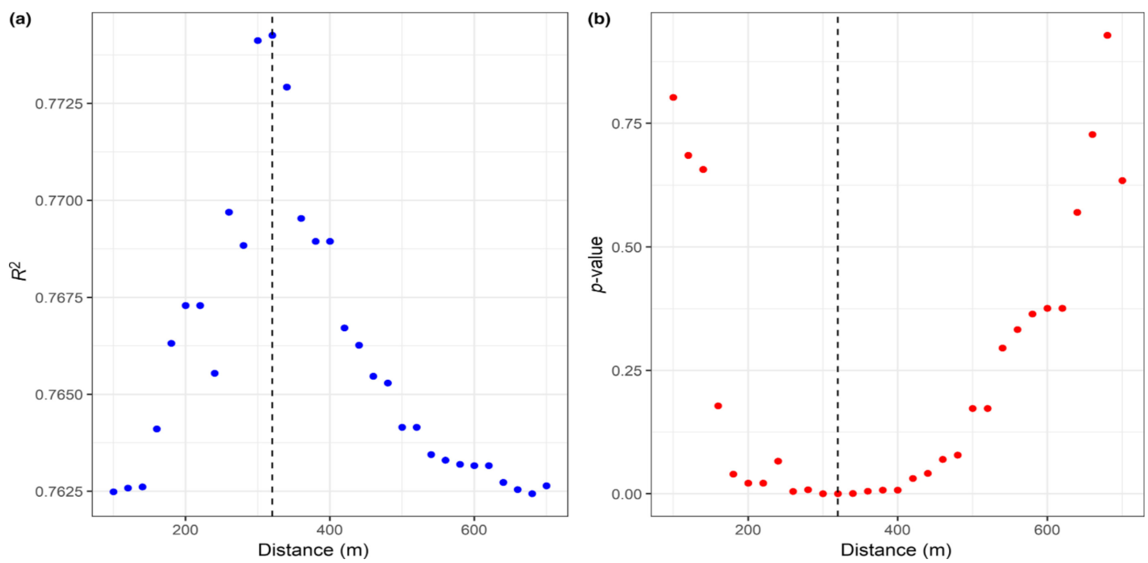
Table 8. Mod1-OLS, Model 1 estimated by OLS with distance to train line. Mod2-OLS, Model 2 estimated by OLS with a threshold of 320 m to the train line.

	Mod1-OLS	**Mod2-OLS**	**Mod2-SAR**	**Mod2-SEM**	**Mod2-RK**
Const.	10.897 (0.000)	10.918 (0.000)	6.412 (0.000)	10.984 (0.000)	10.895 (0.000)
Structural variables:					
Garage	0.117 (0.026)	0.108 (0.035)	0.123 (0.006)	0.126 (0.006)	0.134 (0.003)
Rooms	0.061 (0.038)	0.063 (0.028)	0.043 (0.084)	0.045 (0.059)	0.051 (0.024)
Baths	0.132 (0.000)	0.121 (0.000)	0.117 (0.000)	0.114 (0.000)	0.103 (0.000)
Surface	0.002 (0.000)	0.002 (0.000)	0.002 (0.000)	0.002 (0.000)	0.002 (0.000)
Swimmingpool	0.113 (0.034)	0.103 (0.042)	0.082 (0.063)	0.109 (0.036)	0.086 (0.068)
Age	−0.006 (0.002)	−0.005 (0.002)	−0.004 (0.004)	−0.005 (0.001)	−0.005 (0.001)
Locational variables:					
Dist-University	−0.143 (0.002)	−0.106 (0.020)	−0.082 (0.037)	−0.071 (0.299)	−0.041 (0.610)
Dist-School	−0.367 (0.000)	−0.354 (0.000)	−0.143 (0.007)	−0.349 (0.000)	−0.294 (0.004)
Dist-Train	0.003 (0.912)	—	—	—	—
Buffer-Train	—	−0.239 (0.000)	−0.151 (0.009)	−0.259(0,000)	−0.270 (0.001)
Dist-Sport	−0.154 (0.000)	−0.144 (0.001)	−0.074 (0.053)	−0.193 (0.002)	−0.210 (0.003)

	Mod1-OLS	Mod2-OLS	Mod2-SAR	Mod2-SEM	Mod2-RK
Dist-Beach	−0.023 (0.031)	−0.031 (0.003)	−0.024 (0.008)	−0.021 (0.052)	−0.021 (0.067)
Dist-Park	0.100 (0.174)	0.096 (0.173)	0.068 (0.266)	0.054 (0.600)	0.014 (0.898)
Dist-Super	−0.004 (0.952)	0.031 (0.655)	0.015 (0.804)	0.104 (0.270)	0.164 (0.121)
Wy	—	—	0.386 (0.000)	—	—
Wu	—	—	—	0.522 (0.000)	—
Spatial dependence:					
Moran's *I* error	0.295 (0.000)	0.296 (0.000)	—	—	—
Nugget	—	—	—	—	0.034
Sill	—	—	—	—	0.068
Range	—	—	—	—	1,082.538
Fit model:					
R2	0.762	0.774	0.819	0.825	0.814
AIC	88.310	74.533	26.304	28.362	—
N	270	270	270	270	270

Mod2-SAR and Mod2-SEM are Model 2 estimated with the SAR and SEM specifications and Mod2-RK is Model 2 estimated by RK. p values are shown in paren- theses. AIC is used to compare the relative quality of models.
AIC, Akaike information criterion; OLS, ordinary least squares; RK, regression kriging; SAR, spatial autoregressive model; SEM, spatial error model.

Figure 10. R^2 of Mod2-OLS model (a) and p value of the binary variable buffer-train (b) for different distances from the houses to the train line.



A main issue for a spatial econometric model is to define the spatial connectivity in the W weight matrix (Bivand & Piras, 2015; Dubin, 1998). This work considers different types of spatial connectivity or neighbourhood conceptualisations commonly used in the literature, all of which are based on the location of each house and use coordinates. In particular, we have considered three types of neighbourhood conceptualisation to specify the W matrix: (1) a binary matrix in which the k-nearest neighbour (from two to 10 nearest neighbours in space) are considered neighbours; (2) a binary matrix that considers the threshold or cut-off distance so as to ensure that all observations have at least one neighbour (in our case the cut-off = 1,893.5 m, with an average number of links = 107.192); and (3) the inverse distance and inverse distance squared. In all cases, we standardised the W matrix sum to unity (row standardisation) and used the randomisation method with 999 permutations to perform the spatial autocorrelation tests.

In practice, it is very difficult to select the best W matrix (Anselin et al., 2007). One criterion used in the literature is to compare the value of the Akaike information criterion (AIC) statistic of the spatial models estimated with different W matrices. The model with the lowest AIC value is preferred (Osland, 2010). Another criterion is to select the W matrix for which the highest Moran's I value is reached in combination with a high level of statistical significance (Chi, 2010). We have considered these two criteria, as well as the analysis of the goodness of the estimated spatial model by means of the R2 of the specified spatial model. Table 9 shows Moran's I using different specifications of the W matrix. In all cases, the null hypothesis of no spatial dependence in housing prices is rejected. The highest Moran's I value corresponds to the specification of the W matrix that represents the inverse distance. This specification also has the largest R2 and the smallest AIC.

Table 9. Spatial autocorrelation of housing price (moran's I), R^2 and AIC of SAR model for different conceptualisations of spatial relationships

	***k* nearest**					**Cut-off distance**	**Inv. distance**	
	2	**4**	**6**	**8**	**10**	**1,893.5 m**	**Inv. dist.**	**Inv. dist. squared**
Moran's *I*	0.534	0.515	0.513	0.469	0.430	0.243	0.535	0.453
(*p* value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$R2$	0.801	0.805	0.804	0.807	0.806	0.795	0.814	0.788
AIC	41.605	36.803	37.155	33.583	34.515	50.078	26.304	59.960

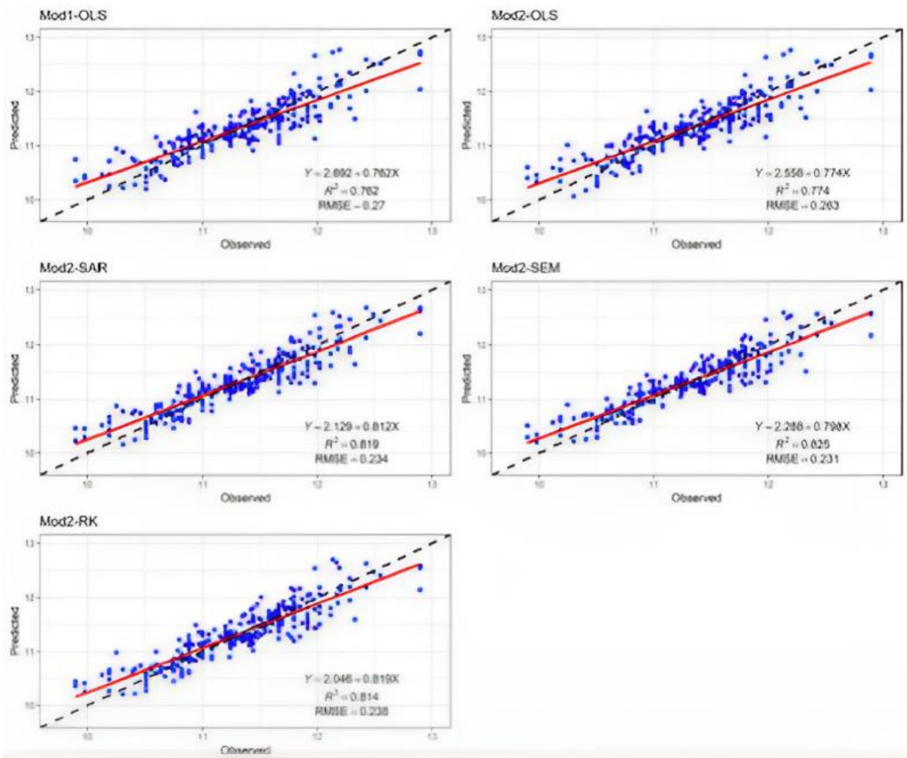
AIC, Akaike information criterion; SAR, spatial autoregressive model.

As can be seen in Table 8, the variables in all the models that are significant at the 5% or 10% level have the expected signs. The table shows the results of the analysis of the presence of spatial autocorrelation in the errors (Moran's I error) in Model 1 (Mod1-OLS) and Model 2 (Mod2-OLS). Since the OLS method provides inefficient estimates in the presence of spatial autocorrelation (Anselin, 1988), two different approaches were used to estimate Model 2: spatial econometrics and geostatistics. In the first case, the SAR (Equation 1) and SEM (Equation 2) specifications of the model were estimated by maximum likelihood, while the RK method was applied in the second case to estimate the parameters using GLS.

To consider the spatial autocorrelation of the model perturbations, we have used a spherical variogram model whose parameters are shown in Table 8. The relative nugget effect of the residual (nugget/sill) is 0.5, thus indicating that spatial dependency is moderate. Furthermore, the difference in housing prices in and outside the impact zone is not only due to the explanatory characteristics considered in the model, but may also be influenced by the omitted variables (Anselin, 1988; Diao et al., 2017). The moderate presence of spatial dependence in the model perturbations could be due to omitted variable problem, such as spatial contagion, which is a non-observable phenomenon. The spatial range of this spatial dependence is approximately 1 km (1,082.538 m), thus indicating the distance at which the effects of these unobservable variables are spatially related.

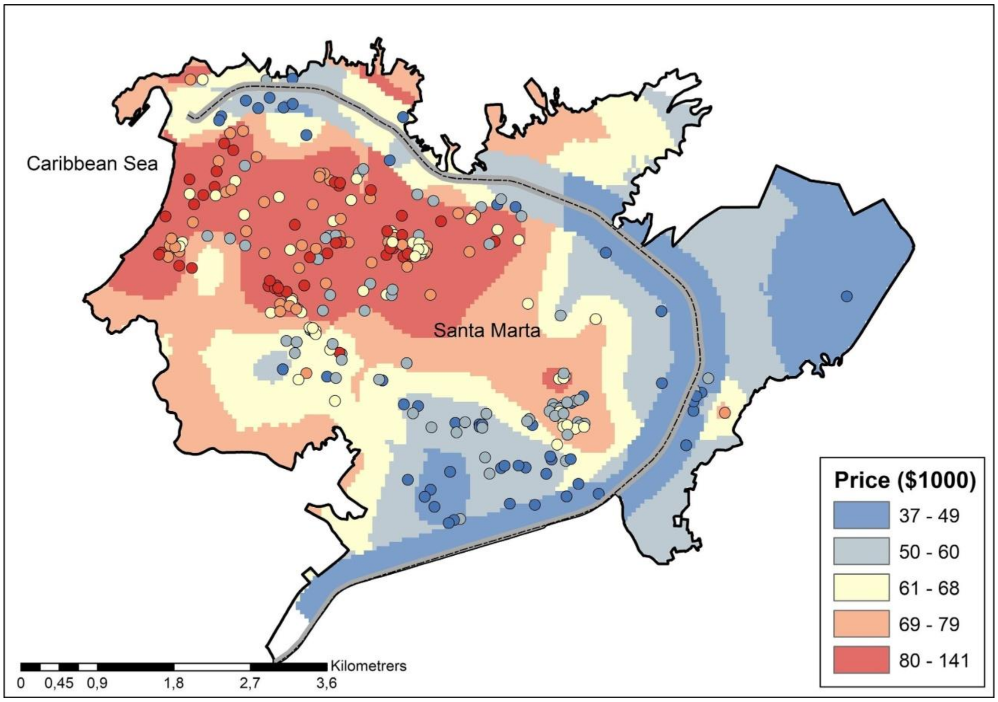
Figure 11 shows the predicted values versus the observed values using the five models in Table 8. As can be seen in the figure, the predictive capacity of the OLS models is lower than that of the SAR, SEM and RK models. The last three models show similar results.

Figure 11. Predicted versus observed values plotted for models 1 and 2 using OLS, SAR, SEM and RK methods. note: the dashed line represents the regression line and the solid line represents line 1:1. OLS, ordinary least squares; RK, regression kriging; RMSE, root mean squared error; SAR, spatial autoregressive model; SEM, spatial error model.



From the standpoint of the housing market, it is clearly of interest to determine housing prices at any location and thus obtain isovalue maps (Clapp, 2004; Yue et al., 2010). To do so, it is first necessary to estimate the price of housing at any point on the map. Although it is possible to determine the value of the locational variables (Dist-University, Dist-School, Buffer-Train, Dist-Sport, Dist-Beach, Dist-Park and Dist-Super) at any point on the plane, it is not possible to know the structural characteristics for these points of the map. Therefore, we have had to define a standard dwelling. A standard dwelling is obtained by assigning the numerical value of the sample average (Surface = 162.374 m2 and Age = 15.504 years) or mode (Garage = 1, Rooms = 3, Baths = 3 and Swimmingpool = 0) of these features to each of the structural characteristics of the dwelling. Figure 12 shows the estimated values of a standard dwelling in the city of Santa Marta.

Figure 12. Housing prices of a standard dwelling estimated with regression kriging (RK).



5.7 Discussion

In this paper we have attempted to quantify the negative externalities of the railway line in two ways: by determining a buffer zone for the freight railway line and by determining the price discount of houses located within the buffer zone.

From an economic standpoint, the second aspect is clearly of greater interest. To obtain the discount we calculated the percentage impact of this variable on housing prices using Equation 6, which is applied to the estimated coefficient of the variable Buffer-Train in the model. This percentage indicates to what extent the price of houses that are inside the buffer zone (determined by the 320 m radius) decreases due to the effect of the negative externalities of the freight railway line. This estimated percentage varies depending on the method used to estimate the hedonic model. The percentage ranges from 14% for the Mod2-SAR model to 23.7% for the Mod2-RK model. This indicates that the cost of an average priced house ($97,160) located within the buffer zone would decrease by from $13,600 to $23,030, depending on the estimation method of the model.

A negative effect on housing prices is observed for dwellings located within the buffer zone of the freight railway line. The estimated price of a standard dwelling located inside the buffer zone ranges from $37,470 to $75,530, with an average price of $50,100. The highest prices are found in the northern area and the lowest prices in the southern area of the buffer zone. This indicates that the prices of this type of dwelling located within the buffer zone are affected not only by their proximity to the train line, but also by their proximity to amenities such as schools, parks or sports facilities, which are more numerous in the north than in the south of the city (see Figure 9).

The isovalues map of the price of a standard dwelling (see Figure 12) obtained by RK can be very useful in the field of real estate analysis, particularly with regard to real estate investment trusts, property taxation, insurance and mortgage management. This map shows the spatial distribution of the price of standard dwellings, which varies in space due to locational characteristics. A negative effect is clearly observed for standard dwellings located inside the buffer zone of the freight line.

5.8 Conclusions

This paper presents a methodology to determine the buffer zone of a factor that may positively or negatively affect the value of real estate and the economic quantification depending on whether the property is within or outside the buffer zone. The main contribution of the study is that it identifies the impact zone of the rail line and accounts for spatial autocorrelation using spatial econometrics and geostatistical approaches. This methodology can be useful for both planning authorities and scholars in the fields of urban studies and economics.

The results of applying this methodology to the particular case of Santa Marta, Colombia show that the freight railway has negative effects (which may reflect noise, pollution, vibrations, hazards, visual impact, barriers, etc.) on the housing prices located in the proximity of the train line. A buffer zone with a width of 320 m was obtained for the railway line by maximising the explanatory capacity of the model. Houses located within the buffer zone are subject to a discount which varies depending on the estimation method that is used, and to other locational influences.

In addition, the results show that, in the case of Santa Marta, the significant variables of the hedonic model have the expected signs and their coefficients are similar for the different estimation methods. From the viewpoint of predictive capacity, the SAR, SEM and RK methods are similar to each other but superior to OLS.

Finally, this study has considered that the width of the buffer zone is constant along the railway line. One issue to consider in future research is that this width may not be constant, since some of the negative factors of the railway line can vary, such as railway level crossing safety (which affects the accident rate) or the installation of screens that reduce noise and foot traffic, among others. Another aspect to consider is the use of other spatial estimation methods such as geographically weighted regression in order to consider the spatial heterogeneity. We believe that both issues could be of interest in future studies.

5.9 References

Anselin, L. (1988). *Spatial Econometrics: Methods and Models* (Vol. 4). Springer Netherlands. https://doi.org/10.1007/978-94-015-7799-1

Anselin, L. (1992). Space and applied econometrics: Introduction. *Regional Science and Urban Economics*, *22*(3), 307–316. https://doi.org/10.1016/0166-0462(92)90031-U

Anselin, L., & Rey, S. J. (2012). Spatial econometrics in an age of CyberGIScience. *International Journal of Geographical Information Science*, *26*(12), 2211–2226. https://doi.org/10.1080/13658816.2012.664276

Anselin, L., Sridharan, S., & Gholston, S. (2007). Using Exploratory Spatial Data Analysis to Leverage Social Indicator Databases: The Discovery of Interesting Patterns. *Social Indicators Research*, *82*(2), 287–309. https://doi.org/10.1007/s11205-006-9034-x

Apparicio, P., Shearmur, R., Brochu, M., & Dussault, G. (2003). The measure of distance in a social science policy context: Advantages and costs of using network distances in eight canadians metropolitan areas. *Journal of Geographic Information and Decision Analysis*, *7*(2), 105–131.

Banister, D., & Thurstain-Goodwin, M. (2011). Quantification of the non-transport benefits resulting from rail investment. *Journal of Transport Geography*, *19*(2), 212–223. https://doi.org/10.1016/j.jtrangeo.2010.05.001

Bivand, R., & Piras, G. (2015). Comparing Implementations of Estimation Methods for Spatial Econometrics. *Journal of Statistical Software*, *63*, 1–36. https://doi.org/10.18637/jss.v063.i18

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. https://doi.org/10.1006/juec.2001.2214

Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, *22*(3), 453–474. https://doi.org/10.1016/0166-0462(92)90039-4

Case, B., Clapp, J., Dubin, R., & Rodriguez, M. (2004). Modeling Spatial and Temporal House Price Patterns: A Comparison of Four Models. *The Journal of Real Estate Finance and Economics*, *29*(2), 167–191. https://doi.org/10.1023/B:REAL.0000035309.60607.53

Cetintahra, G. E., & Cubukcu, E. (2015). The influence of environmental aesthetics on economic value of housing: An empirical research on virtual environments. *Journal of Housing and the Built Environment*, *30*(2), 331–340. https://doi.org/10.1007/s10901-014-9413-6

Chasco, C., & Gallo, J. L. (2013). The impact of objective and subjective measures of air quality and noise on house prices: A multilevel approach for downtown Madrid. *Economic Geography*, *89*(2), 127–148. https://doi.org/10.1111/j.1944-8287.2012.01172.x

Chau, K. W., & Chin, T. L. (2003). A Critical Review of Literature on the Hedonic Price Model. *International Journal for Housing Science and Its Applications*, *27*(2), 145–165.

Chen, W. Y., & Jim, C. Y. (2010). Amenities and disamenities: A hedonic analysis of the heterogeneous urban landscape in Shenzhen (China). *The Geographical Journal*, *176*(3), 227–240. https://doi.org/10.1111/j.1475-4959.2010.00358.x

Chen, W. Y., & Li, X. (2018). Impacts of urban stream pollution: A comparative spatial hedonic study of high-rise residential buildings in Guangzhou, south China. *The Geographical Journal*, *184*(3), 283–297. https://doi.org/10.1111/geoj.12246

Chi, G. (2010). The Impacts of Highway Expansion on Population Change: An Integrated Spatial Approach. *Rural Sociology*, *75*(1), 58–89. https://doi.org/10.1111/j.1549-0831.2009.00003.x

Chica Olmo, J. (1994). Teoría de las variables regionalizadas. Aplicación en economía espacial y valoración inmobiliaria. Universidad de Granada. https://editorial.ugr.es/libro/teoria-de-las-variables-regionalizadas_138704/

Chica-Olmo, J. (1995). Spatial Estimation of Housing Prices and Locational Rents. *Urban Studies*, *32*(8), 1331–1344. https://doi.org/10.1080/00420989550012492

Christensen, R. (1987). *Plane answers to complex questions: The theory of linear models*. Springer-Verlag.

Clapp, J. M. (2004). A Semiparametric Method for Estimating Local House Price Indices. *Real Estate Economics*, *32*(1), 127–160. https://doi.org/10.1111/j.1080-8620.2004.00086.x

Cohen, J. P., & Coughlin, C. C. (2008). Spatial Hedonic Models of Airport Noise, Proximity, and Housing Prices. *Journal of Regional Science*, *48*(5), 859–878. https://doi.org/10.1111/j.1467-9787.2008.00569.x

Conroy, S. J., & Milosch, J. L. (2011). An Estimation of the Coastal Premium for Residential Housing Prices in San Diego County. *The Journal of Real Estate Finance and Economics*, *42*(2), 211–228. https://doi.org/10.1007/s11146-009-9195-x

Corrado, L., & Fingleton, B. (2012). Where Is the Economics in Spatial Econometrics? *Journal of Regional Science*, *52*(2), 210–239. https://doi.org/10.1111/j.1467-9787.2011.00726.x

Cressie, N. (1990). The origins of kriging. *Mathematical Geology*, *22*(3), 239–252. https://doi.org/10.1007/BF00889887

Cressie, N. (1991). *Statistics for Spatial Data*. John Wiley & Sons.

Dai, X., Bai, X., & Xu, M. (2016). The influence of Beijing rail transfer stations on surrounding housing prices. *Habitat International*, *55*, 79–88. https://doi.org/10.1016/j.habitatint.2016.02.008

Debrezion, G., Pels, E., & Rietveld, P. (2007). The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics*, *35*(2), 161–180. https://doi.org/10.1007/s11146-007-9032-z

Debrezion, G., Pels, E., & Rietveld, P. (2011). The Impact of Rail Transport on Real Estate Prices: An Empirical Analysis of the Dutch Housing Market. *Urban Studies*, *48*(5), 997–1015. https://doi.org/10.1177/0042098010371395

Diao, M. (2015). Selectivity, spatial autocorrelation and the valuation of transit accessibility. *Urban Studies*, *52*(1), 159–177. https://doi.org/10.1177/0042098014523686

Diao, M., Qin, Y., & Sing, T. F. (2016). Negative Externalities of Rail Noise and Housing Values: Evidence from the Cessation of Railway Operations in Singapore. *Real Estate Economics*, *44*(4), 878–917. https://doi.org/10.1111/1540-6229.12123

Diao, M., Leonard, D., & Sing, T. F. (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics*, *67*, 64–77. https://doi.org/10.1016/j.regsciurbeco.2017.08.006

Dubé, J., Thériault, M., & Des Rosiers, F. (2013). Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009. *Transportation Research Part A: Policy and Practice*, *54*, 49–66. https://doi.org/10.1016/j.tra.2013.07.015

Dubin, R. A. (1992). Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics*, *22*(3), 433–452. https://doi.org/10.1016/0166-0462(92)90038-3

Dubin, R. A. (1998). Spatial Autocorrelation: A Primer. *Journal of Housing Economics*, *7*(4), 304–327. https://doi.org/10.1006/jhec.1998.0236

Fan, Q., Hansz, J. A., & Yang, X. (2016). The Pricing Effects of Open Space Amenities. *The Journal of Real Estate Finance and Economics*, *52*(3), 244–271. https://doi.org/10.1007/s11146-015-9508-1

Feng, X., & Humphreys, B. (2018). Assessing the Economic Impact of Sports Facilities on Residential Property Values: A Spatial Hedonic Approach. *Journal of Sports Economics*, *19*(2), 188–210. https://doi.org/10.1177/1527002515622318

Forouhar, A. (2016). Estimating the impact of metro rail stations on residential property values: Evidence from Tehran. *Public Transport*, *8*(3), 427–451. https://doi.org/10.1007/s12469-016-0144-9

Gadziński, J., & Radzimski, A. (2016). The first rapid tram line in Poland: How has it affected travel behaviours, housing choices and satisfaction, and apartment prices? *Journal of Transport Geography*, *54*, 451–463. https://doi.org/10.1016/j.jtrangeo.2015.11.001

Grass, R. G. (1992). The estimation of residential property values around transit station sites in Washington, D.C. *Journal of Economics and Finance*, *16*(2), 139–146. https://doi.org/10.1007/BF02920114

Guerra, E., Cervero, R., & Tischler, D. (2012). Half-Mile Circle: Does It Best Represent Transit Station Catchments? *Transportation Research Record*, *2276*(1), 101–109. https://doi.org/10.3141/2276-12

Halvorsen, R., & Palmquist, R. (1980). The Interpretation of Dummy Variables in Semilogarithmic Equations. *The American Economic Review*, *70*(3), 474–475.

Hengl, T., Heuvelink, G. B. M., & Rossiter, D. G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences*, *33*(10), 1301–1315. https://doi.org/10.1016/j.cageo.2007.05.001

Jang, M., & Kang, C.-D. (2015). Retail accessibility and proximity effects on housing prices in Seoul, Korea: A retail type and housing submarket approach. *Habitat International*, *49*, 516–528. https://doi.org/10.1016/j.habitatint.2015.07.004

Kahn, M. E. (2007). Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems. *Real Estate Economics*, *35*(2), 155–182. https://doi.org/10.1111/j.1540-6229.2007.00186.x

Kane, T. J., Riegg, S. K., & Staiger, D. O. (2006). School Quality, Neighborhoods, and Housing Prices. *American Law and Economics Review*, *8*(2), 183–212. https://doi.org/10.1093/aler/ahl007

Kilpatrick, J., Throupe, R., Carruthers, J., & Krause, A. (2007). The Impact of Transit Corridors on Residential Property Values. *Journal of Real Estate Research*, *29*(3), 303–320. https://doi.org/10.1080/10835547.2007.12091201

Le Boennec, R., & Salladarré, F. (2017). The impact of air pollution and noise on the real estate market. The case of the 2013 European Green Capital: Nantes, France. *Ecological Economics*, *138*, 82–89. https://doi.org/10.1016/j.ecolecon.2017.03.030

LeSage, J., & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Chapman and Hall/CRC. https://doi.org/10.1201/9781420064254

Li, R. Y. M., Chau, K. W., Law, C. Y., & Leung, T. H. (2016). Superstition and Hong Kong housing prices: A Hedonic Pricing approach. In *Econometric analyses of international housing markets* (pp. 92–100). Routledge.

Longley, P. A., Goodchild, M., Maguire, D. J., & Rhind, D. W. (2001). *Geographic Information Systems and Science* (1st ed.). John Wiley & Sons.

Matheron, G. (1965). *Les variables régionalisées et leur estimation: Une application de la théorie de fonctions aléatoires aux sciences de la nature*. Masson et Cie.

Matheron, G. (1970). *La Théorie des variables régionalisées, et ses applications*. Ecole Nationale Supérieure des Mines de Paris.

McMillen, D. P. (2004). Airport expansions and property values: The case of Chicago O'Hare Airport. *Journal of Urban Economics*, *55*(3), 627–640. https://doi.org/10.1016/j.jue.2004.01.001

Militino, A. F., Ugarte, M. D., & García-Reinaldos, L. (2004). Alternative Models for Describing Spatial Dependence among Dwelling Selling Prices. *The Journal of Real Estate Finance and Economics*, *29*(2), 193–209. https://doi.org/10.1023/B:REAL.0000035310.20223.e9

Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*, *37*(1/2), 17–23. https://doi.org/10.2307/2332142

Noticias Uno. (2013, February 17). *El tren de la muerte*. Canal 1. https://canal1.com.co/noticias/el-tren-de-la-muerte/

Osland, L. (2010). An Application of Spatial Econometrics in Relation to Hedonic House Price Modeling. *Journal of Real Estate Research*, *32*(3), 289–320. https://doi.org/10.1080/10835547.2010.12091282

Pace, R. K., Barry, R., & Sirmans, C. F. (1998). Spatial Statistics and Real Estate. *The Journal of Real Estate Finance and Economics*, *17*(1), 5–13. https://doi.org/10.1023/A:1007783811760

Perchoux, C., Chaix, B., Brondeel, R., & Kestens, Y. (2016). Residential buffer, perceived neighborhood, and individual activity space: New refinements in the definition of exposure areas – The RECORD Cohort Study. *Health & Place*, *40*, 116–122. https://doi.org/10.1016/j.healthplace.2016.05.004

Pope, J. C. (2008). Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise. *Journal of Urban Economics*, *63*(2), 498–516. https://doi.org/10.1016/j.jue.2007.03.003

Ramos N., N., Jiménez B., W. G., Arcos T., A., Acosta O., D., & Hurtado R., A. (2015). Consideraciones teóricas sobre la calidad de vida en el Distrito de Santa Marta. *Revista Escuela de Administración de Negocios*, *78*, 122–138.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Rouhani, S., Srivastava, R. M., Desbarats, A. J., Cromer, M. V., & Johnson, A. I. (1996). *STP1283-EB, Geostatistics for Environmental and Geotechnical Applications*. ASTM International. https://www.astm.org/stp1283-eb.html

Strand, J., & Vågnes, M. (2001). The relationship between property values and railroad proximity: A study based on hedonic prices and real estate brokers' appraisals. *Transportation*, *28*(2), 137–156. https://doi.org/10.1023/A:1010396902050

Tita, G. E., Petras, T. L., & Greenbaum, R. T. (2006). Crime and Residential Choice: A Neighborhood Level Analysis of the Impact of Crime on Housing Prices. *Journal of Quantitative Criminology*, *22*(4), 299–317. https://doi.org/10.1007/s10940-006-9013-z

Wardrip, K. (2011). Public Transit's Impact on Housing Costs: A Review of the Literature. *Center for Housing Policy: Insights from Housing Policy Research*.

Webster, R., & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*. John Wiley & Sons.

Weinberger, R. R. (2001). Light Rail Proximity: Benefit or Detriment in the Case of Santa Clara County, California? *Transportation Research Record*, *1747*(1), 104–113. https://doi.org/10.3141/1747-13

Wen, H., Bu, X., & Qin, Z. (2014). Spatial effect of lake landscape on housing price: A case study of the West Lake in Hangzhou, China. *Habitat International*, *44*, 31–40. https://doi.org/10.1016/j.habitatint.2014.05.001

Wen, H., Jin, Y., & Zhang, L. (2017). Spatial heterogeneity in implicit housing prices: Evidence from Hangzhou, China. *International Journal of Strategic Property Management*, *21*(1), 15–28. https://doi.org/10.3846/1648715X.2016.1247021

Wilhelmsson, M. (2000). The Impact of Traffic Noise on the Values of Single-family Houses. *Journal of Environmental Planning and Management*, *43*(6), 799–815. https://doi.org/10.1080/09640560020001692

Wu, C., Ye, X., Du, Q., & Luo, P. (2017). Spatial effects of accessibility to parks on housing prices in Shenzhen, China. *Habitat International*, *63*, 45–54. https://doi.org/10.1016/j.habitatint.2017.03.010

Xu, T., & Zhang, M. (2016). Tailoring empirical research on transit access premiums for planning applications. *Transport Policy*, *51*, 49–60. https://doi.org/10.1016/j.tranpol.2016.03.003

Yue, W., Liu, Y., & Fan, P. (2010). Polycentric Urban Development: The Case of Hangzhou. *Environment and Planning A: Economy and Space*, *42*(3), 563–577. https://doi.org/10.1068/a42116

CAPÍTULO 6. Conclusiones finales, implicaciones y líneas futuras

6.1 Conclusiones finales

El objeto principal de este trabajo ha sido realizar un estudio en el sector de bienes raíces, específicamente en las viviendas de tipo familiar, ya que son consideradas como un activo relevante de su patrimonio o de inversión. Los objetivos se enfocaron hacia la estimación del precio de la vivienda de segunda mano y su precio de arrendamiento. Específicamente, se logró determinar los factores que influyen en dicho precio y cuantificar el precio implícito de cada uno de ellos, teniendo en cuenta, no sólo los factores constructivos, sino también la ubicación del inmueble, la proximidad a lugares de interés y, en uno de los trabajos realizados, las externalidades negativas.

A continuación, se presentan las principales conclusiones obtenidas respecto de las hipótesis enunciadas en este trabajo:

La hipótesis 1 planteaba si el precio de venta de la vivienda se encuentra distribuido aleatoriamente en el espacio o, por el contrario, se encuentra autocorrelacionado espacialmente. La presencia de autocorrelación espacial ha posibilitado la utilización de metodologías que permiten obtener mapas de isovalores, en los que se observa una clara evidencia de dependencia espacial en la distribución de los precios de venta y de alquiler. En los capítulos 3, 4 y 5 estos mapas proporcionaron una representación visual de la distribución espacial de los precios de la vivienda en relación con los diferentes factores localizativos analizados en el estudio. Permitieron identificar patrones y tendencias en la relación entre la ubicación y el valor de las propiedades, lo que contribuyó a una comprensión más profunda del mercado inmobiliario objeto de estudio.

Se pueden identificar clústeres o áreas donde las viviendas presentan valores similares, lo que indica la existencia de patrones espaciales en la determinación de los precios de venta y alquiler. Estos resultados obtenidos a través del análisis espacial del precio de venta de las viviendas respaldan la idea de que existen factores espaciales que afectan a los precios de los inmuebles, como por ejemplo, la distancia al centro. Así, en las ciudades monocentricas se observa la tendencia de que las viviendas con mayor precio de venta o arrendamiento se ubican en la zona céntrica, quiere esto decir que los valores inmobiliarios disminuyen a medida que aumenta la distancia al centro (Von Thünen, 1966). Esta tendencia se muestra en los mapas de isovalores, y además, éstos brindan una comprensión más profunda de la estructura espacial de los valores de los bienes inmuebles. Asimismo, permiten identificar áreas con características similares y comprender mejor los factores que influyen en la determinación de los precios en el mercado inmobiliario.

Esta hipótesis fue desarrollada en los capítulos 3, 4 y 5. Los contrastes de autocorrelación espacial aplicados en los distintos trabajos ponen en evidencia la existencia de dependencia espacial en el precio de venta y de alquiler de las viviendas, por lo que se puede afirmar que no se distribuyen aleatoriamente en el espacio. En el capítulo 3 se presentó un modelo hedónico del precio de alquiler de la vivienda en Atenas, explicado con variables constructivas, que fueron estadísticamente significativas, y además se tuvo en cuenta la presencia de autocorrrelación espacial mediante el variograma (Webster & Oliver, 2007). Según los resultados obtenidos mediante Krigeaje Residual los coeficientes de las variables son estadísticamente significativas entre el 5% y el 10%. Por otro lado, el variograma utilizado frecuentemente en la literatura se determinó que los precios se encuentran relacionados con un radio de influencia de 2.635,813 metros. Se concluyó, por tanto, que el precio de arrendamiento no se encuentra distribuido aleatoriamente y que este precio está relacionado con la zona de ubicación del inmueble, debido a la autocorrelación espacial provocada por los factores de ubicación y a la presencia de contagio espacial.

Por otra parte, en los capítulos 4 y 5 se modelizó el precio de la vivienda de segunda mano en Santa Marta, considerando las variables constructivas y localizativas. Se evidenciaron relaciones directas del precio según la ubicación, con significación entre el 5% y 10% de las variables explicativas, excepto en el caso de algunas variables localizativas, como la distancia a la universidad, distancia al parque y distancia al supermercado.

La hipótesis 2 hace referencia a que la incorporación de información relativa a datos ausentes mejora la bondad y significación del modelo, En el capítulo 4, se abordó y desarrolló dicha hipótesis con el fin de solventar la falta de información suministrada por las páginas inmobiliarias en cuanto a la antigüedad de las viviendas. Durante el proceso de recopilación de datos, se encontró que no todas las páginas de portales inmobiliarios en internet proporcionaban la información completa sobre la antigüedad de las viviendas. Ante esta situación, se planteó la necesidad de estimar esta variable en un total de 69 viviendas para poder completar y enriquecer el conjunto de datos.

Para llevar a cabo esta estimación, se utilizó el método de Krigeaje, el cual se basa en un proceso de interpolación espacial para realizar estimaciones cuando la información está incompleta. El método de Krigeaje permite considerar la estructura espacial de los datos disponibles para generar estimaciones confiables de la antigüedad de las viviendas en las ubicaciones donde esta información no estaba disponible inicialmente.

La aplicación del método de Krigeaje en este contexto permitió obtener estimaciones robustas y precisas de la antigüedad de las viviendas en las observaciones donde esta variable no se encontraba registrada. Esto resultó fundamental para garantizar la integridad y la calidad de los datos utilizados en el estudio, ya que la falta información podría haber generado sesgos o limitaciones en los análisis posteriores.

Estas estimaciones, efectivamente, contribuyeron a mejorar la bondad del ajuste (pasando de un $R^2 = 0,76$ a un $R^2 = 0,84$) y la significación de algunas variables explicativas, entre ellas, la antigüedad, que, en un principio no era significativa y pasó a ser significativa al 10%.

La hipótesis 3 plantea la existencia de una tendencia creciente en los precios de la vivienda desde la periferia hacia la costa. En el capítulo 4, se llevó a cabo un análisis para validar esta hipótesis, incorporando en el modelo la variable de la distancia de cada una de las variables a la costa. A través del mapa de isovalores, se pudo observar una clara deriva en los precios, donde se evidenció que en la zona costera los precios eran significativamente más altos, mientras que a medida que se desplazaba hacia la periferia, los precios disminuían y se encontraban viviendas con precios más bajos. Esta situación puede atribuirse a la alta demanda de propiedades en zonas cercanas al mar.

La hipótesis 4 plantea que los precios de la vivienda no están necesariamente siempre relacionados con la cercanía a lugares de interés como universidades, parques y otros puntos de referencia. Los resultados obtenidos en el análisis revelaron que estas variables localizativas no mostraron una significancia estadística en el modelo presentado. Esto indica que no se encontró una relación clara y definida entre la proximidad a estos lugares de interés y los precios de la vivienda. Es posible que este resultado se deba a la presencia de viviendas con precios muy diferentes cercanos a los lugares de interés, lo que podría provocar que la distancia a esos lugares de interés no sea significativa (Tamaris-Turizo et al., 2022). Estos hallazgos sugieren que la relación entre los precios de la vivienda y la proximidad a lugares de interés puede ser más compleja de lo que se suponía inicialmente.

La hipótesis 5 se planteó con el objetivo de investigar si la distancia con respecto a la vía del tren de mercancías afecta negativamente el precio de la vivienda. Este aspecto fue desarrollado en detalle en el capítulo 5, en el que se llevó a cabo la estimación de un modelo de precios hedónicos específico para analizar esta relación negativa. Para realizar la estimación del modelo, se utilizaron distintos métodos de Econometría Espacial (SAR y SEM) y Geostadística (Regression Kriging). Los hallazgos obtenidos, como el signo esperado de la variable Buffer-Train, que resultó negativo y con un nivel de significanción del 1% confirmaron la hipótesis inicial, revelando una relación negativa entre la distancia a la vía del tren y el precio de la vivienda. Esto implica que a medida que la vivienda se encuentra más cercana a la vía del tren de mercancías, su precio tiende a disminuir. Esta relación negativa se puede explicar por los potenciales impactos negativos asociados con la proximidad al tráfico ferroviario de carga, como el ruido, las vibraciones y otros efectos indeseables. Estos factores pueden afectar negativamente la calidad de vida de los residentes y disminuir el atractivo de la ubicación, lo que a su vez se refleja en una reducción en el valor de la vivienda.

El uso del método de RK, como una combinación entre regresión y método de Krigeaje (Chica-Olmo, 1995; Dubin, 1992; Hengl et al., 2007) arrojó los mejores resultados, y permitió cuantificar y validar de manera precisa la relación negativa entre la distancia a la vía del tren de mercancías y el precio de la vivienda.

La hipótesis 6, pretende determinar el radio de influencia que tiene la vía del tren sobre el precio de la vivienda. Esta hipótesis fue analizada y desarrollada en el capítulo 5, aplicando métodos de la Geoestadística y la Econometría Espacial. Para llevar a cabo este análisis, se utilizó un método iterativo basado en la mejora de la bondad del ajuste del modelo para cuantificar la distancia del buffer alrededor de la línea del tren, lo que permitió delimitar una zona de afectación. A través del buffer, una de las herramientas más utilizada por los SIG (Longley et al., 2001), se observó que los precios de las viviendas variaban en relación con su proximidad a la línea del tren.

Los resultados obtenidos revelaron que, en un rango de distancia de 320 metros en torno a la vía del tren, las viviendas presentaban precios significativamente más bajos. Esto indica la existencia de una externalidad negativa asociada al paso del tren de mercancías, la cual afecta negativamente el valor de las propiedades cercanas. El uso de la Geoestadística y la Econometría Espacial en este estudio fue fundamental para comprender y cuantificar el impacto de vivir cerca de la vía del tren en el precio de la vivienda. Además, estos enfoques permitieron capturar la estructura espacial de los datos y considerar la dependencia espacial en el análisis.

6.2 Implicaciones y líneas futuras

La presente tesis doctoral, busca proporcionar aportes académicos relacionado con la estimación del precio de la vivienda de segunda o en situaciones de arrendamiento, así mismo, muestra la tendencia del precio a lo largo del territorio por medio de una cartografía donde se refleja su estructura espacial (plano de isovalores). Además de considerar variables constructivas, como el número de habitaciones, de baños, el área construida, la antigüedad, etc., es imprescindible también la incorporación de variables localizativas, tal como la distancia al centro u otros lugares de interés (distancia a escuelas, hospitales, zonas verdes, deportivas, etc.), las cuales pueden ser determinantes para explicar el precio de los bienes inmuebles (Chica-Olmo et al., 2019).

En particular, uno de los principales resultados obtenidos en esta tesis doctoral es que la aplicación de las técnicas de la Geoestadística, de la Econometría Espacial y de los SIG ha permitido cuantificar el impacto negativo (externalidades) de la cercanía a la vía de un tren de carga sobre el precio de las viviendas circundantes. Estas técnicas permitieron obtener un mapa en el que se muestra la zona de afectación o buffer, en el cual se observa claramente el alcance espacial de la influencia negativa de la vía del tren de carga.

La cuantificación de este impacto negativo proporciona información relevante y valiosa para la planificación urbana y el sector inmobiliario en el desarrollo de proyectos, especialmente en el ámbito del desarrollo inmobiliario residencial de la ciudad de Santa Marta. Al considerar estos resultados en el proceso de diseño y desarrollo de proyectos inmobiliarios, se pueden tomar decisiones más informadas y estratégicas, teniendo en cuenta los posibles efectos adversos asociados a la cercanía de la línea férrea y el paso de los trenes de carga.

Estos hallazgos subrayan la importancia de integrar aspectos relacionados con la ubicación y la infraestructura circundante en el desarrollo de proyectos inmobiliarios residenciales. La presencia de una línea férrea y la actividad de transporte de carga asociada pueden tener un impacto significativo en la calidad de vida de los residentes y en el valor de las propiedades, lo que debe ser considerado en el diseño urbano y en la toma de decisiones de inversión.

Por otro lado, otro resultado de gran interés alcanzado en esta tesis ha sido el derivado del estudio sobre los precios de arrendamiento en la ciudad de Atenas. En concreto, se evidencia que estos precios no se distribuyen espacialmente de manera aleatoria. Por el contrario, se encontró evidencia de un patrón de contagio espacial, donde las viviendas cercanas tienden a tener precios de alquiler similares. Este hallazgo indica que el precio de arrendamiento de una vivienda está influenciado por su ubicación relativa en el espacio en relación con sus vecinos. En otras palabras, la ubicación relativa del inmueble desempeña un papel crucial en la determinación del precio de arrendamiento. Las viviendas que se encuentran próximas entre sí comparten características similares en términos de accesibilidad, servicios y entorno, lo que influye en la oferta y demanda de alquiler en esa área específica. Además, el efecto de contagio espacial implica que existan interacciones y relaciones entre las viviendas en términos de precios de arrendamiento. Las viviendas vecinas pueden afectarse mutuamente en cuanto a la fijación de precios, ya sea a través de la competencia directa o de la percepción de valor influenciada por la ubicación. Este avance en la comprensión del comportamiento espacial del precio de arrendamiento de las viviendas tiene importantes implicaciones para los propietarios, inquilinos, inversores y profesionales del sector inmobiliario. Esto es así porque permite una mejor comprensión de las dinámicas de precios en diferentes áreas geográficas y facilita la toma de decisiones en la inversión inmobiliaria y la elección de la ubicación adecuada para alquilar una vivienda.

Desde el punto de vista económico, las investigaciones desarrolladas en esta tesis aportan información valiosa a los diferentes actores que de alguna u otra forma se relacionan con el sector inmobiliario. Por una parte, los propietarios o inmobiliarias ahora tienen otro punto de referencia para dar una valoración aproximada de sus inmuebles, la cual debe estar relacionada con su ubicación relativa, es decir, con el precio de las viviendas aledañas. Además de tener en cuenta sus variables constructivas, también deben considerarse las amenidades y lugares de interés que se encuentran cerca de ella, así como, de manera contraria, se ha de considerar si la vivienda está afectada por externalidades negativas.

Por otro lado, existen ciertas instituciones que se pueden interesar por esta clase de estudios, como las entidades públicas encargadas de la recaudación del impuesto de bienes inmuebles, puesto que los resultados pueden ser una fuente de apoyo para realizar estimaciones presupuestales sobre la recaudación del tributo.

Por último, también se dan avances en materia de estimación de datos faltantes relacionado con los años de antigüedad de la vivienda, donde en un principio no se contaba con la información para todas las viviendas. Por lo tanto, fue novedoso haber aplicado métodos de la Gesoestadística, tal como el Krigeaje, para solucionar este problema.

Finalmente, y de acuerdo con las implicaciones mencionadas anteriormente, se podrían presentar nuevas líneas futuras de investigación. Por ejemplo, georreferenciar los lugares donde se han presentado hurto o algún tipo de delincuencia que pueda afectar negativamente el precio de la vivienda o el precio del alquiler.

También sería de interés estudiar la relación entre la pobreza y el precio de las viviendas o el precio del arrendamiento. En este caso, el objetivo sería explorar y comprender cómo la pobreza se vincula con la distribución espacial de los precios de la vivienda a lo largo de un territorio determinado. Para lograr este propósito sería fundamental llevar a cabo una caracterización geográfica de las zonas con presencia de pobreza, utilizando indicadores multidimensionales. Estos indicadores pueden incluir variables como nivel de ingresos, tasa de desempleo, nivel de alfabetización, etc. Esta información podría complementarse con otros factores que reflejen las condiciones de vida en esas áreas. Es importante destacar que este tipo de estudio requiere de una metodología rigurosa y una recopilación de datos precisa. Es necesario utilizar herramientas analíticas y técnicas econométricas adecuadas para evaluar la relación entre la pobreza y los precios de la vivienda en un contexto espacial. Al relacionar la caracterización geográfica de las zonas de pobreza con la distribución espacial de los precios de la vivienda, se podrán identificar posibles patrones y tendencias. Es de esperar que se observe una correlación entre la presencia de pobreza y precios más bajos de las propiedades. Este análisis permitirá obtener una visión más completa y profunda de cómo la pobreza y sus dimensiones socioeconómicas impactan en el mercado inmobiliario y en la distribución espacial de los precios de la vivienda. Además, proporcionará información valiosa para los planificadores urbanos, responsables de políticas públicas y otros actores del sector inmobiliario.

Por último, en cuanto al caso de la ciudad de Santa Marta, sería pertinente considerar la inclusión de otro tipo de externalidades negativas que puedan tener un impacto significativo en el precio de la vivienda. Un ejemplo relevante sería la afectación del río Manzanares, un cuerpo de agua que atraviesa la ciudad de Santa Marta y discurre por la parte trasera de algunas viviendas. El río Manzanares se caracteriza por su caudal reducido y su alto nivel de contaminación, lo cual plantea preocupaciones ambientales. Esta situación puede generar efectos negativos en la calidad de vida de los residentes y, por ende, influir en el valor de las propiedades ubicadas a lo largo de la ribera. Por tanto, sería relevante desarrollar un modelo que permita analizar y determinar cómo la afectación del río Manzanares impacta espacialmente sobre el precio de la vivienda en diferentes puntos de la ribera. Esto implicaría considerar variables como la distancia a la ribera, la calidad del agua, la presencia de olores desagradables y otros indicadores de contaminación. Mediante la modelización de estas variables y su relación con los precios de las viviendas, se podría obtener una comprensión más precisa y detallada de cómo el factor ambiental del río afecta al precio de las propiedades. Esto permitiría identificar las áreas más afectadas y evaluar el grado de influencia de esta externalidad negativa en los precios de la vivienda.

Referencias Bibliográficas

Agudelo Torres, J. E., Agudelo Torres, G. A., Franco Arbeláez, L. C., & Franco Ceballos, L. E. (2015). The Effect of a Sports Stadium on Housing Rents: An Application of Geographically Weighted Regression. *Ecos de Economía*, *19*(40), 66–80. https://doi.org/10.17230/ecos.2015.40.4

Basu, S., & Thibodeau, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics*, *17*(1), 61–85. https://doi.org/10.1023/A:1007703229507

Blanco-Romero, A., Blázquez-Salom, M., & Cànoves, G. (2018). Barcelona, Housing Rent Bubble in a Tourist City. Social Responses and Local Policies. *Sustainability*, *10*(6), 2043. https://doi.org/10.3390/su10062043

Chen, W. Y., & Jim, C. Y. (2010). Amenities and disamenities: A hedonic analysis of the heterogeneous urban landscape in Shenzhen (China). *The Geographical Journal*, *176*(3), 227–240. https://doi.org/10.1111/j.1475-4959.2010.00358.x

Chica-Olmo, J. (1995). Spatial Estimation of Housing Prices and Locational Rents. *Urban Studies*, *32*(8), 1331–1344. https://doi.org/10.1080/00420989550012492

Chica-Olmo, J. (2007). Prediction of Housing Location Price by a Multivariate Spatial Method: Cokriging. *Journal of Real Estate Research*, *29*(1), 91–114. https://doi.org/10.1080/10835547.2007.12091188

Chica-Olmo, J., Cano-Guervos, R., & Chica-Olmo, M. (2013). A Coregionalized Model to Predict Housing Prices. *Urban Geography*, *34*(3), 395–412. https://doi.org/10.1080/02723638.2013.778662

Chica-Olmo, J., Cano-Guervos, R., & Tamaris-Turizo, I. (2019). Determination of buffer zone for negative externalities: Effect on housing prices. *The Geographical Journal*, *185*(2), 222–236. https://doi.org/10.1111/geoj.12289

Chica-Olmo, J., Cano-Guervos, R., Moschovaki, M.-D., & Tamaris-Turizo, I. (2021). What is the effect of location on rental housing prices in Athens? *Revista Brasileira de Gestão de Negócios*, *23*(3), 439–453. https://doi.org/10.7819/rbgn.v23i3.4114

Dubin, R. A. (1992). Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics*, *22*(3), 433–452. https://doi.org/10.1016/0166-0462(92)90038-3

Efthymiou, D., & Antoniou, C. (2013). How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. *Transportation Research Part A: Policy and Practice*, *52*, 1–22. https://doi.org/10.1016/j.tra.2013.04.002

Gabe, J., Robinson, S., & Sanderford, A. (2022). The Relationship Between School Quality and U.S. Multi-family Housing Rents. *The Journal of Real Estate Finance and Economics*, *64*(4), 615–645. https://doi.org/10.1007/s11146-020-09814-0

Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, *5*(4), 471–484. https://doi.org/10.1016/0094-1190(78)90004-9

Hengl, T., Heuvelink, G. B. M., & Rossiter, D. G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences*, *33*(10), 1301–1315. https://doi.org/10.1016/j.cageo.2007.05.001

Igan, D. O., & Loungani, P. (2012). *Global Housing Cycles*. International Monetary Fund. https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Global-Housing-Cycles-26229

Kuroda, Y. (2018). The effect of school quality on housing rents: Evidence from Matsue city in Japan. *Journal of the Japanese and International Economies*, *50*, 16–25. https://doi.org/10.1016/j.jjie.2018.05.002

Longley, P. A., Goodchild, M., Maguire, D. J., & Rhind, D. W. (2001). *Geographic Information Systems and Science* (1st ed.). John Wiley & Sons.

Matheron, G. (1970). *La Théorie des variables régionalisées, et ses applications*. Ecole Nationale Supérieure des Mines de Paris.

Nelson, J. P. (2010). Valuing Rural Recreation Amenities: Hedonic Prices for Vacation Rental Houses at Deep Creek Lake, Maryland. *Agricultural and Resource Economics Review*, *39*(3), 485–504. https://doi.org/10.1017/S1068280500007462

Nicholson, W. (1998). *Microeconomic Theory: Basic Principles and Extensions* (7th ed.). The Dryden Press.

Ofori, G. (Ed.). (2011). *New Perspectives on Construction in Developing Countries* (1st ed.). Routledge. https://doi.org/10.4324/9780203847343

Palomino Silva, J., Hennings Otoya, J., & Echevarría Alvarado, V. R. (2017). ANÁLISIS MACROECONÓMICO DEL SECTOR CONSTRUCCIÓN EN EL PERÚ. *Quipukamayoc*, *25*(47), 95–101. https://doi.org/10.15381/quipu.v25i47.13807

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Sadayuki, T. (2018). Measuring the spatial effect of multiple sites: An application to housing rent and public transportation in Tokyo, Japan. Regional Science and Urban Economics, 70, 155–173. https://doi.org/10.1016/j.regsciurbeco.2018.03.002

Tamaris Turizo, I., Chica Olmo, J., & Cano Guervos, R. (2022). The use of geostatistics to estimate missing data in a spatial econometric model of housing prices. *Revista Ingeniería de Construcción*, *37*(3), 423–434. https://doi.org/10.7764/ric.00044.21

Tejada, M., & Vidaurre, R. (2007). ¿Comprar o arrendar una vivienda? Una aplicación para el caso de la ciudad de La Paz. *Desarrollo y Sociedad*, *60*, 199–228. https://doi.org/10.13043/dys.60.6

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234–240. https://doi.org/10.2307/143141

Vásquez, W. F., & Beaudin, L. (2020). On the use of hypothetical price data to estimate hedonic models in a developing country context. *Letters in Spatial and Resource Sciences*, *13*(3), 219–231. https://doi.org/10.1007/s12076-020-00255-0

von Thünen, J. H. (1966). *Isolated State: An English Edition of Der Isolierte Staat*. Pergamon Press.

Webster, R., & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*. John Wiley & Sons.

Wen, H., Zhang, Y., & Zhang, L. (2014). Do educational facilities affect housing price? An empirical study in Hangzhou, China. *Habitat International*, *42*, 155–163. https://doi.org/10.1016/j.habitatint.2013.12.004

Yang, L., Zhou, J., Shyr, O. F., & Huo, (Derek) Da. (2019). Does bus accessibility affect property prices? *Cities*, *84*, 56–65. https://doi.org/10.1016/j.cities.2018.07.005

Zambrano-Monserrate, M. A., & Ruano, M. A. (2021). The spatial effect of estuaries pollution on the housing rental market: Evidence from South America. *Environment and Development Economics*, *26*(1), 69–87. https://doi.org/10.1017/S1355770X20000194