

Cómo citar este trabajo: Molinero-Parejo, R., Aguilera-Benavente, F., & Gómez-Delgado, M. (2021). Geographically Weighted Logistic Regression to identify explanatory factor of land use distribution in future scenarios of urban growth. *Boletín de la Asociación de Geógrafos Españoles*, (88). <https://doi.org/10.21138/bage.3052>

Regresión Logística Geográficamente Ponderada para identificar los factores explicativos de la distribución de usos de suelo en escenarios futuros de crecimiento urbano

Geographically Weighted Logistic Regression to identify explanatory
factor of land use distribution in future scenarios of urban growth

Ramón Molinero-Parejo 

ramon.molinero@uah.es

Francisco Aguilera-Benavente 

f.aguilera@uah.es

Montserrat Gómez-Delgado 

montserrat.gomez@uah.es

*Departamento de Geología, Geografía y Medio Ambiente
Universidad de Alcalá (España)*

Resumen

La expansión urbana es un proceso espacio-temporal que refleja los patrones de localización de la población y sus actividades, pudiendo provocar impactos irreversibles sobre el territorio. La construcción de narrativas y el posterior cartografiado de escenarios futuros se ha revelado como una estrategia de planificación que puede ayudar en la organización de los usos del suelo y el transporte. El Corredor del Henares (Madrid) se ha escogido como caso de estudio,

Recepción: 24.09.2020

Aceptación: 26.11.2020

Publicación: 15.02.2021

representando la evolución espacial de cinco usos del suelo urbano en tres escenarios disruptivos para 2050. El presente trabajo tiene como objetivo determinar qué factores explicativos influyeron en la distribución espacial de los cinco usos que se proyectaron en un previo taller de cartografía colaborativa. De esta forma, se pretende obtener información más precisa acerca de los factores explicativos del cambio de uso en cada escenario, que posteriormente pudiera ser aplicable a nuevos modelos espaciales de simulación urbana. Así, dado el carácter ubicuo de los procesos de urbanización, se empleó Regresión Logística Geográficamente Ponderada al permitir un análisis espacial de las relaciones entre los factores explicativos que los modelos globales no permiten estudiar. Los resultados obtenidos mostraron coincidencias entre los factores más significativos del modelo y las narrativas de los escenarios.

Palabras clave: escenarios; Regresión Logística Geográficamente Ponderada (RLGP); cartografía colaborativa; usos de suelo urbano; factores explicativos.

Abstract

Urban expansion is a spatio-temporal process that reflects the localization patterns of the population and its activities, which can cause irreversible impacts on the territory. The construction of narratives and the subsequent mapping of future scenarios has been revealed as a planning technique, which can help in the management and planning of land use and transport. The Corridor of Henares (Madrid) has been chosen as a case study, representing the spatial evolution of five urban land uses in three disruptive scenarios for 2050. The present research aims to determine which explanatory factors influenced the process of spatialization of the five uses that was projected in a previous mapping workshop. In this way, the aim is to provide more precise information about the driving factors of use changes in each scenario, which could later be applied to new spatial models of urban simulation. Thus, given the ubiquitous nature of urbanization processes, Geographically Weighted Logistic Regression (GWLR) was used as it enables a spatial analysis of the relationships between the explanatory factors that global models do not allow to study. The outcomes showed coincidences between the most significant factors of the model and the scenario narratives.

Key words: scenarios; Geographically Weighted Logistic Regression (GWLR); collaborative mapping; urban land use; explanatory factors.

1 Introducción

Desde mediados del siglo XX, la población mundial que reside en áreas urbanas ha experimentado un crecimiento medio anual del 2,5 % aproximadamente, duplicando la población urbana en ese mismo periodo (United Nations, 2018). En consecuencia, las zonas urbanizadas del planeta se encuentran en continua expansión, presentando ritmos más acelerados en los países en vías de desarrollo (Hou, Wang & Murayama, 2019). En el caso Europeo, el rápido crecimiento de los espacios periurbanos de las áreas metropolitanas está generando crecimientos dispersos frente a un desarrollo histórico más compacto (Ustaoglu, Williams & Petrov, 2017). Más concretamente, en España, este fenómeno ha provocado profundas transformaciones territoriales, traducidas principalmente en intensos procesos de suburbanización y fragmentación del espacio (Olazabal & Bellet, 2019; Pascual Rosa, Aguilera-Benavente & Salado García, 2019). Todo ello ha supuesto impactos irreversibles sobre el territorio, afectando enormemente al conjunto de la sociedad en términos de gestión y acceso a recursos, contaminación, problemas de índole social y económica, etc. (Bharath, Chandan, Vinay & Ramachandra, 2018; Chen, 2007).

En este contexto, el desarrollo de instrumentos de soporte a la toma de decisiones y la planificación urbana se ha visto impulsado en las últimas décadas. Así, el empleo de escenarios futuros en planificación (Börjeson, Höjer, Dreborg, Ekvall & Finnveden, 2006), ha sido uno de los instrumentos que mejor se ha adaptado a la gestión de los usos del suelo, permitiendo conocer, desde diferentes enfoques, cómo afecta la evolución futura de los distintos usos a su distribución espacial (Gallardo Beltrán, 2014). Este instrumento de planificación ha ido adquiriendo especial protagonismo en la literatura científica, marcando un importante desarrollo en la línea del crecimiento urbano (Avin & Goodspeed, 2020; Chotchaiwong & Wijitkosum, 2019; Dadashpoor, Azizi & Moghadasi, 2019; Perveen, Kamruzzaman & Yigitcanlar, 2017; Ustaoglu et al., 2017). Gracias a estos escenarios es posible explorar diferentes alternativas de evolución de los sistemas urbanos, a la vez que permiten examinar los problemas de la expansión urbana en una línea temporal desde distintas perspectivas (Deal, Pan, Timm & Pallathucheril, 2017). En este sentido, ofrecen la capacidad de deliberar sobre el futuro para revelar y comprender los principales problemas de la planificación, utilizando diferentes técnicas de representación espacial como el cartografiado, el dibujo, la realización de esquemas o la utilización de marcadores (Radinsky et al., 2017). Estas ventajas constituyen un sistema funcional de planificación futura que permite actuar sobre eventos más allá de la tendencia, es decir, ofrece cambios plausibles en la trayectoria territorial fuera de la idea de *Business as usual* (BAU)

(van der Heijden, 2000). Tratar con diferentes futuros que escapan del pensamiento común prepara a los planificadores ante el surgimiento de eventos disruptivos e inesperados, identificando las debilidades y fortalezas que pudiesen tener frente a diversas situaciones venideras (Zapata & Kaza, 2015). A dichas ventajas se debe añadir la fácil adaptación y aplicación de los escenarios a distintas escalas y ámbitos territoriales, además de la complementariedad con un amplio abanico de instrumentos de planificación, lo que justifica la implementación de escenarios en la metodología del presente artículo.

Este enfoque de *scenario planning* es por el que apuesta el proyecto TRANSURBAN en el que se enmarca este trabajo. Para abordarlo, entre sus diferentes actividades, se organizó un taller de cartografía colaborativa donde participó un amplio conjunto de expertos adscritos al campo de la planificación urbana y el transporte, con la finalidad de cartografiar tres narrativas que representaban futuros disruptivos para el Corredor del Henares (Madrid). El resultado fue la obtención de tres escenarios que mostraban la configuración espacial de distintas posibilidades de evolución del crecimiento urbano y de la red de transporte, aportando información acerca de la magnitud y ubicación de los cambios asociados a cada escenario (p. ej.: cantidad de crecimiento).

Además del valor que de por sí representan los escenarios obtenidos, es interesante explorar los factores espaciales que se encuentran relacionados con los cambios observados en cada escenario por diversas razones: a) por una parte, el estudio de dichos factores puede ayudar a constatar si las localizaciones de los crecimientos obtenidas en talleres colaborativos operan desde la lógica de los escenarios en los que se basan, es decir, permite reconocer y aceptar el criterio de los expertos mediante técnicas de estadística espacial; b) por otra parte, la identificación de dichos factores explicativos, puede ser de utilidad para el diseño y calibración de modelos de simulación urbana (Barreira-González, Aguilera-Benavente & Gómez-Delgado, 2017; Jokar Arsanjani, Helbich, Kainz & Darvishi Bolorani, 2013), especialmente en la configuración de reglas de transición para aquellos basados en Autómatas Celulares (AC), aportando un mayor realismo a estos (Barreira-González, Gómez-Delgado & Aguilera-Benavente, 2015); y c) permiten la formulación de políticas y estrategias de planificación más sostenibles (Shafizadeh-Moghadam & Helbich, 2015), ya que pueden ayudar a identificar espacios que por su localización sean considerados atractivos para el desarrollo urbano y, a su vez, preservar de esta ocupación aquellos espacios de elevado valor ambiental.

En definitiva, los planificadores urbanísticos y los agentes con capacidad decisoria requieren información precisa y detallada enfocada a comprender los procesos de expansión urbana, que les permita asesorar adecuadamente la creación de planes orientados a la gestión de estos espacios (Jokar Arsanjani et al., 2013). Por consiguiente, el estudio de los factores que dirigen y condicionan el crecimiento urbano es fundamental para formular estrategias futuras de planeamiento sostenible (Jokar Arsanjani et al., 2013; Shafizadeh-Moghadam & Helbich, 2015). Dicho estudio es posible gracias al uso de técnicas enfocadas a evaluar y cuantificar la variabilidad de estos factores.

A la hora de entender las relaciones entre la amplia diversidad de factores, los modelos de regresión globales (Hu & Lo, 2007) no resultan los más idóneos debido a que no tienen en cuenta la componente espacial, pudiendo ocultar relaciones espaciales interesantes (Brunsdon, Fotheringham & Charlton, 1996). Dichos autores demostraron que las estimaciones de los parámetros de regresión pueden variar en el espacio, afirmación a partir de la cual desarrollaron una familia de modelos matemáticos no estacionarios que se ajustasen a este. Ello ha permitido adecuar estos modelos de regresión a los datos geográficos, de tal manera que es posible evaluar y cuantificar los efectos de las variables observando variaciones espaciales (Gao & Li, 2011; Luo & Wei, 2009; Pineda Jaimes, Bosque Sendra, Gómez Delgado & Franco Plata, 2010).

Por ello, los modelos de Regresión Logística Geográficamente Ponderada (RLGP) han experimentado un importante auge en las últimas décadas, mostrando especial interés en el estudio de las relaciones espaciotemporales entre los factores del crecimiento urbano (Guanglong, Erqi & Hongqi, 2017) o en su variabilidad espacial (Shafizadeh-Moghadam & Helbich, 2015).

Aunque han sido aplicados en diversos contextos espacialmente explícitos del transporte y la movilidad, como el cálculo de la demanda en las estaciones de metro (Gutierrez-Puebla, García-Palomares & Daniel-Cardozo, 2012) o la exploración espacial de la variabilidad en el comportamiento y modos de viaje (Nkeki & Asikhia, 2019), el presente artículo los emplea para explorar y comprender el proceso de expansión urbana (Guanglong et al., 2017; Luo & Wei, 2009; Mondal, Das & Dolui, 2015; Shafizadeh-Moghadam & Helbich, 2015), haciendo especial hincapié en su aplicación para mejorar modelos de simulación del crecimiento urbano basados en AC, y más concretamente, en la definición de las reglas de transición y la mejora de la precisión en las simulaciones (Mirbagheri & Alimohammadi, 2017).

Así, la presente investigación tiene como objetivo determinar qué factores explicativos del crecimiento urbano influyeron en la distribución espacial de los usos del suelo para cada escenario generado en el proceso de cartografiado colaborativo desde una doble perspectiva: por un lado, valorar si la reconfiguración del espacio establecida por los participantes del taller se corresponde con las descripciones de los escenarios y, por otra parte, obtener resultados aplicables al diseño y calibración de modelos de simulación, para así configurar una técnica robusta y completa de soporte para la planificación.

El presente artículo se estructura de la siguiente forma: la sección 2 trata el caso de estudio, la información de partida y la metodología desarrollada, exponiendo el Análisis Exploratorio de Datos (AED), las variables independientes y dependientes, el desarrollo y calibración de RLGP y la evaluación de los modelos mediante la comparación de resultados y la aplicación del estadístico ROC; en la sección 3 se exponen los resultados obtenidos; en la sección 4 se discuten los resultados y, por último, en la sección 5, se narran las conclusiones.

2 Materiales y métodos

2.1 Caso de estudio

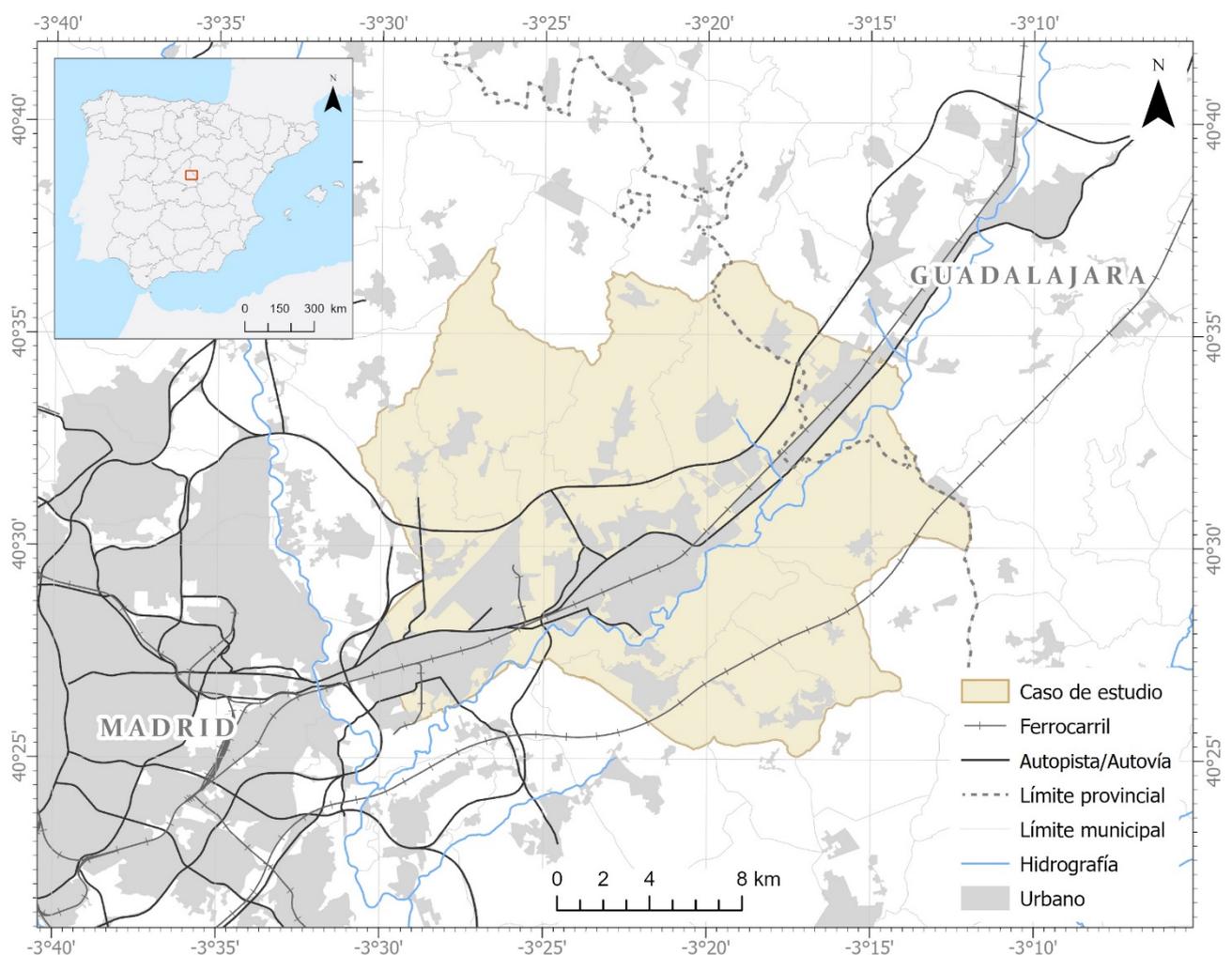
Para alcanzar el objetivo establecido, se utilizó como caso de estudio un sector del Corredor del Henares (Madrid–Guadalajara) (Figura 1). Este se caracteriza por ser un espacio metropolitano estructurado a lo largo del río Henares y otras infraestructuras de transporte como la autovía A-2 y las líneas de AVE y ferrocarril que conectan Madrid– Zaragoza–Barcelona. Su fuerte carácter urbano industrial y su dinamismo demográfico han impulsado las relaciones entre los municipios del eje del corredor con otros de la periferia metropolitana, experimentando sustanciales desarrollos urbanos a lo largo del último siglo. Así, en los últimos 30 años, la población cercana al límite entre Madrid y Guadalajara se ha incrementado en un 40 % como causa del proceso de urbanización (García Rodríguez & Pérez González, 2014).

Ya en el periodo 1988 – 2018, esta zona contaba con un índice de crecimiento medio de superficie urbana del 3,4 % anual¹, destacando el fuerte auge del uso residencial unifamiliar e industrial. Desde el punto de vista de la planificación existente, el gobierno regional de Madrid ha desarrollado diversos planes estratégicos (Plan Estratégico de movilidad sostenible de la Comunidad de Madrid (2013–2025) y Plan Estratégico/Económico Activa Henares (CRTM,

1 Datos propios obtenidos a partir de la cartografía vectorial de la Dirección General del Catastro (2018).

2015), además del surgimiento de otros proyectos logísticos (ALMA Henares) de gran envergadura en el Corredor, con importantes repercusiones territoriales, especialmente en temas de transporte, infraestructura y vivienda. Este fuerte dinamismo hace que el presente caso de estudio resulte de especial interés para aplicar metodologías prospectivas que permitan determinar los principales factores que condicionan los cambios y distribución de los usos del suelo y los sistemas de transporte. Al mismo tiempo, se ofrece la oportunidad de comparar posibles escenarios de desarrollo futuro con la planificación trazada por las entidades competentes.

Figura 1. Localización del caso de estudio



Fuente: elaboración propia sobre datos de BCN200 (2010) y CLC (2018)

2.2 Datos de partida

Los datos de partida fueron obtenidos de un taller de cartografía colaborativa celebrado en junio de 2019, partiendo de tres narrativas para el año 2050, generadas previamente a partir de un conjunto de encuestas a la ciudadanía (Soria-Lara et al., 2021), que ofrecían distintas líneas

futuras de desarrollo del Corredor del Henares, con diferentes grados de disrupción. Un grupo de expertos en diversos campos relacionados con la planificación urbana y el transporte, colaboraron conjuntamente para ubicar espacialmente los posibles desarrollos urbanos asociados a cada narrativa. En la Tabla 1 se describen brevemente estas tres narrativas (y sus implicaciones espaciales).

Tabla 1. Aspectos destacados de las narrativas de los escenarios futuros

Escenario	Descripción
Centros históricos no motorizados	Esta narrativa plantea un corredor con alta mezcla de funciones incluyendo residencia, compras, ocio y áreas industriales o empresariales. Se transformarían las actuales áreas industriales y tecnológicas en lugares más multifuncionales, a la vez que los centros urbanos recuperarían una movilidad netamente no motorizada, lo que propiciaría un mayor desarrollo de viviendas en altura para incrementar la densidad urbana y reducir distancias de viaje. Se establecería una amplia red de espacios verdes que articularían todo el corredor. Las estaciones de transporte público se localizarían en superficie y estarían fuertemente integradas con la red de espacios verdes. Caminar e ir en bicicleta serían los modos preferentes para la población residente. El coche tendría un papel residual, siendo más importante y usado el transporte público .
Un corredor superpoblado	También caracterizado por una alta mezcla de funciones residenciales, ocio, compras e industria y, por tanto, con ausencia de espacios monofuncionales. Se trataría de un territorio polarizado : centros urbanos ocupados por rentas más bajas en bloques de mediana altura y periferia ocupada por rentas más altas en viviendas unifamiliares. Se establecería una amplia red de espacios verdes que articularían todo el corredor. Las estaciones de transporte público estarían en superficie y fuertemente integradas con la red de espacios verdes. Se reduciría el uso del coche privado, mientras que el transporte público tendría niveles similares a los actuales, con preferencia por los sistemas ferroviarios .
Altos niveles de inseguridad ciudadana	En este escenario se produciría una baja mezcla de funciones como residencia, ocio, compras e industria. Predominarían los espacios monofuncionales. En los centros urbanos prevalecería el comercio, los servicios y las viviendas para clases más bajas. Los estratos con rentas más altas de la población aspirarían a vivir en la periferia, en viviendas unifamiliares distribuidas en urbanizaciones cerradas y protegidas. Los espacios verdes públicos no tendrían protagonismo en el corredor. Aumentaría el uso del vehículo privado y disminuirían las zonas de movilidad peatonal o en bicicleta.

Fuente: elaboración propia

La localización de los cambios de uso urbano en el taller fue posible gracias al empleo de una cartografía base compuesta por una cuadrícula de 250 m x 250 m de tamaño de celda, mostrando el uso del suelo que se pretendía representar espacialmente: servicios y equipamientos, industrial, residencial unifamiliar, residencial multifamiliar y residencial mixto. Esta cartografía fue utilizada como lenguaje común para la comunicación entre los diferentes expertos (arquitectos, promotores, planificadores del transporte, etc.) que participaron en el taller (Molinero-Parejo et al., 2021). Tras su desarrollo, se obtuvieron los mapas que incorporaban los nuevos crecimientos, los cambios y las pérdidas de cada uso. Dado que su representación mostraba una información generalizada debido al tamaño de la celda, y con el objetivo de

generar un conjunto de datos a escala local, se digitalizó la información, representándola en el parcelario catastral a partir de la intersección de las nuevas celdas con las parcelas correspondientes, así como las nuevas infraestructuras viarias diseñadas por los expertos (Figura 2). De esta manera, se obtuvieron las variables dependientes que fueron utilizadas para cada modelo (crecimiento y cambios de cada uso). Las variables independientes se derivaron igualmente a partir de la cartografía resultado del taller. En este punto, cabría resaltar que la parcela catastral ha sido utilizada como base para aplicar la metodología desarrollada. Su estructura irregular aporta un mayor realismo en cuanto a la división del espacio geográfico, proporcionando información detallada de la propiedad del suelo y siendo utilizada como unidad de referencia espacial del planeamiento urbanístico español (Barreira-González et al., 2015).

2.3 Metodología

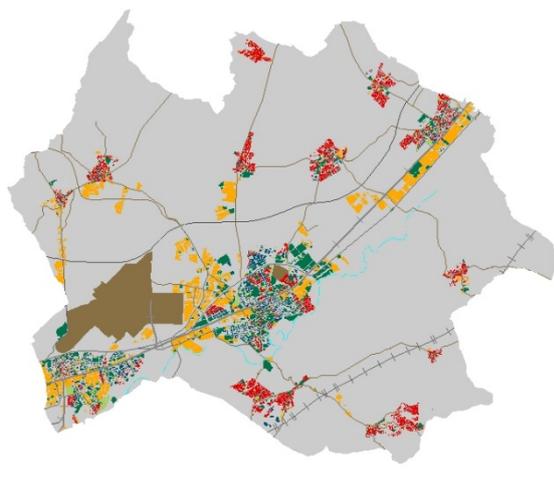
Para llevar a cabo el modelado de los resultados del taller mediante RLGP, es imprescindible concretar dos aspectos fundamentales en el proceso de representación espacial de cambios de usos del suelo, como son la determinación de la cantidad de cambio y la localización del mismo. A este respecto, en el presente trabajo, la determinación de la cantidad de superficie fue obtenida a partir de la opinión de los expertos para cada escenario (Tabla 2), mientras que la localización de los nuevos desarrollos vendrá determinada por el potencial de transición del parcelario para cada uso. Por tanto, para identificar aquellas parcelas de mayor aptitud y determinar qué factores explicativos tuvieron mayor influencia en la localización de los usos del suelo para cada escenario, se aplicó el modelo de Regresión Geográficamente Ponderado (RGP, en inglés *Geographically Weighted Regression*, GWR) desarrollado por Brunsdon et al. (1996). La aplicación de esta metodología se justifica principalmente por el carácter espacial de los datos, ya que los modelos RGP se utilizan, precisamente, para observar relaciones espaciales que los modelos de regresión logísticos no permiten estudiar (Fotheringham, Brunsdon & Charlton, 2002).

Para la aplicación de esta técnica, se crearon modelos únicamente para los usos que experimentaron crecimiento en cada escenario, no considerando aquellos que no presentaron un aumento de superficie (estabilidad o pérdida; ver Tabla 2). De este modo, se desarrollaron ocho modelos, tres para el escenario *Centros históricos no motorizados* (para modelar los usos del suelo “Servicios y equipamientos”, “multifamiliar” y “mixto”), cuatro para el escenario *Un*

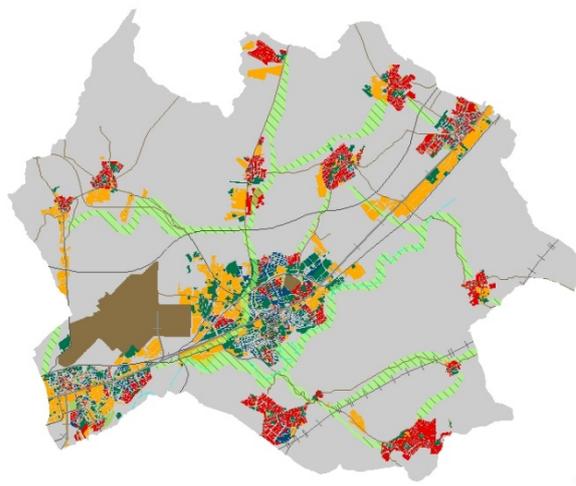
corredor superpoblado (usos "Servicios y equipamientos", "unifamiliar", "multifamiliar" y "mixto") y uno para el escenario *Altos niveles de inseguridad ciudadana* (uso "unifamiliar").

Figura 2. Representación espacial de los escenarios futuros

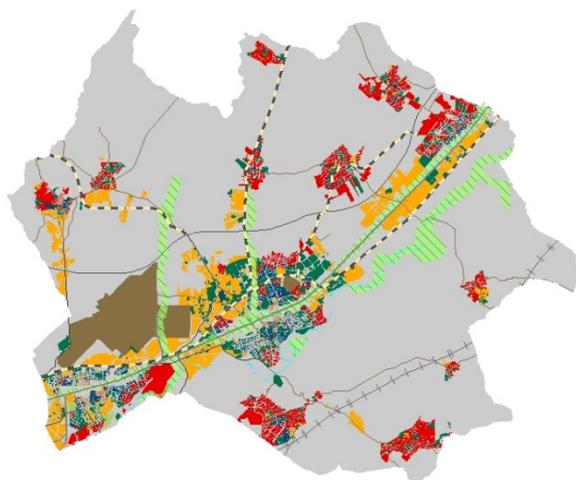
Situación inicial (2018)



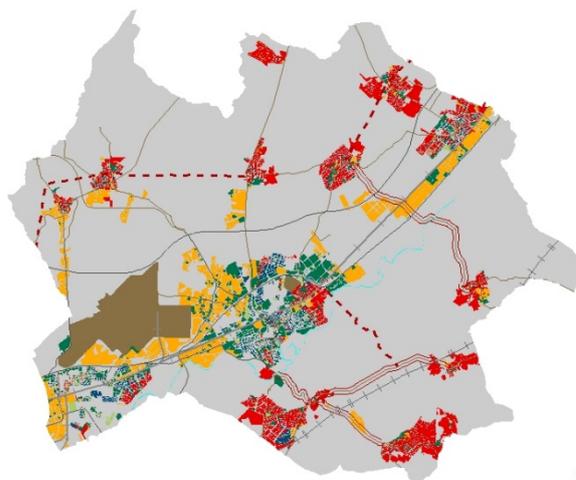
Centros históricos no motorizados (2050)



Un corredor superpoblado (2050)



Altos niveles de inseguridad ciudadana (2050)



Usos del suelo urbano

- | | |
|---------------------------|-------------------|
| Servicios y equipamientos | Residencial mixto |
| Industrial | Militar |
| Residencial unifamiliar | Zonas verdes |
| Residencial plurifamiliar | Corredor verde |

Infraestructuras de transporte

- | | |
|----------------------|-----------------|
| Autopista / Autovía | Línea Bus VAO |
| Carretera autonómica | Nueva carretera |
| Ferrocarril / AVE | Desdoblamiento |
- ETRS 1989 UTM Zona 30N



Fuente: elaboración propia

Tabla 2. Evolución de los usos por escenario para el periodo 2018–2050

Uso	Escenario		
	Centros históricos no motorizados	Un corredor superpoblado	Altos niveles de inseguridad ciudadana
Servicios y equipamientos	Crece moderadamente	Crece moderadamente	Estable
Industrial	Estable	Estable	Pierde ligeramente
Residencial unifamiliar	Pierde ligeramente	Crece moderadamente	Crece intensamente
Residencial multifamiliar	Crece ligeramente	Crece ligeramente	Estable
Residencial mixto	Crece ligeramente	Crece ligeramente	Pierde ligeramente

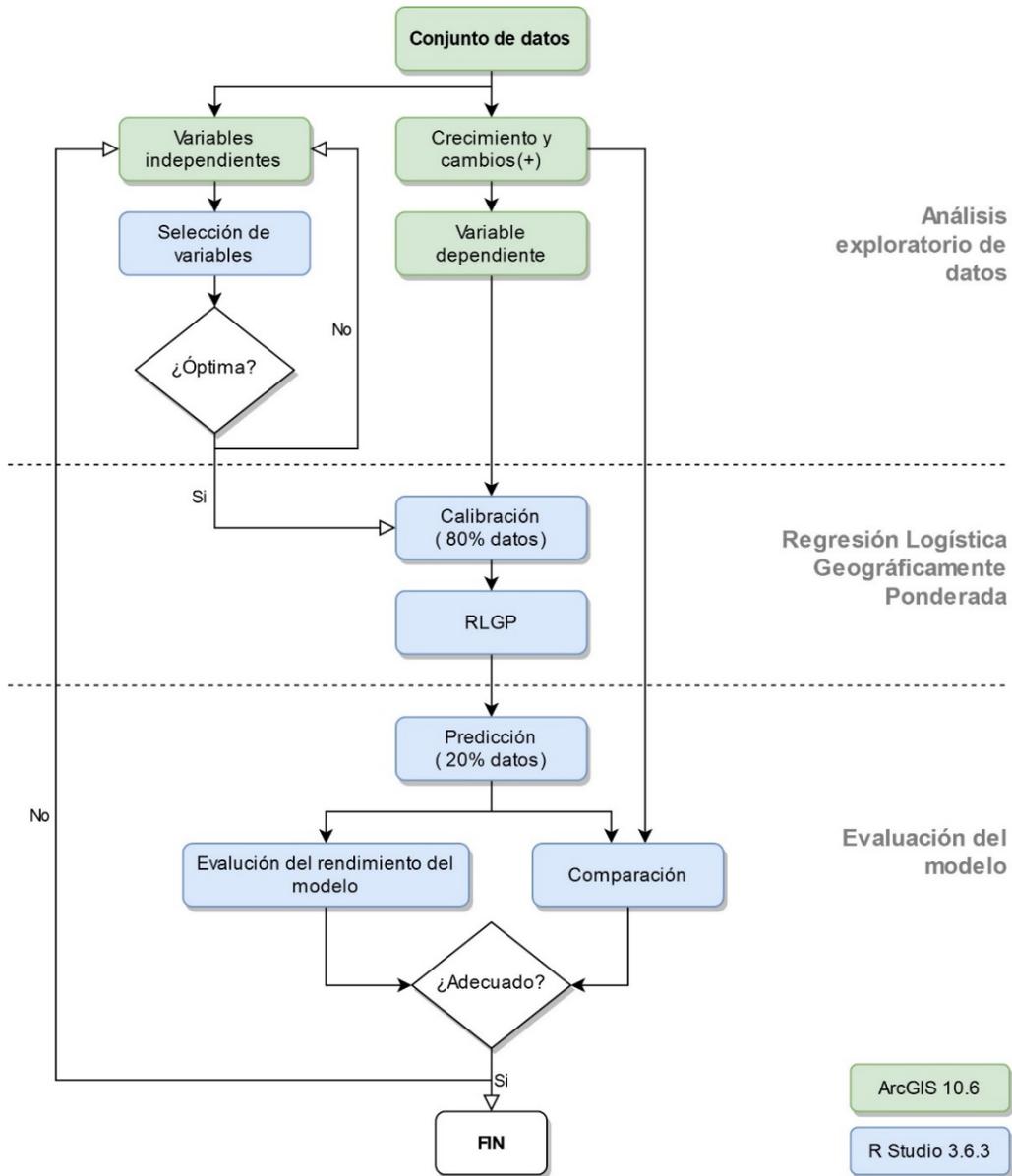
Nota: *En negrita aquellos modelos desarrollados para los usos del suelo con crecimientos

Fuente: elaboración propia

Todos los modelos desarrollados en esta investigación han sido ejecutados mediante R (<https://www.r-project.org/>) en su versión 3.6.3 y su IDE RStudio (<https://rstudio.com/>). Para ello, la principal librería utilizada ha sido GWmodel (Gollini, Lu, Charlton, Brunsdon & Harris, 2015), un paquete que contiene técnicas de una rama de la estadística espacial, concretamente modelos de ponderación geográfica.

La metodología aplicada en la presente investigación sigue la secuencia mostrada en la Figura 3, que se resume en varios bloques: Análisis Exploratorio de Datos previo al análisis de regresión y selección de variables independientes; elaboración de los modelos logísticos geográficamente ponderados; y evaluación de los modelos obtenidos.

Figura 3. Flujograma metodológico



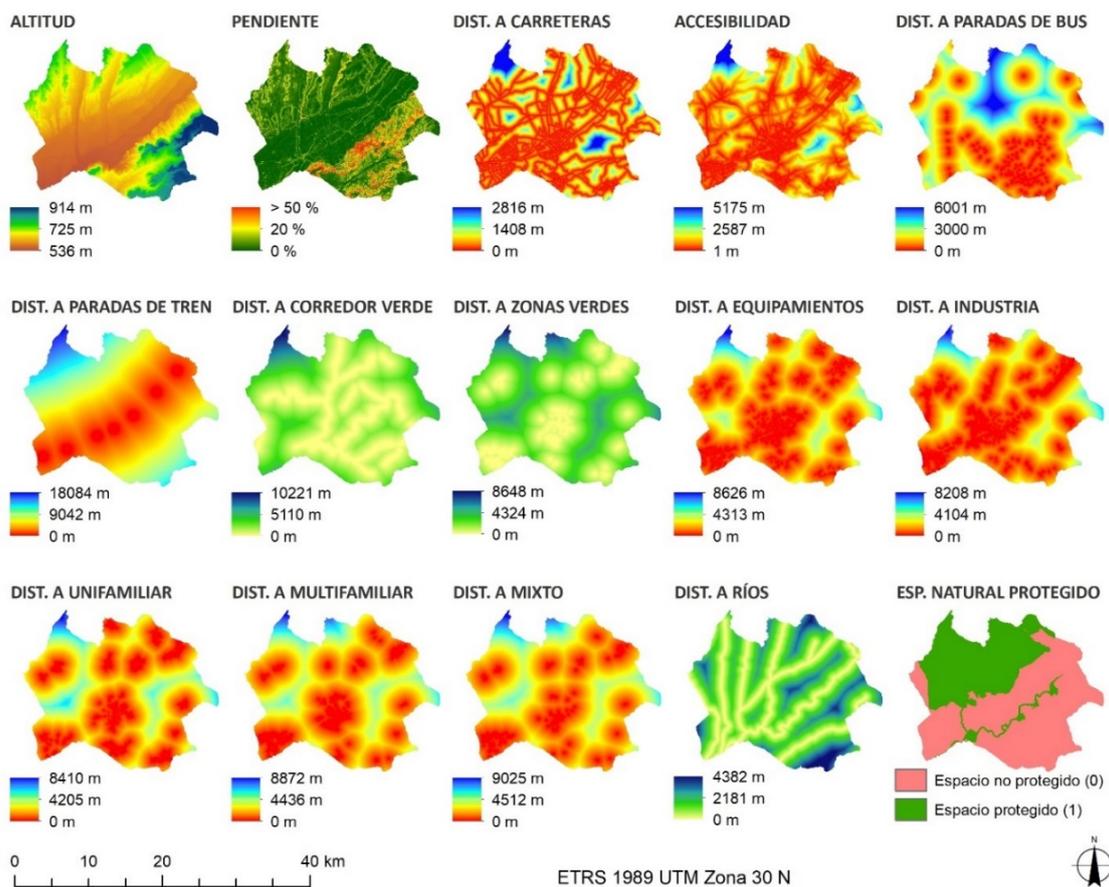
Fuente: elaboración propia

a) *Análisis Exploratorio de Datos (AED) y selección y cartografiado de variables independientes*

La selección de las variables explicativas del fenómeno de expansión urbana se realizó, en primer lugar, a partir de una revisión bibliográfica para identificar los factores que pudiesen expresar las relaciones espaciales con dicho fenómeno para cada uso (Barreira-González et al., 2017; Guanglong et al., 2017; Mirbagheri & Alimohammadi, 2017). Concretamente, este proceso de selección se ha visto apoyado en diversos estudios que identifican un amplio abanico de variables explicativas del crecimiento urbano para la Comunidad de Madrid (Díaz-Pacheco & Hewitt, 2013; Plata Rocha et al., 2011). Variables como la altitud, pendiente o la distancia a las principales vías de comunicación están presentes también en numerosos estudios

internacionales (Guanglong et al., 2017; Hu & Lo, 2007; Shafizadeh-Moghadam, 2019; Shafizadeh-Moghadam & Helbich, 2015). Por su parte, la accesibilidad fue calculada de dos formas: mediante la distancia desde un punto hasta la sección de la red viaria más cercana (variable Dist. a carreteras) y, de una manera más realista, a través de la fricción o coste de desplazamiento en función del nivel jerárquico de las distintas carreteras teniendo en cuenta la velocidad límite establecida (variable Accesibilidad). De manera complementaria, para aportar mayor precisión al modelo, se analizaron otras variables fruto de la discusión entre los participantes del taller colaborativo, así como de la observación de los factores tenidos en cuenta a la hora de asignar un nuevo uso por parte de los expertos. En la Figura 4 se observan todos los factores inicialmente recogidos y cartografiados mediante TerrSet 18.21 y ArcGIS Desktop 10.6.

Figura 4. Variables independientes recogidas



Fuente: elaboración propia

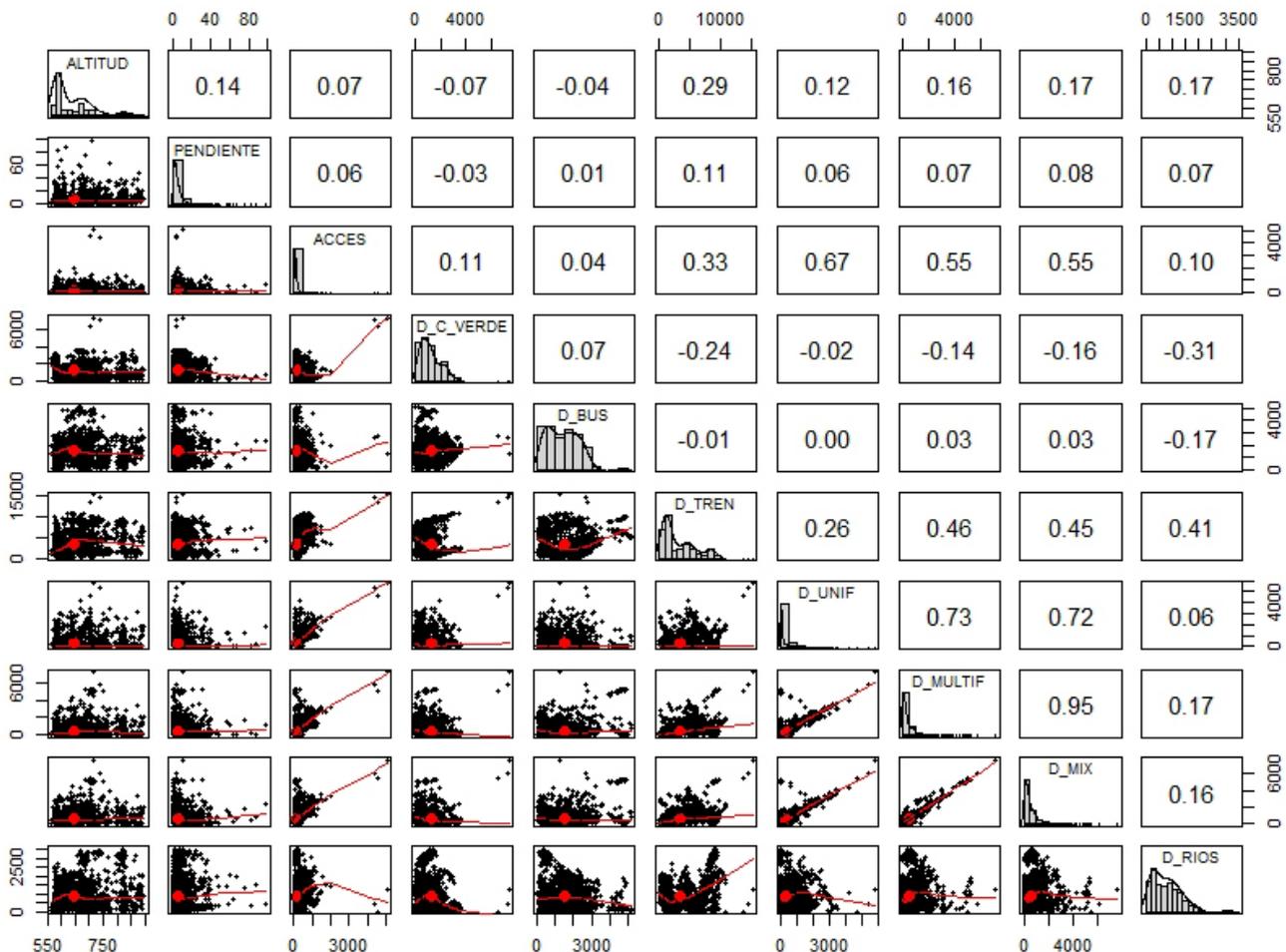
A partir de este conjunto de variables se realizó un AED con la finalidad de generar una base de datos robusta, evitando valores ausentes, repetidos o fuera de rango y buscando aquellas variables que presentaban relación con la variable respuesta, es decir, observando qué factores podrían intervenir en el proceso de crecimiento urbano. Otro aspecto relevante considerado es la presencia o ausencia de autocorrelación espacial en las variables independientes. Esta frecuente característica de los datos espaciales no se recoge en los modelos globales (Duque, Velásquez & Agudelo, 2011), por lo que es determinante a la hora de elegir el modelo de regresión que se debe aplicar.

Mediante una matriz de correlaciones bivariadas (Figura 5) se excluyeron aquellas variables que presentaban una elevada correlación para evitar la multicolinealidad, tomando como valor crítico $\pm 0,7$, umbral a partir del cual el par de variables se consideraría redundante y podría dar lugar a un modelo inestable y poco confiable (Clark & Hosking, 1986). Una correlación moderada o fuerte entre dos variables explicativas revela una elevada interdependencia. En este sentido, dicha correlación de los coeficientes de regresión locales invalidaría cualquier interpretación que pueda hacerse sobre el modelo, pudiendo derivar en conclusiones erróneas (Wheeler & Tiefelsdorf, 2005).

b) Variable dependiente

La variable dependiente de cada modelo se obtuvo de la digitalización de los mapas resultantes del taller, de los cuales se extrajo una muestra que el modelo RLGP pudiese procesar de manera eficiente. Así, el conjunto de datos seleccionado se estructuró en función de la cantidad de crecimiento estipulada para cada escenario. El número de puntos varía desde 435, para el modelo de uso "residencial mixto" del escenario *Centros históricos no motorizados*, a 6900 puntos para el modelo de "servicios y equipamientos" del escenario *Un corredor superpoblado*. Posteriormente, se hizo una división para realizar la calibración del modelo, recogiendo un 80 % de los datos para el entrenamiento, y un 20 % restante para la evaluación de los resultados. Cabe mencionar que los datos procedentes de parcelas catastrales fueron recogidos a partir del centroide de estas, configurando un marco de datos por puntos, a partir de los cuales se extrajo el valor por localización de las variables independientes procesadas en formato ráster.

Figura 5. Matriz de correlaciones bivariadas de las variables independientes



Fuente: elaboración propia

Así, en primer lugar, dichos crecimientos se binarizaron, marcando como (1) los puntos que no tenían uso asignado y habían recibido un nuevo uso concreto (crecimiento) o aquellos que anteriormente pertenecían a un uso, pero este había sido modificado, y (0) aquellos puntos que mostraron persistencia en el periodo en cuestión (2018–2050), ya presentarían un uso concreto o no. Un total de ocho conjuntos de datos fueron seleccionados para la aplicación del modelo.

Una cuestión para tener en cuenta fue la aplicación del modelo, no al total de suelo urbano, sino a los distintos usos del suelo urbanos seleccionados en un espacio metropolitano extenso como el caso de estudio escogido, ya que los puntos de datos que marcan crecimiento, en ocasiones apenas representaban el 1% del total, por lo que se requirieron medidas para adecuar el conjunto de datos al modelo asignado. Así, dada la gran cantidad de parcelas que componen el caso de estudio en referencia a las parcelas que indicaban crecimiento, los conjuntos de datos tuvieron que balancearse. Se recogió el total de parcelas que marcaban crecimiento (valor 1) para cada uso concreto (hasta cierto límite para agilizar el procesado computacional),

seleccionando posteriormente la misma cantidad de parcelas que no aplicaban crecimiento (valor 0) de forma aleatoria. De esta manera, el número de puntos codificados como 1 era equivalente al número de puntos codificados como 0 (Guanglong et al., 2017).

c) Regresión Logística Geográficamente Ponderada (RLGP)

Los modelos globales, como pueden ser los modelos de regresión logística, generan un único conjunto de coeficientes para el área total analizada, expresando que un factor tiene el mismo efecto en todo el área estudiada (Guanglong et al., 2017). Se entiende, por consiguiente, que los tradicionales modelos de regresión logística asumen el mismo poder explicativo para todo el conjunto de datos, presentando homogeneidad en los datos espaciales, independientemente de su localización (Pineda Jaimes et al., 2010). La ecuación del modelo logístico se traduce en:

$$y_i = \ln \left(\frac{P_i}{1 - P_i} \right) = \beta_0 + \sum_n \beta_n x_n + \varepsilon_i$$

donde la variable dependiente y_i es igual al \ln de la probabilidad de ocurrencia del fenómeno de crecimiento de un determinado uso P_i , dividido entre la probabilidad de no ocurrencia $1 - P_i$, lo que se conoce como *Odds*. β_0 es el intercepto, β_n el coeficiente de regresión, x_n el valor de la variable explicativa n y ε_i es el error aleatorio en un punto i .

Asumiendo que las relaciones varían a lo largo del espacio, los modelos anteriormente mencionados pueden ocultar variaciones espaciales potencialmente interesantes, impidiendo un ajuste más preciso de los datos (Fotheringham et al., 2002). En este sentido, los datos espaciales presentarán una cierta variabilidad dependiendo de su ubicación. De este modo, para ajustar la fórmula anterior teniendo en cuenta la ubicación de los puntos de datos según Fotheringham et al. (2002), RLGP se presenta como:

$$y_i = \beta_0(u_i, v_i) + \sum_n \beta_n(u_i, v_i)x_n + \varepsilon_i$$

donde la diferencia con respecto al modelo global se observa en (u_i, v_i) , que representan las coordenadas para un punto i . Este modelo ajusta una regresión para cada punto teniendo en cuenta un número determinado de vecinos cercanos (Fotheringham et al., 2002). La distancia a partir de la cual se establece el número óptimo de vecinos más próximos considerados en el modelo se establece a partir del ancho de banda, el cual puede ser adaptativo o fijo. En el ancho de banda fijo, el modelo es calibrado teniendo en cuenta una distancia fija a partir de cada punto, mientras que el ancho de banda adaptativo recoge el número de vecinos más próximos, por lo que la distancia variará dependiendo de cada punto de regresión local y de la

distribución espacial del conjunto de datos, es decir, el *kernel* presentará un ancho de banda mayor en aquellas ubicaciones donde se dé una dispersión de puntos elevada, y un ancho de banda menor donde se dé una abundancia de puntos (Fotheringham et al., 2002).

Al mismo tiempo, el modelo permite asignar una mayor ponderación a aquellos puntos más cercanos respecto al punto de regresión local (Fotheringham et al., 2002). Para ajustar los pesos asignados, existen distintas funciones *kernel* (*Gaussian*, *Exponential*, *Box-car*, *Bi-square*, *Tri-cube*). En la presente investigación se ha aplicado la función *Gaussian*, que ofrece una ponderación máxima de 1, decreciendo según una distribución de tipo gaussiana conforme aumenta la distancia entre el punto observado y el calibrado (Gollini et al., 2015):

$$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right)$$

donde d_{ij} es la distancia entre el punto i y el punto j y b es el ancho de banda. De esta manera, se obtiene el peso asignado de la diagonal de la matriz de pesos geográficos w_{ij} .

Una vez seleccionado el tipo de ancho de banda y la función *kernel* a aplicar, para calibrar la función de ponderación espacial, las aproximaciones más frecuentes son *cross-validation* (CV) (Cleveland, 1979) y *Akaike Information Criterion* (AIC) (Akaike, 1973). En este caso, debido a su aplicación más generalizada en modelos de tipo *Poisson* y logísticos (Fotheringham et al., 2002), se escogió la minimización del valor AIC, en su versión corregida (AICc), que toma un mejor ajuste ante muestras de pequeño tamaño (Hurvich & Tsai, 1989).

d) Evaluación del modelo RLGP

El primer método para validar el modelo RLGP consistió en el empleo del valor pseudo R^2 de la calibración. Dicha información permite obtener una primera aproximación del grado de ajuste del modelo local (Rodrigues, de la Riva & Fotheringham, 2014). A su vez, para observar si existen problemas de inferencia derivados de la subestimación de los errores estándar de los parámetros (Fotheringham et al., 2002), se analizaron los residuos del modelo aplicando la medida estadística I de Moran.

Por otra parte, y tomando como datos de partida el 20% del conjunto total reservado, se realizó una matriz de confusión para evaluar el desempeño de cada modelo y observar así su capacidad predictiva. Partiendo de la probabilidad de crecimiento estimada por cada modelo, se estableció un valor de corte para determinar el umbral a partir del cual se consideraba valor 0 o valor 1, intentando obtener los valores que mejor se acercasen a la variable observada. Una

vez establecido el umbral y construida la matriz, se calculó una serie de métricas que ayudaron a evaluar mejor cada modelo.

La *exactitud* indica la proporción entre los crecimientos estimados correctamente y el total de casos que el modelo consideraba que crecerían. La *precisión*, por su parte, indica un buen ajuste del modelo, y se traduce en la proporción entre el número de estimaciones acertadas frente al total.

Tras el cálculo de las métricas de la matriz de confusión, se utilizó finalmente como técnica complementaria de validación la curva ROC (*Relative Operating Characteristic*), muy utilizada para evaluar modelos que pretenden clasificar algún tipo de fenómeno o evento (Mas, Soares Filho, Pontius, Farfán Gutiérrez & Rodrigues, 2013), y especialmente para cambios en la cobertura y usos del suelo (Hu & Lo, 2007; Mas et al., 2013).

Si la variable a estimar es binaria, se parte de la idea de que, si no se utilizara ningún modelo, la probabilidad de acierto sería del 50% (0 o 1). Esta situación se ve marcada con una diagonal al graficar la curva ROC, también denominada línea de no discriminación, marcando un índice uniforme (Pontius & Parmentier, 2014). La propia curva se construye a partir de la tasa de verdaderos positivos, indicada por la sensibilidad del modelo, representada frente al valor obtenido al restar una unidad a la especificidad (tasas de falsos positivos). Lo que permite esta técnica es revelar la capacidad discriminatoria del modelo en cuestión, es decir, la habilidad para distinguir las ubicaciones que muestran crecimiento y las que no. Dicho parámetro se calcula a partir del AUC (*Area Under Curve*) y suele utilizarse para comparar mapas de probabilidad (Mas et al., 2013).

3 Descripción y análisis de resultados

3.1 Resultados del AED

Tras seleccionar el conjunto de factores como posibles predictores para cada modelo, se introdujeron aquellos que pasaron el filtro de la metodología aplicada. Todos ellos fueron estandarizados entre 0 y 1, permitiendo la comparación entre factores y la posibilidad de estudiar interacciones en el caso de que existiesen. La Tabla 3 muestra una breve descripción de las variables introducidas. Cabría destacar que estas variables seleccionadas no fueron aplicadas a un mismo modelo, sino que conforman el total de las escogidas en los ocho modelos.

Tabla 3. Descripción y fuente de las variables escogidas para el estudio

Variable	Descripción	Fuente
ALTITUD	Altura sobre el nivel del mar	MDT25 (IGN ¹)
PENDIENTE	Grado de inclinación del terreno en relación con la horizontal	MDT25 (IGN ¹)
ACCES	Coste de desplazamiento desde cualquier punto de la red viaria a cualquier punto del territorio, medido en función del tipo de vía (convencional, doble carril, autovía, autopista)	BCN100 (IGN ¹)
D_C_VERDE	Distancia a nuevos corredores verdes diseñados por los expertos	Taller ²
D_BUS	Distancia a las paradas de autobús (urbanas e interurbanas)	CRTM ³
D_TREN	Distancia a las estaciones y apeaderos de la red ferroviaria de cercanías	CRTM ³
D_UNIF	Distancia al uso residencial unifamiliar	DGC ⁴
D_MULTIF	Distancia al uso multifamiliar (bloques en altura)	DGC ⁴
D_MIX	Distancia al uso residencial mixto (viviendas y comercio o servicios)	DGC ⁴
D_RIOS	Distancia a la red hidrográfica de cursos permanentes	BCN100 (IGN ¹)

Nota: 1 Instituto Geográfico Nacional / 2 Taller de Cartografía Colaborativa. Proyecto TRANSURBAN / 3 Consorcio Regional de Transportes de Madrid / 4 Dirección General del Catastro. Ministerio de hacienda.

Fuente: elaboración propia

De entre las variables recogidas, la distancia a carreteras, la distancia a zonas verdes urbanas, la distancia a los usos servicios y equipamientos e industrial y la presencia/ausencia de espacios naturales protegidos no fueron seleccionadas en ningún modelo debido a que presentaban una elevada correlación positiva o negativa ($> 0,7$ o $< -0,7$) con otras variables escogidas de mayor interés. En determinados casos, es posible observar una elevada correlación entre dos o más variables seleccionadas tras el AED. La distancia al uso residencial mixto y la distancia al uso residencial multifamiliar presentan un valor de correlación de 0,95. Del mismo modo ocurre con la distancia al uso residencial unifamiliar. Esta cuestión se debe a que estas variables no se utilizaron de forma conjunta en un único modelo, sino que se introdujeron en distintos modelos. Así mismo, todas las variables descritas en la tabla tres presentaban autocorrelación espacial.

3.2 Resultados de la RLGP

A continuación, se muestran los resultados principales de los modelos destinados a explicar el crecimiento de cada uno de los usos en función de los tres escenarios.

En la Tabla 4 se observan las variables dependientes seleccionadas para cada modelo, que aportan valor explicativo al proceso de crecimiento de cada uso. En todos los casos se buscó aquel modelo acorde al principio de parsimonia, minimizando el valor *AICc* e intentando

explicar el mayor porcentaje de variabilidad del modelo, haciéndolo lo más sencillo posible. Así, puede observarse que el número de variables utilizadas varía entre 3 y 5.

Tabla 4. Variables utilizadas para cada uso en el modelo RLGP

	Escenario 1			Escenario 2				Escenario 3
	EQP	MLTF	MIX	EQP	UNIF	MLTF	MIX	UNIF
ALTITUD		✓	✓			✓	✓	
PENDIENTE						✓		
ACCES						✓		✓
D_C_VERDE	✓	✓		✓	✓	✓		
D_BUS	✓	✓	✓	✓	✓			
D_TREN			✓	✓	✓		✓	✓
D_UNIF					✓		✓	✓
D_MULTIF				✓				
D_MIX		✓						
D_RIOS	✓	✓	✓	✓				

Leyenda: EQP: Servicios y equipamientos, MLTF: Residencial multifamiliar, MIX: Residencial mixto y UNIF: Residencial unifamiliar

Fuente: elaboración propia

A modo de ejemplo, se han comparado los resultados obtenidos del ajuste del modelo para el uso “residencial unifamiliar” del escenario *Un corredor superpoblado*, tanto para el modelo global como para el modelo RLGP. A este respecto, la Tabla 5 recoge los valores resultantes de las estimaciones del ajuste para el modelo de regresión logística global y la tabla 6 para el modelo RLGP. La notable mejoría del modelo RLGP respecto a la regresión logística global, se observa en los valores pseudo R^2 (0,58 frente a 0,02) y AICc (2851 frente a 4312,63) del modelo tomado como ejemplo. Estos valores demuestran la adecuación de este enfoque para ser implementado en datos espaciales, presentando resultados difícilmente identificables en el modelo logístico global.

Tabla 5. Resultados derivados del modelo de regresión logística global
(Esc. 2 Un corredor superpoblado – Uso residencial unifamiliar)

	Estimado	Error Estd.	Valor z	Pr(> z)	Signif.
Intercepto	0,28	0,09	3,12	0,001	**
D_BUS	-0,86	0,17	-4,93	8,07E-07	***
D_TREN	1,20	0,19	6,3	2,98E-10	***
D_UNIF	1,43	0,32	4,42	9,63E-06	***
D_C_VERDE	-2,73	0,38	-7,25	4,16E-13	***

Leyenda: Códigos de signif.: 0 '***' / 0,001 '**' / 0,01 '*' / 0,05 '.' / 0,1 ''

Fuente: elaboración propia

Tabla 6. Resumen de las estimaciones generalizadas de los coeficientes de RLGP
(Esc. 2 Un corredor superpoblado – Uso residencial unifamiliar)

	Min.	1st Qu.	Median	3rd Qu.	Max.
Intercept	-27,24	-1,70	0,19	2,35	37,04
D_BUS	-31,18	-3,50	-0,59	1,92	39,97
D_TREN	-54,33	-2,38	0,51	2,95	21,91
D_UNIF	-4,57	-4,19	-0,04	4,33	56,68
D_C_VERDE	-58,92	-5,22	-0,73	3,75	83,38

Fuente: elaboración propia

Respecto a los resultados estadísticos de la calibración de los modelos (Tabla 7), el valor de R^2 ajustado varía entre 0,58, correspondiente al modelo que intenta explicar el crecimiento del uso “residencial unifamiliar” del escenario *Un corredor superpoblado*, y 0,84, correspondiente al modelo que intenta explicar el crecimiento del uso “residencial mixto” del escenario *Centros históricos no motorizados*. En todos los casos, con 5 o menos variables, se explica más de la mitad de la variabilidad del crecimiento del uso del suelo urbano en cuestión. También se observa el ancho de banda que, al ser adaptativo, recoge el número de vecinos cercanos tenidos en cuenta para cada punto del conjunto de datos, por lo que la distancia variará dependiendo de la distribución de las observaciones.

Tras aplicar el proceso de calibración, algunos de los resultados más interesantes que ofrecen los modelos RLGP son el cálculo de los coeficientes locales y los *t-value*, parámetros que pueden ser cartografiados y medidos, ofreciendo superficies que representan la variabilidad espacial y las relaciones con la variable dependiente (Matthews & Yang, 2012).

Tabla 7. Resumen de los resultados de los modelos RLGP

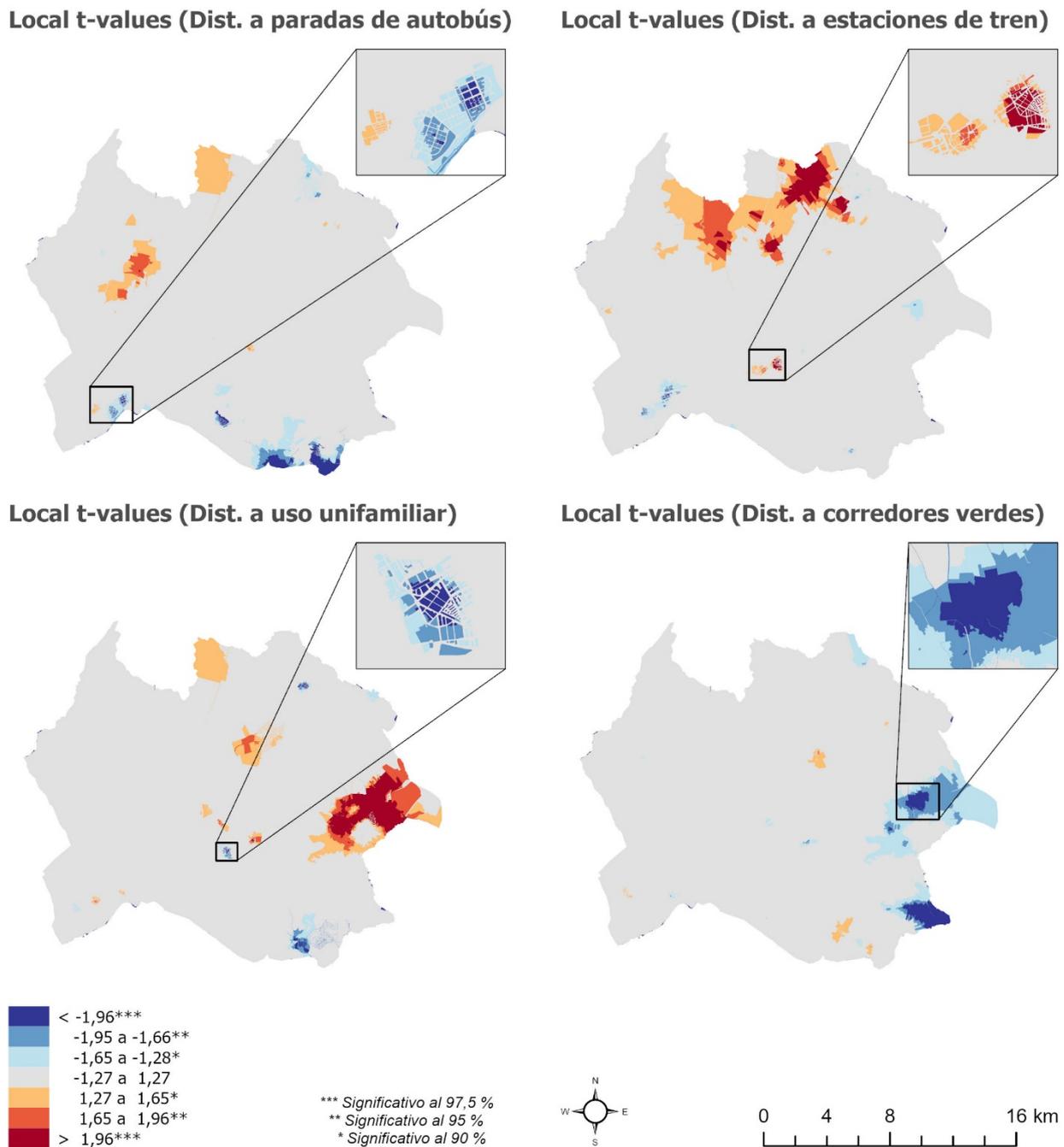
	Escenario 1			Escenario 2				Escenario 3
	EQP	MLTF	MIX	EQP	UNIF	MLTF	MIX	UNIF
<i>AICc</i> global	3735	2283	544	9483	4313	4337	3128	4310
<i>AICc</i> local	1868	1160	191	6390	2851	2702	1712	2372
<i>BW</i>	25	47	37	18	19	19	19	22
Pseudo R ² modelo global	0,02	0,12	0,11	0,01	0,03	0,02	0,01	0,02
Pseudo R ² modelo local	0,66	0,73	0,84	0,60	0,58	0,61	0,64	0,61
I. de Moran	0,06	0,07	0,11	0,01	0,04	0,04	0,03	0,03

Fuente: elaboración propia

Como resultado del ajuste de RLGP, los coeficientes de regresión locales (Anexo I) muestran variaciones espaciales para cada variable independiente. Desde una observación más general, los coeficientes presentan superficies bastante heterogéneas, marcando importantes variaciones sobre el territorio. En el caso concreto de la distancia a las paradas de autobús interurbano, se identifica una marcada dualidad entre la zona sur, y las zonas norte y este, posiblemente condicionada por la distribución de las propias paradas. Por su parte, respecto a la distancia de las estaciones de tren de cercanías, se observa un claro corredor de este a oeste, marcado por la red ferroviaria y la propia ubicación de las estaciones y apeaderos. La distancia a los usos unifamiliares presenta una fuerte influencia negativa en la zona sur, frente a una influencia positiva al norte y al oeste. Por último, los coeficientes locales de la distancia a corredores verdes indican un patrón bastante heterogéneo.

El estadístico local *t-value* representa el peso de la variable en el modelo, es decir, permite resaltar aquellas superficies donde suceden relaciones de interés (Fotheringham et al., 2002). Respecto a las cifras positivas, la probabilidad de crecimiento del uso en cuestión será mayor cuanto más elevado sea el valor de la covariable, siendo estadísticamente significativa en un intervalo de confianza del 97,5% con un *t-value* > 1,96. En contraposición, respecto a los valores negativos, a menor valor de la variable (*t-value* < - 1,96), mayor probabilidad de crecimiento. En el caso del ejemplo mostrado en la Figura 6, se observan determinadas zonas estadísticamente significativas.

Figura 6. *t-values* locales (UNIF. Esc 2)



Fuente: elaboración propia

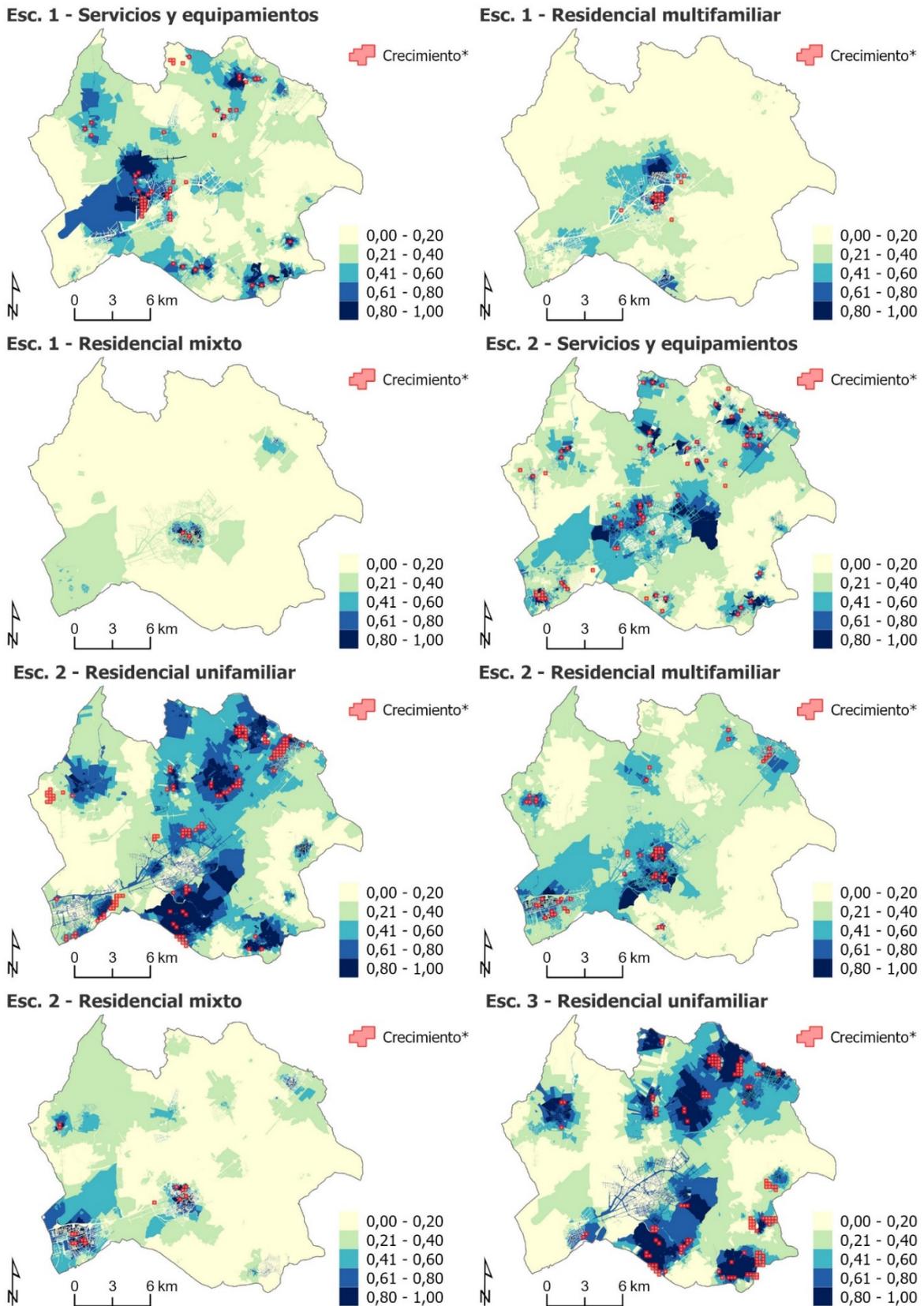
Un análisis visual puede realizarse a partir de la Figura 7, que muestra la probabilidad de crecimiento (valores entre 0 y 1) de los diferentes usos en diversos escenarios. Cabría destacar que, tanto las superficies de coeficientes de regresión locales y *t-values*, como los valores obtenidos de la probabilidad de crecimiento, fueron representados cartográficamente mediante el empleo de *IDW* (*Inverse Distance Weighted*) y la posterior asignación de valores a cada parcela mediante la extracción a través de los centroides de las mismas. El número de

ubicaciones recogidas para calibrar el modelo varía en función de la cantidad de crecimiento de la muestra de datos. Así, los usos “residencial unifamiliar” y “servicios y equipamientos” presentan un mayor crecimiento frente a otros usos, además de observarse una mayor dispersión del crecimiento en el área de estudio. En ambos casos, principalmente en los escenarios *Centros históricos no motorizados* y *Un corredor superpoblado*, la presencia de corredores verdes y la cercanía al transporte público son claros condicionantes de la expansión de ambos usos. Por otra parte, los usos “residencial mixto” y “residencial multifamiliar” muestran un menor crecimiento, ubicándose principalmente en los núcleos urbanos del eje del corredor, localizaciones de escasa elevación y pendiente. El acceso a transporte público, como la red de autobuses interurbanos y la red de tren de cercanías, demuestra una elevada significancia en ambos modelos. Por su parte, el uso “residencial unifamiliar” presente en el escenario *Altos niveles de inseguridad ciudadana* muestra un crecimiento marcado por la accesibilidad por carretera (como se describe en la narrativa), aunque también por la cercanía al transporte en tren y la propia ubicación de otras urbanizaciones o espacios residenciales unifamiliares.

En la misma Figura 7 se representaron los crecimientos de cada uso para 2050, indicados durante el taller de cartografía colaborativa, marcados por celdas rojas de 250 m x 250 m. En la mayoría de los casos analizados, las parcelas que presentan mayor probabilidad de crecimiento (color oscuro), coinciden con los crecimientos proyectados durante el taller (celdas rojas). En el caso del escenario *Centros históricos no motorizados*, los tres usos presentan unos resultados bastante precisos respecto a la ubicación. En el caso de los escenarios restantes, la coincidencia es menor, aunque también presentan buenos resultados. Visualmente, el caso que menor ajuste presenta en términos de similitud de valores estimados frente a los valores recogidos es el uso “residencial unifamiliar” del escenario *Un corredor superpoblado*, con el valor pseudo R^2 más bajo (0,58).

Los valores observados demuestran que la autocorrelación espacial es muy baja al acercarse los valores a 0. Esto indica que los puntos muestrales positivos de los valores medios de desviación multiplicados conjuntamente, es decir, de los productos cruzados, se encuentran bastante equilibrados respecto a los valores negativos. El valor más alto presentado (0,11) se observa en el modelo de crecimiento del uso “residencial mixto” del escenario *Centros históricos no motorizados*.

Figura 7. Probabilidad de crecimiento de los diferentes usos de cada escenario (2050)



Nota: * Crecimiento asignado durante el taller de cartografía colaborativa para 2050.

Fuente: elaboración propia

3.3 Resultados de la evaluación del modelo local RLGP

Respecto a una primera evaluación de los modelos, los valores R^2 presentan cifras por encima de 0,58, valores considerables para aceptar su validez. Con más detalle, extraídos de la matriz de confusión, se presentan los valores de exactitud, tasa de error, sensibilidad, especificidad, precisión, predicción negativa y AUC en la Tabla 8.

Tabla 8. Evaluación de los modelos predictivos para cada escenario

Escenario	Uso	Exactitud	Tasa de error	Sensibilidad	Especificidad	Precisión	VPN*	AUC
1	EQP	0,79	0,21	0,75	0,83	0,81	0,78	0,90
	MLTF	0,85	0,15	0,93	0,77	0,81	0,92	0,97
	MIX	0,83	0,17	0,87	0,79	0,79	0,86	0,96
2	EQP	0,78	0,22	0,83	0,74	0,76	0,82	0,88
	UNIF	0,81	0,19	0,85	0,76	0,79	0,83	0,88
	MLTF	0,81	0,19	0,90	0,74	0,75	0,89	0,92
	MIX	0,86	0,14	0,92	0,80	0,81	0,91	0,91
3	UNIF	0,87	0,13	0,91	0,83	0,81	0,92	0,93

Nota: VPN*.: Valor de Predicción Negativa.

Fuente: elaboración propia

El valor de exactitud más elevado se presenta en el escenario *Altos niveles de inseguridad ciudadana*, en el modelo de uso “residencial unifamiliar”. Por el contrario, la tasa de error más elevada se observa en el escenario *Un corredor superpoblado*, concretamente en el modelo de “servicios y equipamientos”, sin embargo, este no es el que presenta el valor R^2 ajustado más bajo (0,60). La precisión, por su parte, ofrece el valor más elevado en el uso “residencial multifamiliar” del escenario *Centros históricos no motorizados*, siendo de 0,81. Esto indica un buen ajuste del modelo, traducida en la proporción entre el número de estimaciones acertadas frente al total. De los resultados se deduce, pues, que los valores obtenidos marcan una elevada exactitud y precisión del modelo, con bajas tasas de error, por lo que el nivel de acierto de los valores predichos es aceptado para determinar los factores escogidos para cada modelo. En el Anexo II se pueden observar las curvas ROC indicando los resultados del AUC, con valores que rondan el 0,88 y 0,97, lo que se traduciría en una elevada capacidad predictiva de los diferentes modelos al acercarse a 1.

4 Discusión

El planteamiento de distintos escenarios de evolución de los usos del suelo constituye una herramienta adecuada para afrontar la incertidumbre relacionada con la planificación de crecimientos urbanos futuros. En el presente caso de estudio, el criterio y la experiencia de expertos en el campo de la planificación urbana y el transporte ha sido el principal medio para la interpretación y plasmación de narrativas de visiones futuras sobre una base cartográfica. Sin embargo, esta cuestión planteó la siguiente pregunta: ¿es posible justificar las decisiones tomadas por los expertos para validar su criterio en cuanto a la distribución localizada de los distintos usos y la red de transporte? Este artículo ha validado cuantitativamente la interpretación cualitativa de las tres narrativas propuestas, detallando estadísticamente la influencia de determinados factores explicativos en la distribución de los diferentes usos de suelo urbano y la red de transporte.

Entendiendo que el conjunto de factores que explican parte del crecimiento de un determinado uso varía en función de las características para ese mismo uso descritas en cada escenario, los resultados han demostrado estadísticamente que las narrativas que marcaban el comportamiento de los usos del suelo y el sistema de transporte durante el taller de cartografía colaborativa fueron tenidas en cuenta por los expertos, adecuándose al patrón espacial derivado del dinamismo urbano de la región que presenta cada escenario futuro (Tabla 4). En términos generales, los factores derivados de la presente investigación han ofrecido un elevado grado de coincidencia tanto en los aspectos destacados de cada escenario, como en la información detallada de los expertos.

Por su parte, respecto a los factores identificados, los modelos locales RLGP constataron la existencia de variabilidad espacial en los factores que explican la localización de los crecimientos urbanos de acuerdo con los resultados del taller colaborativo. Observando los resultados es posible deducir que, de entre las variables físicas, la distancia a ríos y la altitud han sido relevantes en gran parte de los modelos estudiados, mientras que la variable pendiente no ha presentado tal relevancia a pesar de su aplicación en numerosos estudios (Barreira-González et al., 2017; Hu & Lo, 2007; Plata Rocha, Bosque Sendra & Gómez Delgado, 2011). Por su parte, las variables relativas al transporte público y la movilidad (distancia a estaciones de tren y autobús interurbano) presentan influencia en un mayor número de usos, dejando en un segundo plano la accesibilidad por carretera, a excepción de modelo que ajusta la distribución espacial del uso “residencial unifamiliar” en el escenario *Altos niveles de inseguridad ciudadana* y el uso

“residencial multifamiliar” en el escenario *Un corredor superpoblado*. Tanto la accesibilidad por carretera como la distancia a carreteras principales (esta última excluida por la elevada correlación entre ambas), son factores críticos en la mayoría de los estudios enfocados a comprender el crecimiento urbano (Hamdy, Zhao, Osman, Salheen & Eid, 2016; Li, Li & Wu, 2018), sin embargo, en dos de los tres escenarios no representaron valores significativos. Puede deducirse al respecto que, tanto la presencia de espacios multifuncionales que reducen teóricamente las distancias, como las infraestructuras verdes que articulan el espacio, favorecen una movilidad netamente no motorizada, tal y como describen las narrativas de los escenarios, por lo que ambos factores no resultaron determinantes.

Respecto a la distancia a corredores verdes urbanos, es interesante remarcar su efecto sobre los escenarios *Centros históricos no motorizados* y *Un corredor superpoblado*. Cabe diferenciar dicho factor explicativo de los demás por su presencia relativa en ambos escenarios. Dichas infraestructuras verdes fueron proyectadas durante el taller de cartografía colaborativa, es decir, actualmente no se han planificado, pero su desarrollo se prevé para dos escenarios en el periodo 2018 – 2050. Ello implica que determinados elementos que puedan surgir a lo largo del periodo de tiempo seleccionado podrían afectar al estado final del uso en cuestión en la fecha límite prevista (2050). A este respecto, se deduce que las actividades participativas tienen un importante papel a la hora de determinar posibles elementos que han estructurado, estructuran o podrían estructurar el espacio en un horizonte futuro. Por otra parte, el análisis de investigaciones previas sobre la evolución urbana de la Comunidad Autónoma de Madrid (CAM) permitió contrastar los resultados obtenidos en relación con otros estudios. Así, Gallardo Beltrán (2014) en su selección de variables para explicar los cambios en los usos del suelo en tres escenarios diferentes con horizonte futuro 2025, incluye factores relacionados con los propios usos del suelo, socioeconómicos y biofísicos, donde fueron incorporadas entre 4 y 7 variables, dependiendo del escenario simulado. Se pueden observar numerosas coincidencias como la altitud, pendiente, distancia a ríos o accesibilidad respecto a los usos urbano e industrial-comercial, a pesar de la diferencia de escala entre estudios y de su aplicación a usos del suelo de carácter general, frente a los usos del suelo de carácter urbano tratados en la presente investigación. Por su parte, Díaz-Pacheco & Hewitt (2013) coinciden en la selección de diversos factores para evaluar la probabilidad de cambio de usos del suelo, donde la mayor diferencia se observa en la implementación de la densidad de población. En una escala más local, Barreira-González et al. (2017) aplica un modelo de regresión logística para calcular la aptitud de una parcela respecto a su transición al uso residencial o productivo en el periodo 2000 - 2010. En él

se recogen variables como altitud, pendiente, hidrografía, espacios protegidos, usos del suelo, etc., gran parte de ellas coincidentes. Por último, variables como distancia a carreteras o distancia al uso comercial o industrial son utilizadas por Plata Rocha et al. (2011) con el objetivo de simular los usos de suelo urbano comercial, industrial y residencial para 2020. A pesar de la amplia literatura disponible, la complejidad de los sistemas urbanos dificulta la reflexión sobre cómo los factores explicativos influyen en su dinámica evolutiva. En cierta medida, dependiendo de la escala, la zona de estudio, la metodología aplicada o el enfoque al que se orientan los resultados, las variables a incluir pueden variar (Barreira-González et al., 2017), más aún si se aplican modelos enfocados a comprender fenómenos disruptivos proyectados mediante escenarios futuros.

En lo que respecta al análisis de los resultados, analizando en profundidad los modelos RLGP obtenidos, estos muestran un ajuste elevado (Tabla 6), revelando su adecuación para mejorar la comprensión de los procesos de expansión urbana reproducidos en el taller. No obstante, otros factores que podrían ser determinantes no han sido tenidos en cuenta en los modelos, como variables socioeconómicas, biofísicas, demográficas, etc., u otras difíciles de cuantificar o representar espacialmente (como por ejemplo decisiones políticas en cuanto a legislación urbanística o el propio comportamiento de la población), que influyen en la explicación del crecimiento urbano (Cantergiani & Gómez Delgado, 2020; Plata Rocha et al., 2011). En este sentido, esta delimitada selección de variables se justifica en el contexto del taller de cartografía colaborativa, donde los factores explicativos intentaron ajustarse a los elementos presentados en la cartografía disponibles para los expertos. Por otro lado, como ha ocurrido en la presente investigación y al igual que en otros estudios (Guanglong et al., 2017; Luo & Wei, 2009; Plata Rocha et al., 2011; Shafizadeh-Moghadam & Helbich, 2015), numerosas variables explicativas que podrían haber sido determinantes en la explicación del fenómeno de expansión urbana no han podido ser utilizadas debido a la disponibilidad en una escala no adecuada a la utilizada en el estudio (Gallardo Beltrán, 2014) o a la presencia de multicolinealidad y autocorrelación espacial, evitando así una redundancia en la información del modelo en cuestión (Plata Rocha, Bosque Sendra, et al., 2011).

Respecto al cálculo de aptitud o idoneidad en modelos de simulación urbana, concretamente aquellos basados en AC, la regresión logística ha sido ampliamente utilizada por su carácter empírico y su complementariedad con otras herramientas (Liao et al., 2016). Por el contrario, la aplicación de modelos de ponderación geográfica aún se encuentra poco extendida, a pesar de

ofrecer una mejora en el desempeño de las simulaciones al aplicar RLGP a la definición de las reglas de transición (Mirbagheri & Alimohammadi, 2017).

Finalmente, en relación al análisis de los residuos derivados de la calibración del modelo, para poder evaluar en profundidad dicha información es necesario partir de la hipótesis nula, donde se expone la existencia de una diferencia estadísticamente significativa entre el valor observado y el valor esperado si el atributo que se analiza está distribuido de forma aleatoria entre las entidades del caso de estudio. A pesar de que los valores residuales observados demuestran que la correlación espacial es muy baja, el *p-value* presenta cierta significancia en los modelos, lo que indica que los conjuntos de datos presentan cierta agrupación espacial, posiblemente debido a la presencia de crecimientos compactos que han dificultado la creación de un marco de datos homogéneo en el espacio. En determinadas ocasiones, el grado de correlación de los residuos puede venir dado por la localización absoluta de los datos muestrales, así como la distancia entre sus localizaciones (Fotheringham et al., 2002). Por consiguiente, la distribución de la variable dependiente (*y*) no permitió obtener un conjunto perfectamente equilibrado. Este hecho demuestra una de las dificultades que afrontan los estudios urbanos a escala local, donde la distribución de usos en el espacio presenta una elevada complejidad.

5 Conclusiones

En el contexto de la planificación de escenarios futuros, la combinación de metodologías de análisis estadístico y técnicas de comunicación participativas pueden aportar un doble criterio. Por una parte, se justifica la información derivada del taller de cartografía colaborativa por su carácter experimental y observacional y, por otra, esta investigación demuestra la validez de dichos resultados a través de métodos estadísticos. Por lo tanto, el modelo RLGP fue la metodología escogida para determinar estadísticamente qué factores explicativos influyeron en el proceso de distribución espacial de los usos del suelo urbano y el transporte para cada escenario futuro descrito. De esta manera, aprovechando su aplicación para estudiar relaciones no estacionarias entre las variables, se logró un mayor detalle de la información extraída de los datos espaciales.

En este sentido, la aplicación de RLGP a escenarios futuros ha sido una importante contribución para este estudio al reorientar la utilidad de esta herramienta hacia el análisis de un fenómeno con un horizonte a largo plazo. Hasta ahora, solo se había aplicado al presente inmediato basado en información del pasado. Este nuevo enfoque aporta información de gran valor si se persigue una planificación proactiva, especialmente aquella donde los agentes involucrados en

la toma de decisiones y los planificadores urbanos requieren información más precisa para la aplicación de instrumentos de planificación.

En cuanto al apartado metodológico de la investigación, aunque los resultados muestran un elevado ajuste del modelo local en la explicación del fenómeno de expansión urbana para los distintos usos, es importante tener en cuenta diversos aspectos que dificultan la implementación de RLGP. El proceso de selección de variables independientes es esencial para una correcta predicción, más aún cuando estas intentan explicar un fenómeno en un horizonte futuro. Del mismo modo, la muestra obtenida de la variable dependiente debe ofrecer una distribución espacial y una cantidad significativas para evitar errores durante el desarrollo del modelo. En definitiva, para adaptarse a las observaciones extraídas, se requiere una minuciosa interpretación de los resultados, siempre expresada desde el contexto que ofrece el conjunto de datos de partida.

Para investigaciones futuras, la aplicación de la metodología propuesta proporcionaría un mayor grado de detalle si el parcelario rústico presentara una subdivisión que ofrezca más realismo en cuanto a las parcelas urbanas. De este modo, las posteriores tareas del presente proyecto enfocadas a simular el crecimiento urbano mediante el desarrollo de un modelo basado en AC irregulares lograrían configurar con elevada exactitud y precisión las reglas de transición que determinan los valores de vecindad, accesibilidad y adecuación espacial.

Agradecimientos: Esta investigación se realizó en el contexto del proyecto TRANSURBAN (Simulación de escenarios colaborativos para integrar políticas de transporte urbano sostenible y usos del suelo), financiado por el Ministerio de Economía y Competitividad de España (CSO2017-86914-C2-1-P). El primer autor agradece también la ayuda prestada por el Programa Estatal de Promoción del Talento y su Empleabilidad en I+D+i (2018), financiado por el Ministerio de Ciencia, Innovación y Universidades y el Fondo Social Europeo. Además, los autores desean agradecer a los revisores anónimos sus útiles comentarios y sugerencias.

Declaración responsable: Los autores declaran que no existe ningún conflicto de interés con relación a la publicación de este artículo. Todos los autores han participado en el planteamiento de la metodología y en la redacción del artículo. El primer autor ha realizado el análisis estadístico y la producción cartográfica.

Bibliografía

Akaike, H. (1973) Information Theory and an Extension of the Maximum Likelihood Principle. In B. N. Petrov & F. Csaki (Eds.), *International Symposium on Information Theory* (pp. 267-281). Akadémiai Kiadó.

Avin, U., & Goodspeed, R. (2020). Using Exploratory Scenarios in Planning Practice: A Spectrum of Approaches. *Journal of the American Planning Association*, 0(0), 1-14. <https://doi.org/10.1080/01944363.2020.1746688>

Barreira-González, P., Aguilera-Benavente, F., & Gómez-Delgado, M. (2017). Implementation and calibration of a new irregular cellular automata-based model for local urban growth simulation: The MUGICA model. *Environment and Planning B: Urban Analytics and City Science*, 46(2), 243–263. <https://doi.org/10.1177/2399808317709280>

Barreira-González, P., Gómez-Delgado, M., & Aguilera-Benavente, F. (2015). From raster to vector cellular automata models: A new approach to simulate urban growth with the help of graph theory. *Computers, Environment and Urban Systems*, 54, 119-131. <https://doi.org/10.1016/j.compenvurbsys.2015.07.004>

Bharath, H. A., Chandan, M. C., Vinay, S., & Ramachandra, T. V. (2018). Modelling urban dynamics in rapidly urbanising Indian cities. *The Egyptian Journal of Remote Sensing and Space Science*, 21(3), 201-210. <https://doi.org/10.1016/j.ejrs.2017.08.002>

Börjeson, L., Höjer, M., Dreborg, K. H., Ekvall, T., & Finnveden, G. (2006). Scenario types and techniques: Towards a user's guide. *Futures*, 38(7), 723-739. <https://doi.org/10.1016/j.futures.2005.12.002>

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28(4), 281-298.

Cantergiani, C., & Gómez Delgado, M. (2020). Simulación del crecimiento urbano con AMEBA: Modelo Basado en Agentes para la ocupación residencial. *Boletín de la Asociación de Geógrafos Españoles*, (86). <https://doi.org/10.21138/bage.2910>

Chen, J. (2007). Rapid urbanization in China: A real challenge to soil protection and food security. *CATENA*, 69(1), 1-15. <https://doi.org/10.1016/j.catena.2006.04.019>

- Chotchaiwong, & Wijitkosum. (2019). Predicting Urban Expansion and Urban Land Use Changes in Nakhon Ratchasima City Using A CA-Markov Model under Two Different Scenarios. *Land*, 8(9), 140. <https://doi.org/10.3390/land8090140>
- Clopper, C. J., & Pearson, E. S. (1934). The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, 26(4), 404. <https://doi.org/10.2307/2331986>
- CRTM (2015). Plan Estratégico de Movilidad Sostenible de la Comunidad de Madrid 2013-2025. In *Consorcio Regional de Transportes de Madrid*. Retrieved from https://www.crtm.es/media/397998/pemscam_triptico.pdf
- Dadashpoor, H., Azizi, P., & Moghadasi, M. (2019). Analyzing spatial patterns, driving forces and predicting future growth scenarios for supporting sustainable urban growth: Evidence from Tabriz metropolitan area, Iran. *Sustainable Cities and Society*, 47, 1–15. <https://doi.org/10.1016/j.scs.2019.101502>
- Deal, B., Pan, H., Timm, S., & Pallathucheril, V. (2017). The role of multidirectional temporal analysis in scenario planning exercises and Planning Support Systems. *Computers, Environment and Urban Systems*, 64, 91–102. <https://doi.org/10.1016/j.compenvurbsys.2017.01.004>
- Díaz-Pacheco, J., & Hewitt, R. (2013). Modelado de cambios de usos de suelo urbano a través de redes neuronales artificiales. Comparación con dos aplicaciones de software. *GeoFocus*, 14, 1-22. Retrieved from <http://www.geofocus.org/index.php/geofocus/article/view/298>
- Duque, J. C., Velásquez, H., & Agudelo, J. (2011). Infraestructura pública y precios de vivienda : una aplicación de regresión geográficamente ponderada en el contexto de precios hedónicos. *Ecos de ECONOMÍA*, (33), 95–122.
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). *Geographically weighted regression. The analysis of spatially varying relationships*. West Sussex: JOHN WILEY & SONS, LTD.
- Gallardo Beltrán, M. (2014). *Cambios de usos del suelo y simulación de escenarios en la Comunidad de Madrid*. Universidad Complutense de Madrid.
- Gao, J., & Li, S. (2011). Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using Geographically Weighted Regression. *Applied Geography*, 31(1), 292–302. <https://doi.org/10.1016/j.apgeog.2010.06.003>

- García Rodríguez, M. P., & Pérez González, M. E. (2014). Análisis multitemporal del urbanismo expansivo en el corredor del Henares. Aportación de las imágenes de satélite. *Estudios Geográficos*, 75(277), 597-618. <https://doi.org/10.3989/estgeogr.201419>
- Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel : an R package for exploring spatial heterogeneity. *Journal of Statistical Software*, 63(17), 1–50. <https://doi.org/10.1080/10095020.2014.917453>
- Guanglong, D., Erqi, X., & Hongqi, Z. (2017). Urban Expansion and Spatiotemporal Relationships with Driving Factors Revealed by Geographically Weighted Logistic Regression. *Journal of Resources and Ecology*, 8(3), 277-286. <https://doi.org/10.5814/j.issn.1674-764x.2017.03.008>
- Gutierrez-Puebla, J., García-Palomares, J. C., & Daniel-Cardozo, O. (2012). Regresión Geográficamente Ponderada (GWR) y estimación de la demanda de las estaciones del Metro de Madrid. In *XV Congreso Nacional de Tecnologías de La Información Geográfica* (pp. 1-13). Retrieved from http://www.age-geografia.es/tig/2012_Madrid/ponencia2/GutierrezPuebla_final_par.pdf
- Hamdy, O., Zhao, S., Osman, T., Salheen, M., & Eid, Y. (2016). Applying a Hybrid Model of Markov Chain and Logistic Regression to Identify Future Urban Sprawl in Abouelreesh, Aswan: A Case Study. *Geosciences*, 6(4), 43. <https://doi.org/10.3390/geosciences6040043>
- Hou, H., Wang, R., & Murayama, Y. (2019). Scenario-based modelling for urban sustainability focusing on changes in cropland under rapid urbanization: A case study of Hangzhou from 1990 to 2035. *Science of The Total Environment*, 661, 422–431. <https://doi.org/10.1016/j.scitotenv.2019.01.208>
- Hu, Z., & Lo, C. P. (2007). Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems*, 31(6), 667–688. <https://doi.org/10.1016/j.compenvurbsys.2006.11.001>
- Hurvich, C. M., & Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2), 297–307. <https://doi.org/10.1093/biomet/76.2.297>
- Jokar Arsanjani, J., Helbich, M., Kainz, W., & Darvishi Bolorani, A. (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation*, 21, 265–275. <https://doi.org/10.1016/j.jag.2011.12.014>

- Li, C., Li, J., & Wu, J. (2018). What drives urban growth in China? A multi-scale comparative analysis. *Applied Geography*, 98(June), 43-51. <https://doi.org/10.1016/j.apgeog.2018.07.002>
- Liao, J., Tang, L., Shao, G., Su, X., Chen, D., & Xu, T. (2016). Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations. *Environmental Modelling and Software*, 75, 163–175. <https://doi.org/10.1016/j.envsoft.2015.10.014>
- Luo, J., & Wei, Y. H. D. (2009). Modeling spatial variations of urban growth patterns in Chinese cities: The case of Nanjing. *Landscape and Urban Planning*, 91(2), 51–64. <https://doi.org/10.1016/j.landurbplan.2008.11.010>
- Mas, J.-F., Soares Filho, B., Pontius, R., Farfán Gutiérrez, M., & Rodrigues, H. (2013). A Suite of Tools for ROC Analysis of Spatial Models. *ISPRS International Journal of Geo-Information*, 2(3), 869-887. <https://doi.org/10.3390/ijgi2030869>
- Matthews, S. A., & Yang, T.-C. (2012). Mapping the results of local statistics. *Demographic Research*, 26, 151-166. <https://doi.org/10.4054/demres.2012.26.6>
- Mirbagheri, B., & Alimohammadi, A. (2017). Improving urban cellular automata performance by integrating global and geographically weighted logistic regression models. *Transactions in GIS*, 21(6), 1280-1297. <https://doi.org/10.1111/tgis.12278>
- Molinero-Parejo, R., Aguilera-Benavente, F. Gómez-Delgado, M., & Soria-Lara, J. A. (2021). Mapping disruptive long-term scenarios using a participatory approach. *Journal of Maps* (In press).
- Mondal, B., Das, D. N., & Dolui, G. (2015). Modeling spatial variation of explanatory factors of urban expansion of Kolkata: a geographically weighted regression approach. *Modeling Earth Systems and Environment*, 1(4), 29. <https://doi.org/10.1007/s40808-015-0026-1>
- Nkeki, F. N., & Asikhia, M. O. (2019). Geographically weighted logistic regression approach to explore the spatial variability in travel behaviour and built environment interactions: Accounting simultaneously for demographic and socioeconomic characteristics. *Applied Geography*, 108, 47-63. <https://doi.org/10.1016/j.apgeog.2019.05.008>
- Olazabal, E., & Bellet, C. (2019). De la ciudad compacta a la ciudad extensa. Procesos de urbanización recientes en áreas urbanas españolas articuladas por ciudades medias. *Anales de Geografía de La Universidad Complutense*, 39(1), 149-175. <https://doi.org/http://dx.doi.org/10.5209/AGUC.64681>

- Pascual Rosa, V., Aguilera-Benavente, F., & Salado García, M. J. (2019). Clasificación y análisis de los procesos de cambio en la forma urbana de las áreas metropolitanas españolas: aplicación de métricas espaciales. *Estudios Geográficos*, 80(286), e006. <https://doi.org/10.3989/estgeogr.201926.006>
- Perveen, S., Kamruzzaman, M., & Yigitcanlar, T. (2017). Developing policy scenarios for sustainable urban growth management: A Delphi approach. *Sustainability (Switzerland)*, 9(10). <https://doi.org/10.3390/su9101787>
- Pineda Jaimes, N. B., Bosque Sendra, J., Gómez Delgado, M., & Franco Plata, R. (2010). Exploring the driving forces behind deforestation in the state of Mexico (Mexico) using geographically weighted regression. *Applied Geography*, 30(4), 576–591. <https://doi.org/10.1016/j.apgeog.2010.05.004>
- Plata Rocha, W., Bosque Sendra, J., & Gómez Delgado, M. (2011). Análisis de factores explicativos del crecimiento urbano en la Comunidad de Madrid a través de métodos estadísticos y SIG. *Geografía y Sistemas de Información Geográfica (GEOSIG)*, 3(1), 201–230.
- Plata Rocha, W., Gómez-Delgado, M., & Bosque-Sendra, J. (2011). Simulating urban growth scenarios using GIS and multicriteria analysis techniques: A case study of the Madrid region, Spain. *Environment and Planning B: Planning and Design*, 38(6), 1012–1031. <https://doi.org/10.1068/b37061>
- Radinsky, J., Milz, D., Zellner, M., Pudlock, K., Witek, C., Hoch, C., & Lyons, L. (2017). How planners and stakeholders learn with visualization tools: using learning sciences methods to examine planning processes. *Journal of Environmental Planning and Management*, 60(7), 1296–1323. <https://doi.org/10.1080/09640568.2016.1221795>
- Rodrigues, M., de la Riva, J., & Fotheringham, S. (2014). Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography*, 48, 52–63. <https://doi.org/10.1016/j.apgeog.2014.01.011>
- Shafizadeh-Moghadam, H. (2019). Improving spatial accuracy of urban growth simulation models using ensemble forecasting approaches. *Computers, Environment and Urban Systems*, 76, 91–100. <https://doi.org/10.1016/j.compenvurbsys.2019.04.005>
- Shafizadeh-Moghadam, H., & Helbich, M. (2015). Spatiotemporal variability of urban growth factors: A global and local perspective on the megacity of Mumbai. *International Journal of*

Applied Earth Observation and Geoinformation, 35, 187–198. <https://doi.org/10.1016/j.jag.2014.08.013>

Soria-Lara, J. A., Ariza-Álvarez, A., Aguilera-Benavente, F., Cascajo, R., Arce-Ruiz, R. M., López, C., & Gómez-Delgado, M. (2021). Participatory visioning for building disruptive future scenarios for transport and land use planning. *Journal of Transport Geography*, 90. <https://doi.org/10.1016/j.jtrangeo.2020.102907>

Ustaoglu, E., Williams, B., & Petrov, L. O. (2017). Scenario analysis of alternative land development patterns for the Leipzig-Halle region: Implications for transport-land-use sustainability. *Urban Planning*, 2(1), 108–129. <https://doi.org/10.17645/up.v2i1.838>

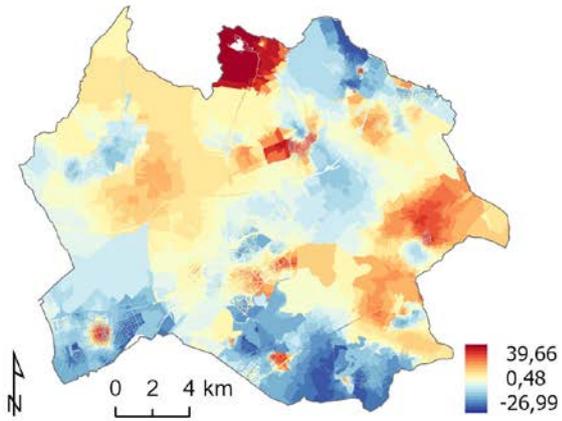
van der Heijden, K. (2000). Scenarios and Forecasting. *Technological Forecasting and Social Change*, 65(1), 31-36. [https://doi.org/10.1016/S0040-1625\(99\)00121-3](https://doi.org/10.1016/S0040-1625(99)00121-3)

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*, 7(2), 161-187. <https://doi.org/10.1007/s10109-005-0155-6>

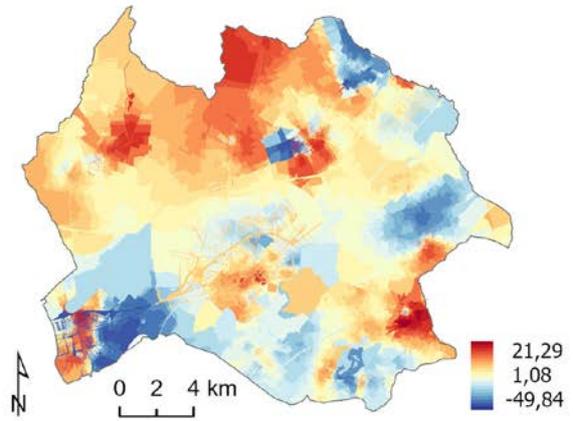
Zapata, M. A., & Kaza, N. (2015). Radical uncertainty: scenario planning for futures. *Environment and Planning B: Planning and Design*, 42(4), 754-770. <https://doi.org/10.1068/b39059>

Anexo I. Coeficientes de regresión locales (UNIF. Esc 2)

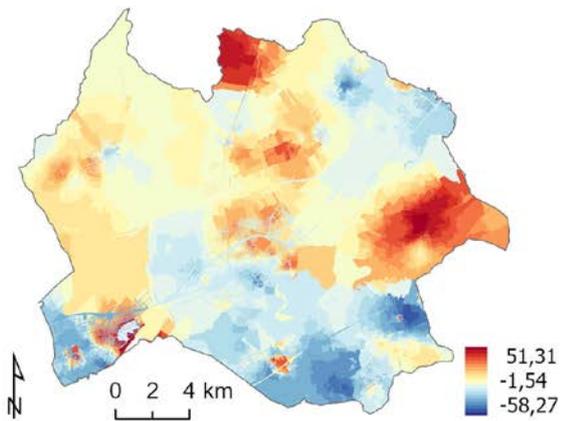
Coeficientes locales β (Dist. a paradas de autobús)



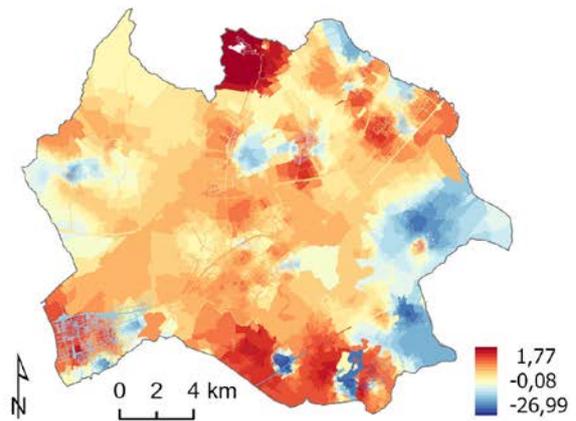
Coeficientes locales β (Dist. a estaciones de tren)



Coeficientes locales β (Dist. a uso unifamiliar)

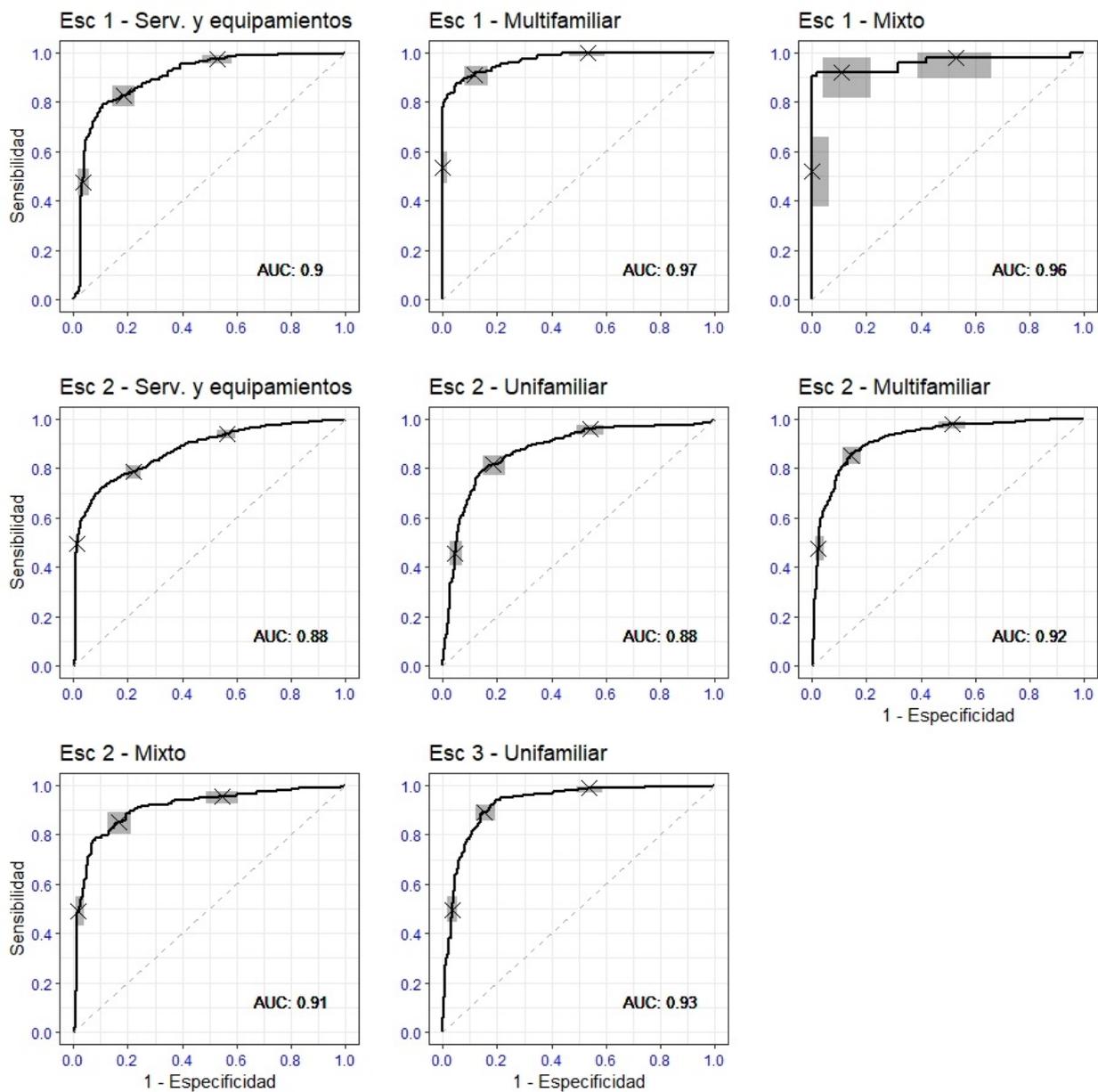


Coeficientes locales β (Dist. a corredores verdes)



Fuente: elaboración propia

Anexo II. Curvas ROC



Nota: *Las zonas sombreadas representan las regiones de confianza (nivel de significación 0.05) para los puntos de la curva ROC usando el método exacto de Clopper y Pearson (1934), recogido en la librería *plotROC*.

Fuente: elaboración propia