Universidad Politécnica de Madrid

Escuela Técnica Superior de Arquitectura de Madrid

**La aplicación de métodos estadísticos
en la valoración masiva de inmuebles:**
Una propuesta de regulación en el marco normativo
del Banco de España

Tesis Doctoral

Ana Isabel Flores Guillén. Arquitecta

2016

Departamento de Construcción y Tecnología Arquitectónicas

Escuela Técnica Superior de Arquitectura de Madrid

**La aplicación de métodos estadísticos
en la valoración masiva de inmuebles:**
Una propuesta de regulación en el marco normativo
del Banco de España

Doctorando: Ana Isabel Flores Guillén. Arquitecta

Director: Federico García Erviti

2016

POLITÉCNICA

Tribunal nombrado por el Sr. Rector Magfco. de la Universidad Politécnica de Madrid, el día.....de.....de 20....

Presidente: _____

Vocal: _____

Vocal: _____

Vocal: _____

Secretario: _____

Suplente: _____

Suplente: _____

Realizado el acto de defensa y lectura de la Tesis el día.....de.....de 20...
en la E.T.S.I. /Facultad.....

Calificación

EL PRESIDENTE

LOS VOCALES

EL SECRETARIO

ÍNDICE

PARTE I: INTRODUCCIÓN.....	12
1.1.1. INTRODUCCIÓN	13
1.1.2. JUSTIFICACIÓN	13
1.1.3. OBJETIVOS.....	14
1.1.4. HIPÓTESIS	14
1.1.5. PROPUESTA METODOLÓGICA.....	15
PARTE II: ANTECEDENTES	19
2.1. MARCO REGULATORIO EN EUROPA	20
2.1.1. RESUMEN	20
2.1.2. INTRODUCCIÓN	21
2.1.3. EMPLEO DE HERRAMIENTAS AVM EN OTROS PAÍSES	23
2.1.4. EL MARCO REGULATORIO DE LOS MODELOS AVM.....	25
2.1.5. MARCO REGULATORIO EN ESPAÑA	28
2.1.6. USO ACTUAL DE MODELOS AVM EN ESPAÑA	33
2.2. ESTADO DEL ARTE.....	36
2.2.1. INTRODUCCIÓN	36
2.2.2. ESTÁNDARES AVM	37
2.2.3. ATRIBUTOS - VARIABLES	38
2.2.4. MÉTODOS.....	43
2.2.5. MODELO	45
2.3. CONCLUSIONES PARCIALES	48
2.3.1. CONCLUSIONES MARCO REGULATORIO.....	48
2.3.2. CONCLUSIONES ESTADO DEL ARTE	49
2.3.3. ESQUEMA CONCEPTUAL.....	50
PARTE III: METODOLOGÍA	52
3.1. INTRODUCCIÓN.....	53
3.2. PLANTEAMIENTO, REQUISITOS Y FUENTES	57
3.2.1. ENTENDIMIENTO DEL PROBLEMA.....	57
3.2.2. REQUISITOS DEL SISTEMA.....	58
3.2.3. DEFINICION DE FUENTES DE DATOS, LOCALIZACION E INTEGRACION	59
3.3. SELECCIÓN DE VARIABLES	62
3.3.1. CLASIFICACIÓN DE VARIABLES	62
3.3.2. VARIABLES DE ESTRUCTURA	62

3.3.3. VARIABLES DE EMPLAZAMIENTO Y BARRIO	68
3.4. PARÁMETROS MÍNIMOS DE DEPURACIÓN Y NORMALIZACIÓN	70
3.4.1. INTRODUCCIÓN	70
3.4.2. TRATAMIENTO DE VARIABLES	71
3.4.3. AJUSTES POR ÍNDICES	72
3.4.4. OTROS AJUSTES	73
3.5. CASO PRÁCTICO DE SELECCIÓN DE VARIABLES, DEPURACIÓN Y NORMALIZACIÓN	75
3.5.1. SELECCIÓN DEL ÁREA DE ESTUDIO	75
3.5.2. SELECCIÓN DE DATOS DE PARTIDA: VARIABLES	78
3.5.3. SELECCIÓN DE MUESTRAS	79
3.5.4. AJUSTE DE PRECIOS SEGÚN LA EVOLUCIÓN DE PRECIOS DE MERCADO	83
3.5.5. FILTRADO PRELIMINAR	86
3.5.6. AJUSTE DE PRECIOS POR NEGOCIACIÓN Y ELIMINACIÓN DE ANEXOS VINCULADOS	87
3.5.7. NORMALIZACIÓN DE DATOS	89
3.5.8. ELIMINACIÓN DE OUTLIERS	91
3.6. ZONIFICACIÓN	93
3.6.1. INTRODUCCIÓN	93
3.6.2. MÉTODOS DE ZONIFICACIÓN	94
3.7. CASO PRÁCTICO DE ZONIFICACIÓN	98
3.7.1. INTRODUCCIÓN	98
3.7.2. SELECCIÓN DE DATOS DE PARTIDA: VARIABLES	99
3.7.3. SELECCIÓN DE MUESTRAS	100
3.7.4. NORMALIZACIÓN DE DATOS	126
3.7.5. ANÁLISIS DE PRINCIPALES VARIABLES	127
3.7.6. ANÁLISIS DE VARIABLES DE DISTANCIA	129
3.7.7. RESULTADO ZONIFICACIÓN	134
3.8. CONSTRUCCIÓN DE MODELOS	138
3.8.1. REGRESIÓN LINEAL	139
3.8.2. INTELIGENCIA ARTIFICIAL	141
3.8.3. CLASIFICADORES K-VECINOS MÁS CERCANOS	144
3.8.4. MODELOS HÍBRIDOS	145
3.8.5. PARTICIÓN DE BASES DE DATOS	147
3.8.6. SEGUIMIENTO Y CALIBRACIÓN	147
3.9. EVALUACIÓN DE MODELOS	149
3.9.1. INTRODUCCIÓN	149
3.9.2. ANÁLISIS DE RATIOS	150
3.9.3. MEDIDAS DE CENTRALIDAD	151
3.9.4. INTERVALOS DE CONFIANZA	152
3.9.5. MEDIDAS DE DISPERSIÓN	152
3.9.6. DIFERENCIAL DE PRECIOS	153
3.10. CASO PRÁCTICO DE EVALUACIÓN DE MODELOS	154

PARTE IV: CONCLUSIONES 156

4.1. CONCLUSIONES GENERALES157

4.1.1. APLICABILIDAD	157
4.1.2. IDONEIDAD DE METODOLOGÍAS	159
4.1.3. SEGMENTACIÓN	160
4.1.4. SELECCIÓN DE VARIABLES	161
4.1.5. PARÁMETROS MÍNIMOS DE DEPURACIÓN DE DATOS	162
4.1.6. CONSTRUCCIÓN Y EVALUACIÓN DE MODELOS	163
4.1.7. PROPUESTA DE REGULACIÓN	165

4.2. PROPUESTA DE NOTAS DE GUÍA DE BUENAS PRÁCTICAS.....166

TÍTULO I. DISPOSICIONES GENERALES	166
Capítulo I. Ámbito, principios y definiciones	166
Capítulo II. Comprobaciones y documentación.....	168
Capítulo III. Advertencias.....	169
TÍTULO II. CÁLCULO DEL VALOR	170
Capítulo I. Requisitos de partida.....	170
Capítulo II. Variables	171
Capítulo III. Segmentación	171
Capítulo IV. Parámetros mínimos de depuración y normalización	172
Capítulo V. Construcción de modelos	173
TÍTULO III. EVALUACIÓN Y PRESENTACIÓN DE RESULTADOS	175
Capítulo I. Rangos de confianza en la valoración	175
Capítulo II. Criterios para la elaboración de informes	176

4.3. FUTURAS LÍNEAS DE INVESTIGACIÓN177

4.4. BIBLIOGRAFÍA178

PARTE V: ANEXOS 187

ÍNDICE DE IMÁGENES

Imagen 2.1. Mapa de valores de la Comunidad de Madrid	34
Imagen 2.2. Cronología modelos de valoración automatizada	36
Imagen 2.3. Ejemplo modelo monocéntrico	46
Imagen 2.4. Ejemplo modelo policéntrico.....	47
Imagen 3.1. Proceso de creación de modelo AVM	54
Imagen 3.2. Fases en la creación de un modelo AVM.....	55
Imagen 3.3. Estrategia de datos.....	59
Imagen 3.4. Chamartín en la zona metropolitana de Madrid	75
Imagen 3.5. Barrios en el distrito de Chamartín	76
Imagen 3.6. Muestras seleccionadas	81
Imagen 3.8. Evolución de precios medios e incrementos anuales.....	85
Imagen 3.9. Mapa de valores del suelo de Madrid.....	96
Imagen 3.10. Muestras seleccionadas	101
Imagen 3.11. Seccionado censal distrito de Chamartín	107
Imagen 3.12. Nivel de ruido en el distrito	112
Imagen 3.13. Distancia a colegio de referencia	118
Imagen 3.14. Distancia a hospital	119
Imagen 3.15. Distancia a zona comercial.....	120
Imagen 3.16. Distancia a zona prime de oficinas.....	121
Imagen 3.17. Distancia a M-30.....	122
Imagen 3.18. Distancia a zonas verdes	123
Imagen 3.19. Distancia a estaciones de metro	124
Imagen 3.20. Gradiente de distancia a la M-30 respecto al precio	133
Imagen 3.21. Análisis de emplazamientos	135
Imagen 3.22. Clusters propuestos para el distrito de Chamartín	137
Imagen 3.23. Regresión lineal.....	139
Imagen 3.24. Red neuronal artificial.....	141
Imagen 3.25. Clasificador K-vecinos	144
Imagen 3.26. Modelo híbrido loosely-coupled.....	146
Imagen 3.27. Modelo híbrido tightly-coupled.....	146
Imagen 3.28. Partición de datos.....	147

ÍNDICE DE TABLAS

Tabla 3.1. Base de datos con información inicial	82
Tabla 3.2. Evolución de precios en el distrito de Chamartín	83
Tabla 3.3. Base de datos con precios actualizados a Q3 2015	84
Tabla 3.4. Filtrado por tipología de vivienda.....	86
Tabla 3.5. Negociación de precios de oferta	88
Tabla 3.6. Base de datos con variables normalizadas	90
Tabla 3.7. Base de datos depurada	92
Tabla 3.8. Renta familiar disponible per cápita	103
Tabla 3.9. Nacionalidad de los residentes en el distrito de Chamartín	103
Tabla 3.10. Población por distrito y secciones censales	104
Tabla 3.11. Base de datos con información suplementaria.....	113
Tabla 3.12. Colegios.....	115
Tabla 3.13. Hospitales.....	115
Tabla 3.14. Zonas comerciales	116
Tabla 3.15. Zonas prime de oficinas	116
Tabla 3.16. M-30	116
Tabla 3.17. Zonas verdes	117
Tabla 3.18. Estaciones de metro	117
Tabla 3.19. Base de datos con variables de distancia	125
Tabla 3.20. Coeficientes de regresión de las principales variables.....	128
Tabla 3.21. Rangos coeficiente de dispersión.....	153
Tabla 3.22. Ejemplo evaluación de modelos.....	154

RESUMEN

Las valoraciones automatizadas (AVM según sus siglas en inglés) son modelos matemáticos de valoración que permiten estimar el valor de un inmueble o una cartera de inmuebles mediante la utilización de información de mercado previamente recogida y analizada. Las principales ventajas de los modelos AVM respecto a las valoraciones tradicionales son su objetividad, rapidez y economía.

Existe regulación al respecto y estándares profesionales en otros países, no obstante en España los criterios de valoración mediante modelos AVM siguen siendo excesivamente básicos, al no establecer procesos concretos ni distinguir entre los diversos métodos estadísticos existentes y la aplicación adecuada de los mismos. Por otra parte, desde la publicación de la Circular 3/2008 del Banco de España se ha extendido el uso de este tipo de valoraciones en España para la actualización de valor de inmuebles en garantía de préstamos hipotecarios.

La actual desregularización en nuestro país en cuanto a normativa de valoraciones automatizadas nos permite plantear propuestas y metodologías con el fin de aportar un punto de vista nuevo desde la investigación, la experiencia en otros países y la práctica empresarial más reciente.

Este trabajo de investigación trata de sentar las bases de una futura regulación en España en materia de valoraciones masivas adaptadas al marco hipotecario español.

ABSTRACT

The Automated Valuation Models (AVM) are mathematically based valuation models that allow to estimate the value of a property or portfolio of properties by using market information previously collected and analyzed. The main advantages of the AVMs in comparison to traditional valuations are the objectivity, speed and economy.

In other countries there is regulation and standards regarding the use of AVMs. However, in Spain the norms that apply to AVM valuations are still too basic, since these norms do not define specific processes and do not distinguish between the different statistical approaches available and their correct use. On the other hand, following the ratification of the 3/2008 Bank of Spain Circular the use of AVM models in Spain for the value update of mortgaged properties has increased.

The current deregulation in Spain regarding Automated Valuation Models allows to provide propositions and methodologies in order to offer a new point of view based on the research, experience in other countries and the most recent professional practice.

The present research thesis aims to provide basic principles and criteria for a future Spanish regulation on massive valuation in the context of the Spanish mortgage framework.

PARTE I: INTRODUCCIÓN

1. 1.1. INTRODUCCIÓN

La actual desregularización en nuestro país en cuanto a normativa de valoraciones automatizadas nos permite plantear propuestas y metodologías innovadoras con el fin de aportar un punto de vista nuevo desde la investigación, la experiencia en otros países y la práctica empresarial más reciente.

Este trabajo de investigación trata de aportar metodologías y procesos que podrán mejorar la actual regulación en materia de valoraciones masivas y controlar de una forma más eficiente las valoraciones realizadas mediante dichos métodos.

1.1. 2. JUSTIFICACIÓN

Mediante la correcta organización y explotación de la información, es posible apoyarse en métodos estadísticos para valorar aquellos inmuebles que cumplan los siguientes requisitos:

- I. Susceptibles de producción repetida
- II. Ubicados en zonas donde exista un mercado representativo de inmuebles comparables
- III. Inmuebles de los que se disponga de información suficiente que permita efectuar apropiadamente este tipo de valoraciones.

El Banco de España ha venido regulando de forma elemental la utilización de dichos métodos de valoración para entidades de crédito.

En la circular 3/2008 se menciona por primera vez la posibilidad de utilizar métodos estadísticos para actualizar el valor de dichos bienes hipotecados. La circular 3/2010 subraya la necesidad de que las valoraciones por métodos estadísticos sean emitidas por una sociedad de tasación homologada por el Banco de España.

En diciembre del año 2012, El Banco de España anunció un proyecto de circular para la modificación de la circular 3/2008, entre otras. En dicho borrador se establecían ciertas recomendaciones a las sociedades de tasación respecto a las valoraciones por métodos estadísticos, que profundizaban mucho más en cuanto a metodología, limitaciones y aplicabilidad del método.

No obstante, la regulación sigue siendo excesivamente básica y permisiva, al no establecer procesos concretos ni distinguir entre los diversos métodos estadísticos existentes y la aplicación adecuada de los mismos.

1.1.3. OBJETIVOS

Los objetivos de esta investigación son los siguientes:

- Adaptar las metodologías de valoración automatizada a las necesidades del Banco de España para actualización de valor de inmuebles hipotecados.
- Redactar un borrador o aproximación normativa a lo que podría considerarse una regulación adecuada en la materia, con base en la normativa existente en otros países y a la regulación internacional.
- Establecer un procedimiento que pueda ser utilizado por el propio Banco de España como control de calidad interno para evitar sobrevaloración de activos inmobiliarios.

1.1.4. HIPÓTESIS

Es posible desarrollar un marco teórico de regulación de las valoraciones estadísticas para finalidades hipotecarias realizadas en España que permita:

- a. mejorar la calidad y fiabilidad de las valoraciones estadísticas,
- b. promover la armonización con estándares internacionales,
- c. delimitar justificadamente el campo de aplicación y
- d. normalizar la aplicación de metodologías estadísticas

La regulación es además necesaria para evitar el uso de metodologías de insuficiente fiabilidad y sería conveniente formularla a través de un estándar o guía de buenas prácticas.

1.1.5. PROPUESTA METODOLÓGICA

La metodología a seguir durante el desarrollo de la tesis parte del análisis de los estándares de valoración AVM existentes en otros países desde el punto de vista de su aplicabilidad al mercado hipotecario español.

Por tanto, primeramente se analizan los antecedentes incluyendo la revisión del marco regulatorio y estándares para valoraciones AVM que han sido publicados en otros países, así como el estado del arte y literatura científica sobre valoraciones automatizadas o ciertos aspectos de las mismas.

Los capítulos referentes a metodología se desarrollan siguiendo el mismo esquema que se seguiría para desarrollar un modelo de valoración, complementando los apartados teóricos con casos prácticos enfocados en la selección de variables, depuración de bases de datos, zonificación y evaluación de rangos de confianza.

Tras las conclusiones generales, se concluye la tesis con una propuesta de regulación que responde a los objetivos e hipótesis propuestos.

Se describen a continuación los aspectos a tratar en cada uno de los capítulos de la tesis:

Antecedentes:

1. Marco regulatorio en Europa y en España

Análisis de la legislación en otros países y en España. Estudio y análisis de la legislación vigente en materia de valoraciones estadísticas en otros países donde el marco regulatorio para estas valoraciones está más regulado.

2. Estado del arte

Resumen de las principales referencias bibliográficas consultadas y su aplicabilidad en la presente investigación.

3. Conclusiones parciales.

Aplicabilidad del método y limitaciones según el análisis de la literatura científica y del marco regulatorio actual en nuestro país.

Selección de las finalidades, casos y tipologías para los cuales sería apropiada la utilización de métodos estadísticos de valoración.

Metodología:

1. Introducción

2. Planteamiento, requisitos y fuentes

Definición de los requisitos de partida, pautas y normas básicas para la construcción de modelos y su posterior evaluación. Análisis de las fuentes de datos disponibles.

3. Selección de variables

Análisis de las variables que se consideran obligatorias para la realización de la valoración. También se proponen variables que es recomendable incluir pero que no sería práctica su obligatoriedad ya que en ciertos casos pueden ser variables no informadas.

4. Parámetros mínimos de depuración y normalización

Una de las fases clave de la valoración estadística es la depuración y normalización de datos. Mediante el estudio profuso de los procesos existentes de depuración y normalización de datos se seleccionarán aquellos parámetros mínimos que deberían llevarse a cabo con carácter obligatorio (tales como eliminación de *outliers*)

5. Caso Práctico de selección de variables, depuración y normalización

6. Zonificación

Análisis de metodologías de segmentación en base a la ubicación del inmueble.

7. Caso Práctico de zonificación

8. Construcción de modelos

Dentro del marco de las valoraciones estadísticas existen diferentes metodologías que permiten calcular el valor de un inmueble. No obstante, no todas las metodologías existentes se consideran válidas para las finalidades reguladas por el Banco de España. En este apartado se analizarán las diferentes metodologías existentes con el fin de delimitar el conjunto de metodologías adecuadas al presente marco regulatorio.

Además, las condiciones del mercado inmobiliario son cambiantes en el tiempo. Se establecerán procedimientos para el seguimiento, calibración y actualización del modelo en el horizonte temporal, que aseguren que el modelo se mantiene actualizado y ajustado al mercado actual durante todo su ciclo de vida.

9. Evaluación de modelos

Actualmente no existe un estándar global de medición de rangos de confianza en las valoraciones automatizadas. Por tanto se considera necesario que la evolución de modelos responde a un estándar global y unificado. Por este motivo, se trasladan en este capítulo las bases de evaluación de modelos establecidas por la I.A.A.O., consideradas un referente a nivel mundial.

Se incluye además un caso práctico de evaluación de modelos basado en el análisis de ratios.

Conclusiones

1. Conclusiones generales

En las conclusiones se exponen justificadamente los principales criterios que se considera necesario introducir en la propuesta de regulación y se sientan las bases para el borrador de propuesta de regulación.

2. Propuesta de notas de guía de buenas prácticas

Borrador de propuesta legislativa que regule la aplicación de metodologías de valoración estadística y recoja todos los parámetros, pautas y normas establecidas en los capítulos anteriores con el fin de mejorar y armonizar la aplicación de estas metodologías en el marco regulatorio del Banco de España.

Se trata de una propuesta de redacción de notas de guía de buenas prácticas adaptadas a las necesidades, características y particularidades del mercado hipotecario en España.

La presente propuesta aborda temas que a día de hoy están aún por resolver en nuestro marco legislativo, aportando de este modo conocimiento significativo en este campo de aplicación.

PARTE II: ANTECEDENTES

2.1. MARCO REGULATORIO EN EUROPA

2.1.1. RESUMEN

La actual ausencia de una normativa reguladora de las valoraciones masivas en nuestro país, requiere el desarrollo de propuestas que permitan ir progresivamente avanzando en la utilización de estas herramientas de valoración dentro del marco hipotecario.

Otros países europeos llevan años utilizando herramientas de valoración masiva para distintas finalidades relacionadas con el mercado hipotecario, tales como control de calidad de valoraciones, monitorización y ajuste de valores o incluso valoraciones iniciales para garantía hipotecaria de determinados tipos de inmuebles.

A nivel europeo, se utilizan actualmente modelos de valoraciones masivas en el marco hipotecario para determinar el valor de inmuebles en Dinamarca, Noruega, Alemania, Italia, Irlanda, Países Bajos, Suecia y Reino Unido, además de España. No obstante, los países que cuentan con una normativa más completa relacionada con este tipo de valoraciones son Dinamarca, Alemania y Reino Unido. A nivel mundial, los países con mayor experiencia en la aplicación de modelos de valoración masiva son Estados Unidos, Suecia y Canadá.

Cabe mencionar que estos métodos de valoración tienen gran aceptación para la determinación de la base imponible de impuestos que gravan los bienes inmuebles; la práctica ganada en este campo también puede ser de gran utilidad en el marco hipotecario.

La experiencia en otros países nos sirve como punto de partida para analizar cómo debería ser en España la normativa que regule este tipo de valoraciones, para qué

finalidades y usos es factible su utilización y cómo han evolucionado estas normativas desde sus inicios.

También es relevante analizar qué metodologías concretas se están utilizando actualmente en otros países. Además de los modelos hedónicos, que han sido usados tradicionalmente para la valoración masiva de inmuebles, hay otras metodologías que permiten la valoración de inmuebles por métodos estadísticos. Las redes neuronales, por ejemplo, presentan el problema de la llamada “caja negra”, es decir, que la interpretación de los resultados resulta especialmente compleja, por lo que resulta necesario analizar la aceptación de estas metodologías en el mercado.

Por último se analizan las fuentes de información disponibles en cada país, ya que la fiabilidad y calidad de las valoraciones masivas depende en gran medida de la calidad y grado de depuración de las bases de datos. En aquellos países en los que el mercado inmobiliario es más transparente, y los datos de los registros de la propiedad son públicos, es posible obtener valoraciones con un mayor grado de fiabilidad.

Como conclusión se resumen aquellos aspectos de aplicación generalizada y las propuestas que serían de fácil aplicación en nuestro país.

2.1.2. INTRODUCCIÓN

Dada la actual situación económica y la incertidumbre que reina en el mercado, es especialmente importante poder actualizar los valores de los activos inmobiliarios de forma frecuente y reducir el coste imputable a las valoraciones en las operaciones hipotecarias.

Existe una demanda creciente de valoraciones realizadas con modelos AVM ya que son mucho más económicas que las valoraciones tradicionales y también se pueden realizar en plazos de tiempo mucho más reducidos. Estas ventajas permiten un seguimiento continuado del valor de los activos inmobiliarios y, en el caso de

inmuebles que sirven como cobertura de préstamos hipotecarios, permiten conocer el Loan to Value – la relación entre valor y préstamo - durante toda la vida del préstamo.

Estos modelos de valoración automatizada también permiten realizar un control de calidad de las valoraciones realizadas al inicio del préstamo hipotecario.

Según el borrador de la nota informativa del Royal Institution of Chartered Surveyors¹ relativa a los modelos de valoración automatizada, la definición de AVM es la siguiente: *“Automated Valuation Models use one or more mathematical techniques to provide an estimate of value of a specified property at a specified date, accompanied by a measure of confidence in the accuracy of the result, without human intervention post-initiation”*

La valoración automatizada, de acuerdo con esta definición, es un sistema que permite estimar el valor de un inmueble mediante el uso de modelos matemáticos en una fecha concreta sin intervención humana tras iniciar el proceso de valoración. La valoración debe ir acompañada del nivel de confianza en el valor proporcionado que proporcione una indicación de la fiabilidad del resultado.

Una de las ventajas de los AVM es que las valoraciones son completamente objetivas e imparciales al no intervenir el valorador humano en la selección de comparables ni en la homogeneización de los mismos.

¹ El Royal Institution of Chartered Surveyors es una asociación profesional fundada en Londres en 1792 que acredita a profesionales del sector inmobiliario en los campos de la construcción, la propiedad inmobiliaria y el suelo. El Libro Rojo de valoración que publica RICS (RICS Appraisal and Valuation Standards) contiene reglas obligatorias, consejos acerca de la buena práctica y comentarios para todos los miembros de la RICS que realicen valoraciones de activos.

Antes de la implantación de los modelos AVM, el uso de los índices de precios era bastante generalizado en el sector, especialmente en Reino Unido y Suecia. Mediante la utilización de estos índices de precios se actualizaban los valores de tasaciones previas permitiendo realizar el seguimiento del activo e identificar activos que necesitan ser valorados de nuevo. Las principales desventajas de los índices de precios son la falta de precisión, ya que no se toman en cuenta los inmuebles concretos a valorar, y que es requisito indispensable disponer de una valoración anterior.

Dentro de los AVM empleados en la actualidad se aplican principalmente los modelos hedónicos, los sistemas basados en inteligencia artificial y los clasificadores de selección de comparables (Sales Comparison models), tales como el KNN (K Nearest Neighbour).

2.1.3. EMPLEO DE HERRAMIENTAS AVM EN OTROS PAÍSES

Actualmente los modelos AVM solamente están establecidos de forma sólida en tres países a nivel mundial: Suecia, Estados Unidos y Canadá.

A nivel europeo se utilizan modelos de valoraciones masivas en el marco hipotecario para determinar el valor de inmuebles en Dinamarca, Noruega, Alemania, Italia, Irlanda, Países Bajos, Suecia y Reino Unido, además de España. No obstante, los países que cuentan con una normativa más completa relacionada con este tipo de valoraciones son Dinamarca, Alemania y Reino Unido.

En Italia e Irlanda su utilización está en una fase aún inicial, mientras que en Suecia tienen una larga trayectoria en la utilización de estos modelos de valoración, favorecida en gran parte por el carácter público de sus bases de datos y registros inmobiliarios.

En Dinamarca, el FSA aprobó en 2005 el uso de modelos AVM de valoración para ciertas tipologías de inmuebles hipotecados. Se utilizan los modelos AVM para ciertas

valoraciones iniciales de préstamos hipotecarios, como control de calidad de valoraciones y para realizar un seguimiento del valor del inmueble durante la vigencia del préstamo.

En Noruega, la mayor parte de las entidades financieras del país utilizan un mismo proveedor de servicios de AVM que fue implantado en el año 2000 por los 4 mayores bancos del país. La base de datos del modelo se actualiza diariamente con la información recibida del registro de la propiedad y de la mayoría de agentes de la propiedad inmobiliaria. A día de hoy la mayoría de los bancos en Noruega utilizan estos servicios de valoración automatizada para valoraciones al inicio del préstamo y para realizar el seguimiento del Loan to Value durante la duración del préstamo.

En Alemania, si bien estos modelos están en fase de desarrollo, su uso es limitado ya que aún hay desconfianza entre ciertas entidades financieras. Se utiliza AVM para ciertas valoraciones iniciales de préstamos hipotecarios y como control de calidad de valoraciones.

En Holanda se utilizan los modelos AVM principalmente como control de calidad de valoraciones presenciales. Se espera que en los próximos años se extienda su uso a la valoración de carteras.

En Reino Unido se comenzaron a utilizar los modelos AVM aproximadamente en el año 2000. Se utilizan principalmente para actualizaciones de valor y valoraciones de carteras. Se están comenzando a usar también en préstamos iniciales.

La utilización de modelos AVM en Suecia tiene una larga trayectoria. La utilización de estos modelos de valoración automatizada en Suecia se basan en información recogida para cálculo de impuestos sobre bienes inmuebles, que es de carácter público. Los registros de la propiedad inmobiliaria recogen datos identificativos de los inmuebles, así como datos de transacciones, características de los inmuebles e información sobre la propiedad y derechos sobre la propiedad, siendo todos estos datos accesibles para cualquiera que los solicite.

Los bancos suecos han estado utilizando modelos AVM desde los años 90 y los aceptan ampliamente como una herramienta eficiente de valoración en los préstamos iniciales. Los modelos AVM se utilizan principalmente en aquellas valoraciones de inmuebles residenciales cuyos valores de mercado son bajos o cuyo ratio entre préstamo y valor del inmueble es bajo.

Las valoraciones tradicionales se limitan, por tanto, en Suecia a aquellos inmuebles cuyo valor de mercado es elevado. a operaciones con un Loan to Value elevado, o a aquellos casos en los que la inspección del inmueble resulta ineludible. Para inmuebles no residenciales los modelos AVM no se consideran adecuados.

La utilización de AVM para valoración de portfolios o control de calidad de préstamos iniciales es la más extendida en todos los mercados. En mercados evolucionados, se extiende además a valoraciones para préstamos iniciales y revaloraciones.

El acceso a información de transacciones en el mercado inmobiliario es muy importante en el desarrollo de los AVM. En países en los que las bases de datos de información inmobiliaria son públicas, es más fácil implementar modelos de valoración automatizada. En países en los que son privadas es más complicado implementar estos modelos de valoración, si bien algunos bancos pueden disponer de información suficientemente relevante de transacciones como para desarrollar un modelo propio, como ocurre por ejemplo en Suecia y Alemania. En Reino Unido actualmente las entidades bancarias se plantean facilitar información de sus bases de datos de transacciones inmobiliarias a sus proveedores de forma que con información más transparente y real la utilización y desarrollo de los AVM pueda avanzar más rápidamente.

2.1.4. EL MARCO REGULATORIO DE LOS MODELOS AVM

Los países en los que existen notas de guía o estándares profesionales para la utilización de modelos AVM son Dinamarca, Alemania y Reino Unido.

En los mercados más evolucionados se ha aplicado la auto-regulación desde las asociaciones de valoradores y desde las asociaciones hipotecarias. Las restricciones inicialmente impuestas por la legislación, agencias de rating y reguladores financieros han ido evolucionando hasta quedar suplidos por estándares profesionales y notas de guía de buenas prácticas.

En España no es posible la utilización de modelos AVM en valoraciones iniciales con finalidad hipotecaria, ya que la orden ECO 805/2003 establece requisitos mínimos incompatibles con la utilización de estos métodos, como por ejemplo la inspección del inmueble y comprobación de superficies.

Por otra parte, en España las bases e datos de catastro y registro aún no están vinculadas y los datos de superficies consignados en dichas bases de datos frecuentemente es incorrecto, por lo cual la comprobación de superficies es imprescindible.

En otros países en los que los datos de superficies están correctamente consignados en el registro de la propiedad y en las bases de datos para el cálculo de impuestos sobre bienes inmuebles, la inspección y comprobación de superficies para valores iniciales de préstamos hipotecarios solo es imprescindible en ciertos casos en los que, debido a la singularidad del inmueble o a que su valor es elevado, es necesario realizar una valoración más detallada.

No obstante de aplicarse ciertos cambios normativos, en un futuro cabría la posibilidad de utilizar métodos híbridos (Appraiser Assisted AVM) con intervención de valoradores expertos para ofrecer un servicio intermedio entre una valoración tradicional y una valoración por métodos estadísticos, de forma que el valorador puede completar la valoración AVM con información adicional o determinar si en casos puntuales es recomendable realizar una valoración completa debido a las características específicas del inmueble.

Esto supondría una importante reducción de costes para las entidades financieras y una simplificación de los procesos en la mayoría de los casos. Permitiría que los inmuebles más estandarizados o con valores de mercado bajos se puedan valorar de forma semi-automatizada, con una intervención mínima de los valoradores expertos. La participación del valorador experto también permite identificar de forma eficaz aquellos inmuebles que requerirían una valoración más detallada.

En mercados con un mayor desarrollo de los modelos AVM la tendencia actual es facilitar el uso de estos modelos de valoración mediante la publicación de estándares y notas de guía de buenas prácticas en las que se detalla en qué tipo de inmuebles se recomienda la utilización de AVM y el grado de supervisión necesaria por parte del valorador.

Recientemente los mayores proveedores de servicios AVM en Europa se han asociado, creando la European AVM Alliance. A esta organización pertenecen por el momento tres empresas procedentes de Holanda, Noruega y Reino Unido respectivamente. Uno de sus objetivos es crear estándares uniformes a nivel europeo que regulen la utilización de los modelos AVM, así como los requisitos de calidad en aspectos tales como la precisión, las metodologías empleadas o la presentación de los resultados obtenidos de la valoración.

El acuerdo de Basilea 2 se ha implementado en Europa a través de la Capital Requirements Directive [2], la cual establece como requisito que las valoraciones sean revisadas por un valorador independiente. En Reino Unido la interpretación de Basilea 2 se ha realizado asumiendo que se permite el uso de AVM siempre que un valorador independiente controle su funcionamiento y precisión, y confirme que los valores obtenidos mediante la aplicación de modelos AVM son iguales o inferiores a los de mercado.

No obstante, en España el proyecto de Circular del Banco de España de modificación de la Circular 3/2008 deja claro que en el caso de inmuebles cuyo valor exceda los 3.000.000 de euros, la valoración del inmueble deberá ser revisada de acuerdo a lo exigido en la orden ECO 805/2003. Esto supone una novedad con respecto a la

normativa actual (Circular 3/2008), en la que se indica únicamente que dichas valoraciones deberán ser revisadas por un tasador independiente.

2.1.5. MARCO REGULATORIO EN ESPAÑA

De acuerdo con lo recogido en la Orden ECO 805/2003, es importante que los responsables de realizar las valoraciones dispongan de numerosos datos sobre transacciones y ofertas ajustadas que permitan identificar parámetros adecuados para la homogeneización de activos inmobiliarios comparables, así como disponer además de datos históricos a nivel nacional, que permitan estimar la evolución de precios de venta en cada mercado local.

Mediante la correcta organización y explotación de dicha información, es posible apoyarse en métodos estadísticos para valorar aquellos inmuebles que cumplan los siguientes requisitos:

- Susceptibles de producción repetida
- Ubicados en zonas donde exista un mercado representativo de inmuebles comparables
- Inmuebles de los que se disponga de información suficiente que permita efectuar apropiadamente este tipo de valoraciones.

En ningún caso se deberán valorar por métodos estadísticos los inmuebles singulares, inmuebles no susceptibles de producción repetida o inmuebles que requieran un análisis particularizado.

Tal y como hemos visto anteriormente en el marco regulatorio en Europa, los modelos AVM se utilizan en los mercado más avanzados para las siguientes finalidades en relación con el negocio hipotecario:

- Valoración de préstamos iniciales
- Análisis inicial de riesgo de crédito por parte del banco
- Refinanciaciones y segundas hipotecas sobre un inmueble
- Valoración de portfolios

- Actualizaciones de valor para control de calidad y supervisión del Loan to Value

En Suecia, país europeo con mayor trayectoria en el uso de AVMs, estos modelos de valoración se utilizan en los siguientes casos:

- Valoración de préstamos iniciales en inmuebles residenciales (excepto LTVs altos)
- Valoración de portfolios
- Actualizaciones de valor para control de calidad y supervisión del Loan to Value

En España la actual regulación de las valoraciones para finalidad hipotecaria no permite que las valoraciones iniciales para préstamos se realicen mediante modelos AVM, al ser indispensable inspeccionar el inmueble entre otras comprobaciones mínimas. Por tanto, en el marco regulatorio hipotecario, sugerimos su utilización en las actualizaciones de valor de inmuebles hipotecados para el control de calidad de los préstamos, así como para la valoración de portfolios. En todos los casos se recomienda su uso exclusivamente en inmuebles residenciales.

La directiva europea 2006/48/CE en su parte 2. Apartado 1.4.b indica lo siguiente:

“b) Supervisión de los valores de los bienes raíces

El valor del bien será supervisado con frecuencia y como mínimo una vez al año, cuando se trate de bienes raíces comerciales, y una vez cada tres años en el caso de bienes raíces residenciales. Se efectuará una supervisión más frecuente cuando las condiciones del mercado puedan experimentar variaciones considerables. Podrán utilizarse métodos estadísticos a fin de supervisar el valor del bien y determinar los bienes que precisen una nueva valoración. La valoración del bien será revisada por un tasador independiente cuando la información disponible indique que su valor puede haber disminuido de forma significativa con respecto a los precios generales del mercado. Para los préstamos que superen 3.000.000 EUR o el 5 % de los fondos propios

de la entidad de crédito, la valoración del bien será revisada por un tasador independiente al menos una vez cada tres años.

Por «tasador independiente» se entiende una persona que posea las cualificaciones, capacidades y experiencia necesarias para efectuar una tasación y que sea independiente del procedimiento de decisión crediticia.”

La Circular 3/2008 del Banco de España, que es la transposición a la regulación española de la Directiva europea 2006/48/CE, indica lo siguiente en su artículo 4.2 referente a los Requisitos específicos de los derechos reales sobre inmuebles:

“b) Valoración de los inmuebles.

Al inicio de la operación, la valoración del inmueble radicado en España deberá ser realizada por una entidad de tasación homologada, y de acuerdo con los requisitos establecidos en la Orden ECO/805/2003, de 27 de marzo, sobre normas de valoración de bienes inmuebles y de determinados derechos para ciertas finalidades financieras. En el caso de los inmuebles radicados en otros países, la valoración corresponderá a las entidades o personas reconocidas en los respectivos países para llevar a cabo dichas valoraciones con fines hipotecarios.

Posteriormente, el valor del bien hipotecado será revisado con una frecuencia mínima de un año, cuando se trate de inmuebles comerciales, y de tres años, en el caso de inmuebles residenciales. Esta revisión podrá realizarse mediante métodos estadísticos, salvo en el caso de inmuebles singulares o no susceptibles de producción repetida.

Adicionalmente, siempre que las condiciones del mercado u otras circunstancias indiquen que los precios pueden estar experimentando variaciones significativas con respecto a los precios generales del mercado que pudieran afectar a la reducción del riesgo de crédito conseguida mediante las garantías de derechos reales sobre inmuebles, se deberán efectuar valoraciones adicionales del conjunto de inmuebles vinculados a los préstamos afectados, por un tasador independiente que podrá utilizar procedimientos muestrales, salvo en el caso de inmuebles

singulares o no susceptibles de producción repetida; el tasador deberá cumplir también los requisitos del primer párrafo de esta letra b).

Para los préstamos que superen los 3 millones de euros o el 5% de los recursos propios de la entidad de crédito, la valoración del bien hipotecado será revisada, al menos una vez cada tres años, por un tasador independiente; el tasador deberá cumplir también los requisitos del primer párrafo de esta letra b).

En las revisiones a que se refiere el párrafo precedente, no será necesaria la inspección ocular del inmueble cuando el valor del préstamo, a causa de su amortización, sea inferior al 50% del valor de la garantía en el caso de inmuebles residenciales, o al 40% en el caso de inmuebles comerciales.”

Por tanto, la Circular 3/2008 del Banco de España permite la aplicación de modelos AVM para la actualización de valor de inmuebles residenciales y locales comerciales. No obstante habrá que tener especial cuidado en la valoración mediante modelos masivos de locales comerciales, ya que la propia Circular establece que sólo serán aplicables estos modelos a inmuebles susceptibles de producción repetida.

En el caso de los locales comerciales, el importante peso en el valor que tienen características como la ubicación específica en determinado eje comercial, los ratios entre fachada/fondo o la distribución de superficies en varias plantas hace muy complicada su valoración mediante modelos AVM, especialmente en inmuebles ubicados en ejes comerciales, que pueden llegar a alcanzar valores muy altos sujetos a la ubicación específicas y características del local.

Por este motivo este trabajo no considera recomendable la aplicación de métodos estadísticos en locales comerciales, excepto aquellos que se ubiquen en mercados muy secundarios, en los que el impacto de las características específicas en el valor no sea especialmente relevante. En algunos casos en los que se disponga de una valoración anterior reciente puede ser incluso más recomendable la actualización de un valor anterior mediante índices, ya que de este modo sí que se estarían teniendo en cuenta las características intrínsecas del inmueble.

Entendemos que en garajes, al poder considerarlos como anexos residenciales, también serían de aplicación los modelos AVM.

Cabe destacar que la Circular no hace mención a circunstancias especiales que puedan hacer que un inmueble residencial no deba valorarse por comparación, tales como la existencia de algún tipo de protección que establezca un precio máximo de venta o la existencia de contratos de arrendamiento que afecten al valor del inmueble.

Para la valoración de inmuebles con contrato de arrendamiento, en algunos países ya existen modelos AVM basados en la capitalización directa. No obstante, estos modelos presentan la dificultad añadida de tener que calcular rentabilidades iniciales y la necesidad de ajustes adicionales en el caso de inmuebles arrendados por encima o por debajo de mercado. En el caso de España los modelos AVM aún no están excesivamente desarrollados, por lo que siempre que la entidad solicitante pueda facilitar un resumen con las condiciones del contrato, la opción más viable sería realizar la valoración mediante la aplicación de AVMs por comparación complementando dicha valoración con la intervención de valoradores expertos que analicen el valor desde el punto de vista de la capitalización de sus ingresos netos.

Circular 3/2010 BdE NORMA PRIMERA apartado 3.i.

“Para las operaciones inferiores a 500.000 euros con garantía de primera carga sobre vivienda terminada, o activos adquiridos en pago de deudas que procedan de esta clase de operaciones, la entidad podrá utilizar como mejor estimación de su valor actual el menor importe entre el 80% de la última tasación disponible y el que resulte de la actualización de esa tasación, con antigüedad inferior a un año, obtenida mediante métodos estadísticos por una sociedad de tasación que reúna las condiciones señaladas en el apartado h).”

El 12-12-2012 el Banco de España publicó un proyecto de Circular² para modificar la circular 3/2008 entre otras donde se recogían además recomendaciones respecto a las valoraciones por métodos estadísticos.

El 15-10-2013 se publica un nuevo proyecto de Circular³, similar al anterior pero sin modificar la circular 3/2008 y con requisitos en cuanto a valoraciones AVM únicamente para la entidad de crédito que encarga la valoración.

Finalmente se aprueba la Circular 3/2014 de 30 de julio, basada en las dos anteriores, pero en la que no se hace mención alguna a las valoraciones por métodos estadísticos.

La valoraciones AVM siguen por tanto sin una regulación específica por parte del Banco de España, salvo lo especificado en la Circular 3/2008.

2.1.6. USO ACTUAL DE MODELOS AVM EN ESPAÑA

Actualmente las sociedades de tasación utilizan modelos AVM principalmente en los siguientes casos:

- Actualización de valor de inmuebles hipotecados, según lo indicado en la Circular 3/2008 del Banco de España.
- Valoración de portfolios de activos adjudicados
- Valoración a efectos contables de ciertos inmuebles para la Sareb, según lo establecido en la Circular 5/2015 del Banco de España.
- Control de calidad
- Valoración online de viviendas individuales

^{2, 3} Se incluye en anexos los Proyectos de Circular de 2012 y 2013 al no haberse publicado en el B.O.E.

Las entidades financieras también utilizan modelos AVM, principalmente como herramienta de control de calidad de los inmuebles valorados.

Por su parte las administraciones utilizan modelos AVM con fines fiscales para el cálculo de valores de bienes inmuebles. La Gerencia del Catastro de la Comunidad de Madrid utiliza una red neuronal para la valoración de inmuebles de toda la región. Utiliza para ello valores de inmuebles declarados por notarios y registradores.

Imagen 2.1. Mapa de valores de la Comunidad de Madrid

Fuente: Julio Gallego⁴

⁴ Presentación realizada por Julio Gallego en el II Congreso Nacional Catastro que tuvo lugar en 2010 en San Luis, Argentina.

Las entidades de valoración están implantando en los últimos años modelos más avanzados basados en comparables, en detrimento de los modelos basados en índices, especialmente a raíz de la publicación de la Circular 3/2008 del Banco de España. La coyuntura en los primeros años de crisis y la regulación de las actualizaciones de valor de inmuebles hipotecados son los factores que han intensificado su uso en España por parte de las sociedades de tasación.

El borrador de Circular del Banco de España publicado en 2012, aunque no se llegó a aprobar con la redacción inicial, ya indicaba las limitaciones de los métodos basados en índices y pretendía restringir su uso a los casos en los que la misma Sociedad de Tasación había realizado la valoración inicial.

2.2. ESTADO DEL ARTE

2.2.1. INTRODUCCIÓN

Muchos aspectos metodológicos de las valoraciones masivas han sido ampliamente estudiados a través de diferentes enfoques y la literatura es, por tanto, profusa y diversa. Sin embargo, los aspectos relacionados con la legislación aún deben ser considerados con mayor profundidad, especialmente en países como el nuestro en los que la regulación al respecto es mínima.

Tal y como muestra la literatura, el tema de las valoraciones masivas ha sido tradicionalmente abordado a través de diferentes métodos de análisis hedónico. Las referencias en la literatura a modelos de regresión es exhaustiva en comparación con las referencias a modelos de inteligencia artificial y clasificadores basados en distancias, ya que estos modelos son más recientes, mientras que el uso de la regresión lineal ha sido más extenso.

Imagen 2.2. Cronología modelos de valoración automatizada

VALORACIÓN AUTOMATIZADA

Desde 1990 Regresión lineal como ayuda al sistema tradicional

Desde 2003 Regresión exponencial e investigación en Redes Neuronales y otros algoritmos (k-vecinos, k-means, etc.)

2004 primeras valoraciones masivas con Redes (uso interno)

2004 – 2008 múltiples valoraciones masivas con Redes

2008 primeras pruebas con Lógica Difusa

2009 valoraciones masivas y mapas de valor con Lógica Difusa

⁵ Presentación realizada por Julio Gallego en el II Congreso Nacional Catastro que tuvo lugar en 2010 en San Luis, Argentina.

Es un hecho corroborado que la literatura tiende a ser extensa en el estudio de variables, construcción de modelos y metodologías. Sin embargo, los aspectos legislativos suelen quedar fuera de las investigaciones. Este aspecto menos estudiado es por este motivo de interés para la actual investigación, permitiendo además avanzar en cuestiones normativas que podrían tener una aplicación directa en nuestro país.

La evaluación de mejoras en el medio construido y el estudio de los factores que ponen en valor emplazamientos específicos aportan datos cuantificables acerca del valor económico de cada atributo o variable que contribuye en el precio. Las cualidades relacionadas con emplazamiento y barrio contribuyen de forma decisiva a determinar las especificaciones de la vivienda, y su importancia se puede a su vez traducir a términos económicos.

Además, con el fin de obtener un modelo válido, es esencial seleccionar las variables relevantes y construir una base de datos que afronte correctamente la heterogeneidad del problema. Por lo tanto, es importante conocer los diferentes tipos de atributos y diferentes enfoques que se pueden seguir en la fase de modelado de datos. Estas cuestiones también son tratadas en las próximas secciones.

2.2.2. ESTÁNDARES AVM

Los principales estándares de valoración automatizada a nivel mundial son los publicados por el International Association of Assessing Officers⁶ en 2003 (“Standard on Automated Valuation Models”).

⁶ La International Association of Assessing Officers fue fundada en 1934 y es una asociación profesional reconocida a nivel mundial que promueve la educación, investigación e innovación en el campo de la valoración inmobiliaria vinculada a los impuestos sobre bienes inmuebles. Fue una de las 9 organizaciones que en los años 80 promovieron la publicación de los “Uniform Standards of Professional Appraisal Practice”, los estándares de valoración que se aplican en Estados Unidos y Canadá.

El I.A.A.O. ha publicado además otros estándares en relación a las valoraciones AVM, como son el “Standard on Ratio Studies” sobre criterios a seguir en la medición de la fiabilidad de los modelos AVM, o el “Standard on Mass Appraisal of Real Property” que se centra en las valoraciones masivas para finalidades fiscales.

El Royal Institution of Chartered Surveyors publicó en 2013 una nota informativa sobre valoración automatizada titulada Automated Valuation Models (AVMs). Su aplicación no es por tanto obligatoria para los miembros de RICS cuando realicen valoraciones AVM, sino que la finalidad del documento es ofrecer información a sus miembros sobre temas relevantes o de actualidad.

En España, la Asociación Española de Análisis del Valor⁷ publicó en 2013 sus estándares sobre métodos estadísticos de Valoración Masiva de Inmuebles. La finalidad es establecer los criterios básicos que deben cumplir sus miembros al realizar valoraciones mediante métodos estadísticos y son, por tanto, de obligado cumplimiento para sus miembros. No obstante, se trata de unos estándares de requisitos mínimos que no entran en detalle en aspectos metodológicos.

2.2.3. ATRIBUTOS - VARIABLES

Las propiedades residenciales situadas en el ámbito metropolitano tienen tres peculiaridades: la complejidad de sus atributos, el vínculo directo con su emplazamiento y la durabilidad en el tiempo (esta última característica hace posible por ejemplo el fuerte mercado de segunda mano). Estas peculiaridades son la base para comprender qué variables podrían ser determinantes para la estimación del precio de la vivienda.

⁷ La Asociación Española de Análisis de Valor (AEV) es una asociación profesional fundada a finales de 2012 por tres de las principales Sociedades de Tasación españolas, SOCIEDAD DE TASACIÓN, TINSA Y SIVASA, y en la actualidad integra a un amplio grupo de esas mismas entidades y consultoras inmobiliarias. La AEV está adherida a la Asociación Hipotecaria Española.

Los atributos urbanos de cada vivienda pueden ser o bien inherentes al emplazamiento o bien externalidades positivas. Small y Steimetz (2006) clasifican a su vez las externalidades en económicas y tecnológicas. Las externalidades económicas son aquellas que simplemente representan un beneficio económico (efecto indirecto) mientras que las externalidades tecnológicas afectan también al bienestar de los residentes vecinos “per se” y, como efecto colateral, los residentes están dispuestos a pagar más (efecto directo)

La investigación de Small y Steimetz (2006) verifica que sólo en el caso de externalidades tecnológicas es posible maximizar la precisión del método de auto regresión espacial. Este método, sin embargo, no es adecuado para tratar el impacto de externalidades positivas de tipo económico. En la siguiente sección se ofrece un mayor acercamiento a los diferentes métodos que se aplican para analizar estas externalidades.

Además de las categorías nombradas, los atributos de la vivienda suelen comúnmente clasificarse en tres tipos básicos: Estructural (referido a la construcción de la propia vivienda), de barrio y de emplazamiento. Esta clasificación representa de forma clara y precisa los diferentes tipos de características que pueden ser incluidos en un análisis de regresión con el fin de obtener precios hedónicos descompuestos. Chin y Chau (2003) identificaron las siguientes variables clave para las categorías de barrio y emplazamiento:

Atributos de barrio:

- Nivel económico de los residentes
- Distancia a un colegio de referencia
- Distancia al hospital
- Distancia a lugares de culto
- Tasa de criminalidad
- Nivel de ruido debido al tráfico

- Distancia a zona comercial
- Distancia al bosque
- Calidad medioambiental

Atributos de emplazamiento:

- Distancia al distrito financiero
- Vistas a lago, río o mar
- Vistas a valles, montañas, etc.
- Obstrucción de vistas
- Duración del contrato de arrendamiento

Otros estudios recientes han usado métodos hedónicos para estimar los beneficios derivados de los siguientes factores:

- Calidad del aire. Kim et al. (2003) y Anselin y Lozano-Gracia (2008)
- Ruido aéreo. Cohen y Coughlin (2008)
- Ruido de vehículos o trenes. Andersson et al. (2009)

Waddell et al. (1993) incluyeron los siguientes atributos en su lista de variables:

- Distancia al aeropuerto
- Distancia a la autopista
- Distancia a núcleos comerciales
- Distancia a centros educativos
- Distancia a hospitales
- Usos del suelo colindantes
- Superficie media de las unidades residenciales en el mismo barrio

- Nacionalidad de los residentes en el barrio

Las variables de emplazamiento incluidas en el estudio de Anselin y Lozano-Gracia (2008) en sus estudios son las siguientes:

- Tasa de criminalidad
- Calidad de los centros educativos
- Distancia a zonas verdes
- Distancia a la autopista
- Duración del trayecto al centro de trabajo
- Porcentaje de población en el umbral de la pobreza
- Ingresos por núcleo familiar
- Calidad del aire

Wen et al. (2005) también consideraron como características de barrio los centros educativos (guarderías, colegios e institutos) y los establecimientos y/o equipamientos tales como supermercados, oficinas postales, hospitales o bancos situados a una distancia inferior a 1.000 metros.

Richardson et al. (1974) consideran que otros atributos favorables son la ausencia de industria, el acceso a zonas verdes, la presencia de escuelas de referencia y la presencia de zonas de alto nivel socio-económico, mientras que entre los factores desfavorables se incluyen el ruido, la polución, barrios degradados y proximidad a polígonos de industria pesada.

Es de esperar que algunas variables, como por ejemplo la distancia a la autopista, tengan un efecto positivo en el precio; sin embargo es importante tener en cuenta que en distancias excesivamente cortas suele tener el efecto contrario debido al ruido, la polución, etc. Varios estudios demuestran que las

distancias cortas a espacios abiertos tales como parques y zonas verdes, suelen hacer que los precios de las viviendas aumenten.

Los estudios demuestran que las ejecuciones hipotecarias en el barrio pueden tener un efecto negativo en los precios de las viviendas próximas. En este sentido, el barrio se define como un bien público que puede ser devaluado o revalorizado a causa de actuaciones de los residentes respecto al espacio común. La devaluación atribuible a una ejecución hipotecaria se traduciría así como un descenso en la calidad de vida del barrio.

También cabe destacar que aquellos atributos relacionados con el año de construcción del edificio y los relacionados con la distancia al centro de la ciudad o casco histórico a menudo quedan distorsionados dada la forma en la que las ciudades crecen (desde el centro hacia fuera) Esto puede llegar a provocar una distorsión en los coeficientes debido al alto grado de correlación entre las variables.

El estudio de Wadell et al. (1993) revela que a mayor extensión total del barrio, menor es el precio de las viviendas en él situadas, debido a la idea generalizada de que los barrios más pequeños tienden a ser más exclusivos.

Contrariamente, la superficie media de las viviendas tiene un efecto positivo en los precios de viviendas próximas, dada la disposición de los compradores a pagar más por vivir en un barrio supuestamente más pudiente.

La investigación de Bourassa et al. (2007) concluye que las vistas a río, mar o lago pueden incrementar el precio de la vivienda en un 10% y que la calidad del paisaje circundante puede también tener efectos notables en el precio final.

Wen et al. (2005) afirman que para una vivienda estándar, el grado de contribución al precio de la vivienda de los atributos de estructura es del 60%, la

contribución de los atributos de barrio es del 16,5% y la contribución de los atributos de emplazamiento es del 19,8%.

2.2.4. MÉTODOS

El enfoque hedónico establece una relación entre los precios implícitos y los atributos que ofrece dicha vivienda. Los múltiples atributos que forman parte de la unidad residencial pueden ser medidos y capitalizados para obtener el valor total del inmueble.

Los modelos hedónicos pueden ser utilizados para hacer estimaciones acerca del porcentaje sobre el precio final atribuible a cada variable. La teoría hedónica se debe a dos principios básicos:

- La demanda de un producto determinado está directamente relacionada con sus características específicas. Sin embargo, el análisis de características individualizadas con métodos más tradicionales es demasiado compleja. El precio agregado se puede calcular de forma simple haciendo el sumatorio de precios hedónicos relativos a variables finitas.
- Basándose en las características de los productos que ofrece el Mercado, debe darse un equilibrio entre la oferta y la demanda para que puede existir un mercado ideal. Los proveedores de cada bien deben identificar si hay o no demanda para determinadas características.

El primer modelo hedónico fue desarrollado por A. T. Court en 1939. Este primer modelo fue utilizado para ilustrar el precio de un vehículo en función del precio de sus componentes. Uno de los primeros académicos en extrapolar el uso de modelos hedónicos a la investigación inmobiliaria fue Ridker en 1967.

La vivienda es un bien puramente heterogéneo y por tanto un método de análisis centrado en la medida y clasificación de las múltiples variables es a menudo adecuado para el análisis de precios. Sin embargo, es importante apreciar que

los modelos hedónicos requieren abundante información y un modelado específico para funcionar correctamente.

Los modelos hedónicos pueden incorporar tanto el efecto del precio medio de la vivienda en la zona (efecto indirecto) como los atributos y características de la vivienda y su emplazamiento (efecto directo). Los valores en renta de las viviendas, por ejemplo, quedan afectados por las características de viviendas próximas y entre estas características, especialmente por el precio. La influencia de estas externalidades puede mostrarse en un modelo econométrico de retardo espacial (space-lag model) Con este método, los atributos son ponderados en relación a un multiplicador espacial.

La diferencia principal entre el modelo de retardo espacial y el modelo de error espacial es que el segundo no tiene en cuenta efectos indirectos. Como consecuencia, la idoneidad de un método u otro depende en gran medida de la información que se pretenda extraer del análisis. En el caso de la actual investigación, el modelo de error espacial podría ser útil al omitir variables que no son parte fundamental del estudio y simplificar así el modelo ofreciendo una visión más clara de las variables a analizar.

Por otra parte, la diferencia entre modelos geoestadísticos y modelos de regresión basados en métodos hedónicos es la forma en la que cada uno de ellos interpreta la heterogeneidad espacial. El modelo hedónico muestra sólo variaciones de primer orden, mientras que el modelo geoestadístico también tiene en cuenta efectos de segundo orden.

La investigación de Bourassa et al. (2007) concluye que el análisis con mínimos cuadrados (Ordinary Least Squares), el cual es mucho más fácil de implementar que los métodos geoestadísticos, proporciona resultados con igual precisión en los supuestos en los que se incluyen variables dicotómicas. Los modelos de mínimos cuadrados tienen un rendimiento adecuado para valoraciones masivas y pueden por tanto ser una alternativa propicia a los complejos modelos geoestadísticos.

Páez (2009) menciona que la aplicación de enfoques econométricos es más habitual que la de enfoques geoestadísticos, siendo posible que la predisposición a los métodos econométricos esté inducida por el conocimiento más amplio de estos métodos entre los profesionales del sector inmobiliario.

Según Hoover (2005), la mayoría de los estudios econométricos tienen un enfoque positivista, ya que su fuente de conocimiento científico se basa en deducciones lógicas o en la observación empírica.

Partiendo de una epistemología positivista, los datos a utilizar para el estudio de relaciones econométricas deberán ser por norma general de tipo cuantitativo. Este acercamiento para la toma de datos comenzó a desarrollarse cuando la econometría se estableció como campo independiente en 1933, ya que el principal objetivo de la Sociedad Econométrica fue la “unificación de los enfoques teórico-cuantitativo y empírico-cuantitativo para resolver problemas económicos”. (Frisch, 1933).

A pesar de que el razonamiento deductivo asociado a la filosofía positivista y a los datos de tipo cuantitativo a priori parte de la teoría, Hoover (2005) demostró que un enfoque que parte de los datos para desarrollar una teoría también es posible en investigaciones del campo de la econometría.

2.2.5. MODELO

Páez (2009) subraya la importancia que tiene producir un modelo adecuado para conseguir resultados precisos.

Bourassa et al. (2007) opinan que los conceptos de dependencia espacial y definición de áreas locales están íntimamente relacionados. El concepto de área local se basa en la posibilidad de reemplazamiento. Elementos reemplazables son aquellos para los cuales el incremento de precio en uno conlleva un incremento de demanda en el otro. Cuando existe un equilibrio, el precio es igual

para ambos. Dentro de un área local, el precio de la vivienda se estabiliza porque en dicha zona existen varios inmuebles reemplazables.

Para conseguir resultados precisos es importante controlar las definiciones de mercados locales dentro del modelo. Los mercados locales normalmente se agrupan en áreas geográficas que pueden ser identificadas por especialistas en el mercado y valoradores. Algunos métodos, como por ejemplo el geoestadístico, crean franjas de distancia que varían de inmueble a inmueble de tal forma que el efecto depende de la distancia a dicho inmueble. Estos modelos consideran el efecto de proximidad al inmueble más importante que el efecto sobre un barrio predefinido.

De igual modo, los modelos monocéntricos se basan en la dependencia entre distancia al distrito financiero y precio de los inmuebles residenciales.

Imagen 2.3. Ejemplo modelo monocéntrico

En las zonas metropolitanas de ciudades contemporáneas sin embargo, los centros tienden a tener múltiples nodos y los precios que se apoyan en este factor se pueden perfilar como gradientes con puntos focales en cada uno de los centros de negocios, ocio, etc. La influencia en el precio tiende a estar limitada a un área pequeña, ya que el efecto desaparece normalmente en distancias superiores a 20 kilómetros.

Imagen 2.4. Ejemplo modelo policéntrico

Waddell et al. (1993) afirman que el efecto espacial de la proximidad a carreteras, zonas comerciales, universidades y colegios varía en forma y distancia. Muchos puntos de interés para los que el acceso es considerado beneficioso exhiben un efecto negativo si las distancias son excesivamente cortas debido a externalidades tales como ruido, polución o tráfico.

Finalmente, es esencial reconocer la extrema sensibilidad de muchos métodos a valores atípicos o extremos. Los valores atípicos corresponden a propiedades cuyas especificaciones o atributos varían sustancialmente con respecto al valor medio para dicha variable. Estos valores atípicos pueden confundir al sistema y disminuir la precisión del modelo; por este motivo es importante identificarlos y eliminarlos de la base de datos. Otros factores que pueden alterar de forma dramática el funcionamiento del modelo son la no-linealidad del modelo, la multicolinealidad de variables o la heteroscedasticidad en las variaciones.

2.3. CONCLUSIONES PARCIALES

2.3.1. CONCLUSIONES MARCO REGULATORIO

Como conclusión principal cabe destacar que la regulación en el marco de las valoraciones hipotecarias es escasa. En los mercados más evolucionados se ha aplicado la auto-regulación desde las asociaciones de valoradores y desde las asociaciones hipotecarias.

La totalidad de los estándares profesionales que han sido consultados establecen la necesidad de incluir en el informe una indicación del rango de confianza de la valoración.

Los estándares profesionales también recomiendan por lo general que el uso de modelos AVM se limite a inmuebles residenciales ubicados en entornos urbanos, es decir, aquellas zonas donde pueda existir un mercado homogéneo y suficientemente representativo de inmuebles comparables.

Los modelos AVM son especialmente adecuados en el caso de actualizaciones de valor de inmuebles, es decir, aquellos casos en los que ya se ha realizado una primera valoración presencial, así como en los casos en los que el Loan to Value es bajo y por lo tanto el riesgo derivado de la utilización de sistemas menos exactos es inferior.

Las principales áreas de negocio futuro en el campo de las valoraciones masivas serían aquellas que integrasen servicios de valoración con servicios para el control del riesgo en el marco hipotecario. Este tipo de servicios integrados ya se están llevando a la práctica en Estados Unidos.

El principal aspecto a mejorar en el corto plazo sería la estandarización a nivel internacional de la forma de medir los rangos de confianza de las valoraciones, de forma que los usuarios finales pudiesen comparar de forma fiable los resultados de diferentes proveedores y de forma que se estableciesen criterios uniformes de rangos mínimos de confianza para que la valoración sea fiable.

2.3.2. CONCLUSIONES ESTADO DEL ARTE

El estudio del estado del arte ha sido productivo como punto inicial para esta investigación por diversos motivos: Primero, ha proporcionado una estructura básica para el estudio. Además, los artículos revisados han sido una excelente fuente de información acerca de atributos relevantes en el precio de la vivienda.

Uno de los primeros rasgos teóricos que ha marcado la estructura de este trabajo de investigación es la categorización de características en estructura, barrio y emplazamiento. Los artículos estudiados muestran de forma unánime y regular esta misma estructura en sus estudios y por lo tanto la utilización de estas tres categorías en el campo de la valoración masiva de inmuebles parece coherente. Esta clasificación, en la que dos de las tres categorías tienen una relación directa con los atributos urbanos, refuerza la hipótesis de que dichos atributos constituyen un factor fundamental en el precio de la vivienda.

Tal y como Páez (2009) puntualiza, actualmente hay un amplio consenso en la interpretación de los efectos económicos de variables espaciales en el mercado inmobiliario. Los atributos relacionados con el emplazamiento han demostrado tener similitudes entre barrios. Por el contrario, las heterogeneidades existentes podrían explicar las diferencias en la demanda actual.

Es un hecho corroborado que la literatura tiende a ser extensa en el estudio de las características de estructura y, en menor medida, de barrio. Sin embargo, la relevancia de las características relativas al emplazamiento específico tiende a ser relegada a un segundo plano.

El efecto de las características de barrio y emplazamiento desde el punto de vista del gradiente de precios aún no ha sido adecuadamente estudiado en la mayor parte de la literatura existente.

Finalmente, las investigaciones relativas al precio ponderado de los atributos se pueden desarrollar según dos tipos de modelos estructurales: o bien modelos basados en la oferta y la demanda o bien modelos basados en los atributos individuales que componen la entidad a valorar. En éste último enfoque se base el método hedónico. Por otra parte, el primer enfoque citado no ha sido abordado en tanta profundidad en el mercado de la vivienda como en otros sectores, dada la heterogeneidad y singularidad de las diferentes unidades de vivienda; por tanto, el estudio de la demanda para determinadas características de la vivienda podría también proporcionar información práctica para el mercado de vivienda si bien su estudio sería de gran complejidad.

2.3.3. ESQUEMA CONCEPTUAL

En los mercados evolucionados las restricciones inicialmente impuestas por la legislación, agencias de rating y reguladores financieros han ido evolucionando hasta quedar suplidos por estándares profesionales y notas de guía de buenas prácticas. Éstos tienen la ventaja de poder ser mucho más específicos sin que ello suponga restringir otras prácticas o criterios en determinados casos. Por este motivo se considera adecuada la publicación de unas notas de guía de buenas prácticas en España que permitan flexibilidad de uso, pero promoviendo las mejores prácticas en el uso de modelos AVM.

En cuanto a la aplicabilidad de los modelos, si bien ciertos estándares profesionales en mercados avanzados permiten la valoración de diversos usos mediante modelos AVM, en el caso de España la falta de transparencia en el sector, la dificultad de acceso a fuentes de información fiables y el importante peso de la vivienda colectiva en el volumen total de operaciones hipotecarias hace que el uso de estas metodologías sea indicado en la valoración de vivienda colectiva y garajes (al poder ser los garajes anexos vinculados a viviendas).

La aplicación de estas metodologías en locales comerciales y vivienda unifamiliar es más compleja. En el caso de las viviendas unifamiliares, éstas tienen a ser más singulares por lo que en un proceso automatizado se complica el proceso de identificación y ajuste de comparables. En la valoración de locales comerciales, los atributos de emplazamiento son clave en la determinación del valor del inmueble y es muy difícil trasladar estos atributos de emplazamiento a un modelo automatizado.

Se recomienda por tanto que la aplicación de estos modelos de valoración en nuestro país se enfoque por el momento en vivienda plurifamiliar y garajes. También se debería limitar el uso de modelos AVM a inmuebles ubicados en entornos urbanos y a actualizaciones de valor o valoración de inmuebles con Loan to Values bajos, si bien por el momento en España no es posible la utilización de modelos AVM en valoraciones hipotecarias iniciales al ser incompatible con lo dispuesto en la Orden ECO/805/2003.

Por otra parte, el acceso a la información de transacciones en el mercado inmobiliario es muy importante en el desarrollo de los AVM. Es en aquellos países con acceso libre a bases de datos públicas en los que más ha evolucionado el uso de los modelos automatizados de valoración. Además, ciertos métodos requieren de volúmenes importantes de comparables para llevar a cabo las valoraciones, tales como los clasificadores K-vecinos. Por ello es importante que desde las Administraciones se promueva el uso de estos modelos de valoración facilitando el acceso a bases de datos públicas de transacciones inmobiliarias.

Por último, el imprescindible incluir en el informe de valoración una indicación del rango de confianza de la valoración. Actualmente no hay consenso a nivel mundial sobre la forma de medir la fiabilidad de los modelos. No obstante los estándares AVM más extendidos son los publicados por la I.A.A.O., por lo que se recomienda que para facilitar la estandarización se apliquen las recomendaciones del I.A.A.O. en relación a los rangos de confianza de las valoraciones automatizadas.

PARTE III: METODOLOGÍA

3.1. INTRODUCCIÓN

La metodología para realizar la modelización parte del entendimiento del problema y sus requisitos de partida para sentar las bases de construcción del modelo. Posteriormente se deben establecer cuáles van a ser las fuentes de obtención de datos y cómo va a ser el proceso de análisis de dichos datos.

Se establecen las variables a tener en cuenta así como el nivel geográfico, que dependerá del proceso de zonificación llevado a cabo. El análisis de los datos conlleva procesos de normalización, ajuste y depuración que permitirán aumentar la fiabilidad de los datos de partida y mejorar el funcionamiento del modelo.

Se procede entonces a la construcción del modelo, escogiendo cuáles serán los algoritmos más idóneos en función de los datos de partida y las características del mercado inmobiliario a analizar.

Tras la modelización se deberán seguir un proceso de seguimiento y ajustes que permita mejorar el rango de confianza de los resultados. Al tratarse de un proceso continuo, en el que cada cierto tiempo se debe incluir nueva información actualizada de comparables en la base de datos, deben establecerse procesos de retroalimentación continuos para la explotación de los datos.

Imagen 3.1. Proceso de creación de modelo AVM

Fuente: Indra

A continuación se muestra de forma resumida y teórica la secuencia de fases en la creación de un modelo automático de valoración:

Imagen 3.2. Fases en la creación de un modelo AVM

Fuente: Indra

Universo objetivo: Para cada uno de los modelos predictivos a realizar por tipo de inmueble, se determinan los filtros que deben cumplir para poder ser valorados. Adicionalmente se definen las variables o características de los inmuebles que influyen en la valoración y si es necesario realizar algún tipo de muestreo para garantizar el tiempo de proceso de los datos para modelar.

Análisis descriptivo: Es necesario realizar un análisis en profundidad de las variables que inciden directamente en el valor del inmueble, para garantizar la estabilidad histórica, y para que no existan errores de aprendizaje de los datos.

Manipulación: Tratamiento posterior de los datos, que incluye, entre otros, clasificación, normalización, depuración y transformación de los datos iniciales. Este proceso suele ser necesario para encontrar variables transformadas que proporcionan un mejor ajuste con respecto a la variable objetivo (valoración).

Modelización : En la fase de modelización, primero hay que particionar los datos en un grupo de entrenamiento, validación y test, de forma que sean conjuntos independientes. Los dos primeros forman parte en la modelización y sirven para validar la fórmula a través de parámetros estadísticos mientras que el último es un conjunto independiente que sólo se utiliza para valorar el error del modelo. Después se probarán distintas técnicas estadísticas de modelización con las variables elegidas como influyentes en la valoración.

Evaluación: Se evalúan y comparan los diferentes modelos estadísticos, estableciendo cuál de ellos resuelve de manera más eficaz el problema planteado, con el menor error.

3.2. PLANTEAMIENTO, REQUISITOS Y FUENTES

3.2.1. ENTENDIMIENTO DEL PROBLEMA

El mercado inmobiliario español carece de un estándar de valoración de activos que permita hacer valoraciones masivas con un grado de precisión y confianza suficientes. Las bases para crear modelos de valoración automatizada deben ser un conocimiento profundo del sector inmobiliario y una metodología analítica adecuada.

Se trata de crear un proceso automático que permita obtener el valor económico para inmuebles residenciales plurifamiliares, con base en las variables de las cuales depende directamente el valor de dichos inmuebles.

Mediante la correcta organización y explotación de dicha información, es posible apoyarse en métodos estadísticos para valorar aquellos inmuebles que cumplan los siguientes requisitos:

- Susceptibles de producción repetida
- Ubicados en zonas donde exista un mercado representativo de inmuebles comparables
- Inmuebles de los que se disponga de información suficiente que permita efectuar apropiadamente este tipo de valoraciones.

Siguiendo las directrices marcadas por el Banco de España en la Circular 3/2008, en ningún caso se deberán valorar por métodos estadísticos inmuebles singulares, inmuebles no susceptibles de producción repetida o inmuebles que requieran un análisis particularizado.

3.2.2. REQUISITOS DEL SISTEMA

Uno de los requisitos de partida es conocer en detalle los requisitos asociados a las valoraciones a realizar, tales como la finalidad de las valoraciones a realizar y las hipótesis de partida sobre los inmuebles. El marco en el que se basa esta propuesta de estándar es para aquellas finalidades hipotecarias relacionadas con la revisión periódica de valor de los inmuebles sobre los que se ha concedido anteriormente una hipoteca, en línea con lo indicado en la Circular 3/2008 del Banco de España.

Otro de los requisitos es establecer el número de modelos analíticos a desarrollar, en función de las tipologías de inmuebles a valorar. En este caso y tomando en consideración tanto la experiencia en otros países europeos como las características del mercado inmobiliario en nuestro país se considera adecuado establecer un modelo para residencial plurifamiliar y otro para garajes, al ser estos anexos vinculados habitualmente a las viviendas plurifamiliares. La Circular 3/2008 del Banco de España permite además valorar otros tipos de inmuebles, ya sean locales comerciales o viviendas plurifamiliares. Sin embargo, se considera más apropiado comenzar únicamente con viviendas plurifamiliares y sus anexos vinculados, ya que estas tienden a un mayor grado de estandarización y el volumen de inmuebles comparables es mayor que en otros tipos de inmuebles.

Por último, es necesario establecer un mínimo de inmuebles comparables para que se considere que existe un mercado representativo y se pueda valorar un cierto inmueble mediante métodos estadísticos. Dado que el número de comparables que establece la orden ECO 805/2003 es de 6, es lógico que el número mínimo de inmuebles comparables para considerar que existe un mercado representativo sea de al menos 6.

Si en la zonificación creamos *clusters* o zonas geográficamente homogéneas en términos inmobiliarios, el número mínimo de inmuebles comparables que se establezca deberán encontrarse dentro del mismo *cluster* al del inmueble a valorar.

Por este motivo es importante ajustar la extensión de cada *cluster* en función de la densidad de cada zona geográfica y de las características de cada mercado local, de forma que los *clusters* ocupen extensiones más amplias de terreno en zonas rurales, donde no existen determinantes geográficos tan relevantes y donde además la densidad tanto de viviendas como de transacciones inmobiliarias tiende a ser mucho menor que en las principales ciudades. En núcleos urbanos densos por el contrario la concentración tanto de viviendas como de transacciones inmobiliarias aumenta exponencialmente y se pueden encontrar diferencias importantes de valor con base en la ubicación específica de cada inmueble; por lo tanto en esta zonas más densas se deben establecer *clusters* de tamaño más reducido de forma que abarquen zonas homogéneas dentro del tejido urbano.

También se deberán establecer rangos dentro de cada variable para determinar qué tipo de inmuebles se consideran singulares y, por tanto, no susceptibles de ser valorados por métodos estadísticos.

3.2.3. DEFINICION DE FUENTES DE DATOS, LOCALIZACION E INTEGRACION

Imagen 3.3. Estrategia de datos

Fuente: Elaboración propia

Es necesario definir la estrategia de datos para ser utilizados y establecer un modelo de datos, con la inclusión de los datos históricos necesarios para modelar y la definición de los *target* de partida.

El nivel geográfico más bajo disponible se establecerá como resultado de la fase de zonificación, en la cual quedarán definidos los *cluster* o zonas homogéneas en términos inmobiliarios.

Las fuentes de datos dependerán del acceso a datos de cada sociedad de tasación, pero en general podrán incluir transacciones reales, bases de datos propias de tasaciones anteriores realizadas por la sociedad así como testigos utilizados en tasaciones anteriores y testigos procedentes de búsqueda automatizada en portales inmobiliarios.

Es muy importante incluir a ser posible comparables de transacciones reales, ya que estos comparables nos servirán para establecer el rango de negociación en un determinado mercado local y un determinado momento del tiempo si analizamos la diferencia entre estas transacciones reales y otros comparables de oferta. Si los proveedores de servicios tuviesen acceso ilimitado a bases de datos de transacciones reales, tal y como ocurre en otros países, sólo se debería utilizar este tipo de información. No obstante, dado que actualmente el acceso a este tipo de información no es posible a gran escala, es necesario incluir información de oferta para obtener un volumen de comparables suficientemente representativo.

Las bases de datos propias y sus testigos ofrecen información que ha sido previamente filtrada por valoradores expertos, por lo que el grado de fiabilidad de estos comparables será por lo general muy alto.

Por último, los comparables de búsqueda automatizada pueden ofrecer un gran volumen de muestras para la base de datos, por lo que incluir este tipo de comparables es importante con el fin de construir una base de datos de inmuebles extensa y que cubra geográficamente todo el territorio de forma representativa.

Es importante considerar que probablemente cada sub-base de datos necesitará ciertos procedimientos de pre-filtrado y normalización específicos en función de su procedencia y nivel de confianza. Por ejemplo, en la base de datos de transacciones reales no será necesario el ajuste por negociación. En la base de datos de búsqueda automatizada sería adecuado realizar una eliminación de *outliers* previa para eliminar

aquellos comparables con información deficiente o incongruente, puesto que la fuente de datos es menos fiable.

El modelo de datos una vez realizados los ajustes correspondientes podrá ser único, con una única base de datos en la que se incluyan todas las muestras de inmuebles comparables. La estrategia de datos para la base de datos única contempla:

- Definición del modelo de datos único integrando toda la información
- Definir y garantizar la profundidad histórica de los datos para la realización de modelos (La IAAO no recomienda utilizar comparables recogidos hace más de un año. En general se recomienda que la profundidad histórica sea lo más corta posible para que el ajuste de precio a fecha de hoy sea mínimo)
- Definición de la frecuencia o periodicidad de las actualizaciones (Por ejemplo, trimestral)
- Procesos de carga y actualización periódica de la información

Por último, y tras haber realizado un análisis de las variables de las que depende el valor del inmueble así como las variables de las cuales se dispone de información en función de las fuentes de datos seleccionadas, se establecen las variables para viviendas plurifamiliares y para garajes que definen la fórmula matemática del modelo. Entre estas variables podrán incluirse variables transformadas que engloben varias características o variables de emplazamiento específico para cada *cluster* en función de si se trata de un inmueble en costa, en núcleo urbano, etc.

3.3. SELECCIÓN DE VARIABLES

3.3.1. CLASIFICACIÓN DE VARIABLES

El estudio del estado del arte ha sido productivo como punto inicial para esta investigación por diversos motivos: Primero, ha proporcionado una estructura básica para el estudio, proporcionando un esquema de categorización de atributos específico para este tipo de estudios. Además, los artículos revisados han sido una excelente fuente de información acerca de atributos que probablemente tengan un efecto en el precio de la vivienda.

Uno de los primeros rasgos teóricos que ha marcado la estructura de este trabajo de investigación es la categorización de características en estructura, barrio y emplazamiento. Los artículos estudiados muestran de forma unánime y regular esta misma estructura en sus estudios y por lo tanto la utilización de estas tres categorías en el campo de la valoración masiva de inmuebles parece coherente.

De cara a la presente investigación, se han tratado por una parte los atributos de emplazamiento y barrio como medio para establecer una zonificación adecuada, que permita delimitar el área donde buscar comparables para cada inmueble a valorar. Por otra parte y una vez identificada la zona homogénea en la que se ubica el inmueble a valorar, los atributos de estructura se utilizan para determinar el valor del inmueble con base en la comparación de las características del inmueble a valorar en relación a la de inmuebles similares ubicados dentro de la misma zona homogénea.

3.3.2. VARIABLES DE ESTRUCTURA

Según la orden ECO/805/2003 los Comparables se definen de la siguiente manera: Son inmuebles que se consideran similares al inmueble objeto de valoración o adecuados para aplicar la homogeneización, teniendo en cuenta su localización, uso, tipología, superficie, antigüedad, estado de conservación, u otra característica física relevante a dicho fin.

Según los estándares de modelos AVM de la IAAO (IAAO, 2003), las variables a incluir en modelos hedónicos pueden representar características del inmueble (tales como superficie útil o antigüedad del edificio), información de la ubicación, información demográfica o variables derivadas de características del inmueble (por ejemplo superficie multiplicada por un índice de calidad).

Según el Standard on Mass Appraisal de la IAAO (IAAO, 2013) las siguientes variables se incluyen entre las que son normalmente relevantes para determinar el valor de inmuebles residenciales (no se han incluido en el listado aquellas que corresponden únicamente a viviendas unifamiliares):

- Superficie útil
- Calidad constructiva
- Antigüedad del inmueble o estado de conservación
- Estilo constructivo
- Zonas secundarias, tales como sótanos, garajes, porches cubiertos o balcones
- Equipamiento del inmueble, como por ejemplo número de baños o existencia de aire acondicionado.
- Instalaciones del inmueble
- Mercado local en el que se ubica el inmueble
- Submercado o barrio en el que se ubica el inmueble
- Servicios o ventajas de la zona, tales como vistas, primera línea de mar o golf.
- Molestias que dependen de la ubicación, tales como tráfico, ruido del aeropuerto o proximidad a zonas comerciales.

Según los estándares de RICS para las valoraciones AVM (RICS, 2013) , entre los datos de partida para realizar la valoración se pueden incluir los siguientes:

- Estilo arquitectónico
- Tipología del inmueble
- Año de construcción o antigüedad
- Ubicación
- Planta en la que se ubica el inmueble

- Número de estancias
- Número de dormitorios
- Número de baños
- Superficie de parcela
- Porche
- Construcciones anexas
- Parking/garaje
- Calidades o estado de conservación del inmueble

Para desarrollar un modelo válido es esencial seleccionar las variables relevantes y crear una base de datos que muestre correctamente la heterogeneidad del problema. De ahí surge la importancia de entender los múltiples atributos que pueden ser relevantes y conocer los diferentes enfoques a considerar en la fase de modelado.

Los atributos de estructura acumulan el mayor porcentaje de contribución al valor de un inmueble, correspondiendo a este conjunto de atributos aproximadamente el 60% del valor del inmueble según se indica en el estado del arte (Wen et al., 2005).

Aparte de la superficie construida, a la cual puede llegar a corresponder más del 50% del precio final, la tipología del inmueble, la existencia de anexos, la calidad constructiva o la antigüedad otros atributos de estructura relevantes en el cálculo del precio de la vivienda.

La variable más relevante, que es la superficie construida, muestra habitualmente una relación lineal con el valor del inmueble. Esta tendencia lineal funciona de forma especialmente adecuada para las viviendas de menor superficie, dado que estas tienden a cierta estandarización originada por la limitación espacial, mientras que las viviendas unifamiliares de mayor superficie tienden a ser únicas y más exclusivas.

El número de atributos se deberá reducir a aquellos de los cuales depende directamente el valor del inmueble y se recomienda seguir las siguientes directrices en la selección de atributos o variables:

- La inclusión de atributos de los cuales no depende directamente el valor puede provocar problemas de aprendizaje y errores predictivos.
- La relación de atributos a incluir deberá estar compensado sin duplicar ni omitir características relevantes.
- Habrá que tener en cuenta el tratamiento posterior de los datos a incluir (clasificación, normalización, depuración y transformación de los datos). Los datos deberán poder transformarse en valores numéricos, dado el tipo de algoritmos a utilizar.
- Se propone utilizar dos tipos de variables:
 1. Atributos excluyentes: deberán ser idénticos en comparables e inmueble a valorar. Se considera como atributo excluyente la zonificación.
 2. Otros atributos: deberán ser similares a los del inmueble a valorar, pero no necesariamente idénticos.

A continuación se detallan las características consideradas relevantes para una valoración automatizada, teniendo en cuenta las fuentes de información disponibles actualmente para obtener información de inmuebles comparables de forma recurrente:

1. ATRIBUTOS EXCLUYENTES:

- Zonificación (atributos de barrio y emplazamiento)
- Uso: Vivienda plurifamiliar o Garaje

2.A. RESTO DE ATRIBUTOS VIVIENDA

Edificio:

- Año de edificación o última rehabilitación edificio
- Bloque abierto o manzana cerrada
- Zonas comunes (ascensor, piscina, zonas verdes, gimnasio, etc.). > Variable transformada

Vivienda:

- Tipología: ático, planta baja, dúplex, piso, etc.

- Planta de la vivienda
- Exterior o interior
- Número de dormitorios
- Antigüedad (última reforma elemento)
- Calidades constructivas

Superficies:

- Superficie vivienda
- Superficies exteriores privativas (terraza cubierta, terraza descubierta o jardín)
- Superficie de anexos: garajes y trasteros vinculados

2.B. RESTO DE ATRIBUTOS GARAJE

Edificio:

- Parking normal o robotizado

Plaza:

- Planta

Superficies:

- Superficie útil

Algunos de los atributos seleccionados deben ser tratados o transformados para poder incluirlos en la base de datos, ya que todos los atributos deben ser expresados de forma numérica. Los atributos que expresan categoría, por tanto, deben ser transformados para adecuarlos al formato numérico. Además, el número total de atributos debe estar compensado, sin duplicar características similares y sin omitir características determinantes para el precio final.

Se considera que las fuentes de datos comúnmente utilizadas en el mercado español para la construcción de bases de datos de inmuebles comparables son las siguientes:

- Tasaciones de viviendas realizadas por la propia compañía
- Comparables de mercado de transacciones reales
- Comparables de mercado de oferta
- Comparables de mercado obtenidos mediante aplicaciones de búsqueda automatizada inteligente

Debido a que es necesario recabar toda la información de variables informadas de todos los comparables que se incluyan en la base de datos, la falta de transparencia en el mercado español puede ser un obstáculo a la hora de obtener cierta información de inmuebles comparables o puede comprometer el volumen de inmuebles comparables en la base de datos.

En cualquier caso, se considera que las variables de estructura mínimas serían las siguientes:

1. ATRIBUTOS EXCLUYENTES:

- Zonificación (atributos de barrio y emplazamiento)
- Uso (vivienda plurifamiliar o garaje)

2.A. RESTO DE ATRIBUTOS VIVIENDA

Edificio:

- Año de edificación

Vivienda:

- Tipología: ático, planta baja, dúplex, piso, etc.
- Calidades constructivas

Superficies:

- Superficie vivienda
- Existencia de anexos: garajes y trasteros vinculados

2.B. RESTO DE ATRIBUTOS GARAJE

Superficies:

- Superficie útil

En cualquier caso, es interesante utilizar en la fase de selección de variables herramientas de selección matemática que ayuden a detectar variables correlacionadas o variables no relevantes para el modelo.

3.3.3. VARIABLES DE EMPLAZAMIENTO Y BARRIO

Se considera adecuado agrupar las variables de barrio en una variable transformada que integre todas las características relevantes de una zona en relación en el precio de los inmuebles que se ubican en la misma. Se tratan estas variables en el apartado de zonificación.

Se considera que las variables de emplazamiento son relevantes y tienen un impacto relevante en el precio de los inmuebles; no obstante, se considera que estas variables no deben ser obligatorias por la dificultad que supone incluirlas en un modelo de valoración automatizada. Además en la Orden ECO 805/2003 no se hace referencia a variables de emplazamiento específicas en su definición del término “comparable”, por lo que es un aspecto que debe quedar a la elección del valorador y que se deberá establecer caso por caso.

En el caso de incluir dichas variables en el modelo, será necesario identificar variables de emplazamiento específicas en función de la ubicación de los inmuebles. Es decir, deberán establecerse variables de emplazamiento diferentes según se trate de

inmuebles en costa, en núcleos urbanos, en entornos rurales, etc. De este modo la fórmula que relaciona las variables con el valor del inmueble será mucho más ajustada en cada caso y tomará en cuenta aspectos específicos de emplazamiento de cada submercado local.

Estas variables, dependiendo del tratamiento específico que se otorgue a cada una de ellas y a las peculiaridades de cada variable, se podrían incluir en la fase de zonificación junto con las variables de barrio, o bien incorporar directamente a la selección de variables dándoles un tratamiento similar a las variables de estructura.

3.4. PARÁMETROS MÍNIMOS DE DEPURACIÓN Y NORMALIZACIÓN

3.4.1. INTRODUCCIÓN

Una parte importante del experimento consiste en el preprocesado de los datos. El conjunto de datos disponible incluye valores atípicos que deben ser eliminados y otros valores que no nos interesan o que los programas de minería de datos no son capaces de procesar. Tras la recogida de datos es necesario realizar un análisis de los datos de partida con el fin de clasificar, depurar y normalizar toda la información.

Además de los datos incluidos inicialmente, es importante seguir modificando las bases de datos de forma regular con el fin de ir disminuyendo el porcentaje de error en cada algoritmo, añadiendo más comparables que complementen a los anteriores o eliminando aquellos comparables que pueden llevar a error.

Partimos de la base de que las sociedades de tasación podrán incorporar a la base de datos comparables que han pasado ya por un filtrado previo de calidad, como pueden ser sus propias tasaciones de inmuebles o comparables que hayan sido utilizados anteriormente en tasaciones. Este filtrado previo garantiza un mínimo nivel de calidad a los datos de partida, ya que estos datos han sido previamente revisados y depurados por valoradores.

En el caso de incluir en la base de datos comparables de fuentes menos fiables, será recomendable realizar un filtrado previo de las mismas eliminando aquellas muestras en las que la información sea incongruente, incompleta en sus variables principales o que muestre diferencias sustanciales con el resto de muestras para la misma zona.

3.4.2. TRATAMIENTO DE VARIABLES

El tratamiento de variables debería incluir los siguientes pasos de normalización y depuración:

- Autocompletado de los datos. El sistema deberá ser capaz de completar datos de testigos que falten, tales como número de dormitorios o antigüedad, con base en los datos medios de testigos similares.

Se recomienda que sólo se autocompleten variables secundarias, que afecten al valor del inmueble en menos medida. Variables tales como la superficie de la vivienda o su ubicación tienen tal relevancia en el valor del inmueble que si no disponemos de esta información del comparable, deberemos desecharlo.

Algunos algoritmos, como las redes neuronales, son capaces de realizar el autocompletado de datos automáticamente sin necesidad de depuración previa.

- Asignación de la variable de zonificación. Con base en la ubicación del inmueble y a la zonificación realizada (que se explica en detalle en el capítulo de zonificación) se deberá asignar a cada inmueble de la base de datos un *cluster* o variable de zonificación.
- Todas las variables se deben expresar de forma numérica. Por tanto es necesario transformar las variables no numéricas en una secuencia de números que sea procesable. En ciertas variables puede ser necesaria la discretización o agrupación de datos en escalones o rangos de valor.
- Normalización de variables. Para optimizar el rendimiento de la red, las cifras con las que esta trabaja han de estar comprendidas dentro de un intervalo cuya amplitud debe ser lo más reducida posible para evitar problemas de bucles o no aprendizaje. También se pueden utilizar tablas o fórmulas específicamente creadas para la normalización de variables.
- Creación de variables transformadas. Algunas variables pueden ser transformadas para englobar atributos interrelacionados y/o reducir el número de variables.

Se incluyen a continuación un ejemplo de cómo se podría realizar la normalización y creación de ciertas variables transformadas:

- Año de edificación o última rehabilitación (se puede computar el año más reciente de entre estas dos variables)
- Zonas comunes (ascensor, piscina, zonas verdes, gimnasio, etc.). Se pueden englobar en una única variable computando el número de anexos ponderados por importancia.
- Se pueden englobar la planta en la que se ubica la vivienda y si es exterior o interior en una única variable ponderando el número de planta en función de si es interior o exterior.
- Antigüedad y calidades de reforma: Se puede utilizar una tabla de ratios para englobar estas dos variables.
- Superficie: Se pueden crear tablas que ayuden a realizar el cómputo terrazas descubiertas y jardines. También se puede crear una variable transformada que tenga en cuenta tanto la superficie construida con comunes como la útil según el ratio entre ambas superficies.

3.4.3. AJUSTES POR ÍNDICES

Si bien los índices no se consideran apropiados para realizar valoraciones automatizadas, dado que existen otros métodos más fiables, las actualizaciones por índices pueden ser de gran ayuda en la fase de depuración de datos.

En primer lugar, es necesario establecer una profundidad histórica máxima para las bases de datos, de forma que con cada nueva actualización se eliminen los registros que superen dicha antigüedad máxima y se incorporen comparables recientes. Dentro de la base de datos existirán por tanto comparables introducidos a fecha actual así como otros con cierta antigüedad, por ejemplo comparables con 3 meses de antigüedad o 6 meses de antigüedad.

El necesario actualizar los precios de los comparables a fecha actual para que todos los datos sean comparables y se refieran a la fecha actual, lo cual se puede realizar mediante índices de evolución de precios en cada zona.

En el caso de incorporar comparables de oferta en las bases de datos, habrá que realizar un ajuste de precios por negociación. Este ajuste se puede realizar separando por una parte los comparables de transacciones reales de la base de datos de aquellos comparables que sean de oferta. Una vez separados los datos se puede calcular la diferencia de valores medios y por tanto estimar el porcentaje de negociación que habrá que restar a los comparables de oferta.

3.4.4. OTROS AJUSTES

Además de la normalización de variables y el ajuste de precios, es necesario realizar los siguientes ajustes en la fase de depuración de datos de partida:

- Establecer un rango de valores para cada atributo con el fin de evitar problemas de bucles o no aprendizaje. Estos rangos también nos ayudan a determinar qué consideramos como *outliers*. Aquellos inmuebles que se salgan del rango establecido se denominan *outliers* y por tanto el sistema los deberá eliminar automáticamente de la base de datos. Se pueden establecer los rangos de eliminación de *outliers* con base en la horquilla de percentiles de distribución.
- Se recomienda valorar los anexos de forma independiente, salvo que en el *cluster* en cuestión la gran mayoría de comparables incluyan anexos vinculados.

En viviendas con garajes vinculados se propone por tanto separar los garajes, estableciendo un modelo de datos para vivienda colectiva y otro modelo diferente para garajes. El modelo agrupado de viviendas con garaje sólo se recomienda en aquellos *clusters* específicos en los que el mercado en la zona esté dominado por viviendas con anexos vinculados.

Hay que tener en cuenta que si se decide hacer un modelo de viviendas con garaje para determinados *cluster*, no será posible valorar viviendas sin garaje en dichos *cluster*.

- Por último sería recomendable realizar una comprobación de superficies con catastro para detectar incongruencias significativas. En caso de diferencias significativas entre el dato de partida y el dato que figura en catastro habría que desechar el comparable.

3.5. CASO PRÁCTICO DE SELECCIÓN DE VARIABLES, DEPURACIÓN Y NORMALIZACIÓN

3.5.1. SELECCIÓN DEL ÁREA DE ESTUDIO

El primer paso es la selección de un área de estudio delimitada, en este caso el distrito de Chamartín.

Imagen 3.4. Chamartín en la zona metropolitana de Madrid

Fuente: Elaboración propia

Chamartín se compone de seis barrios:

- Castilla
- Hispanoamérica
- Ciudad Jardín
- Nueva España
- El Viso
- Prosperidad

Imagen 3.5. Barrios en el distrito de Chamartín

Fuente: Elaboración propia

El proceso de urbanización del distrito comenzó en el barrio de Prosperidad hacia 1860, cuando el barrio de Prosperidad no era más que un pequeño municipio a las afueras de Madrid. En 1925 las colonias de viviendas unifamiliares proliferaron en Ciudad Jardín y El Viso. Las colonias de El Viso, al contrario que las pertenecientes a Ciudad Jardín, fueron construidas especialmente para profesionales autónomos, como médicos, ingenieros y arquitectos.

El Plan General de Ordenación de 1941 recoge la inclusión de Chamartín dentro del área metropolitana de Madrid. El plan planteaba un cambio de uso de vivienda unifamiliar a vivienda colectiva con alta densidad de población en gran parte del distrito.

Es importante remarcar la heterogeneidad del distrito; sus seis barrios presentan importantes diferencias demográficas, sociales, culturales y económicas. Prosperidad, Ciudad Jardín y El Viso presentan orígenes muy diferentes entre sí y con barrios como Hispanoamérica y Nueva España, de más reciente creación. El barrio de Castilla es el más atípico, conviviendo en él estratos sociales muy variados.

El alto nivel socioeconómico, la diversidad de densidades de población y las grandes diferencias existentes entre los barrios hacen que Chamartín sea un área de estudio interesante. Además, la proximidad a zonas prime de oficinas, la inclusión de la franja este de la Castellana dentro del distrito, la proximidad a la M-30 y otros factores hacen posible la creación de gradientes de distancia a diferentes puntos de interés o servicios.

En el área de actuación, se llevan a cabo las fases de recogida de datos, análisis del mercado local, análisis de zonas homogéneas y análisis de puntos de interés y servicios públicos.

3.5.2. SELECCIÓN DE DATOS DE PARTIDA: VARIABLES

Para desarrollar un modelo válido, es esencial seleccionar las variables relevantes y crear una base de datos que muestre correctamente la heterogeneidad del problema.

Algunos de los atributos seleccionados deben ser tratados o transformados para poder incluirlos en la base de datos, ya que todos los atributos deben ser expresados de forma numérica. Los atributos que expresan categoría, por tanto, deben ser transformados para adecuarlos al formato numérico de la base de datos.

Además, el número total de atributos debe estar compensado, sin duplicar características similares y sin omitir características determinantes para el precio final. Este hecho revela la problemática de estudiar características secundarias en las categorías de emplazamiento y barrio, tales como distancia al distrito financiero, nivel de ruido y polución o densidad de población.

Los atributos se pueden clasificar en tres categorías: estructura, barrio y emplazamiento. Se tratarán posteriormente las variables relacionadas con Barrio y emplazamiento incluyéndolas dentro de la variable zonificación. Se indican a continuación las variables consideradas relacionadas con la estructura en este caso práctico:

Atributos de estructura:

- Calidad constructiva
- Superficie construida
- Antigüedad del edificio
- Número de dormitorios
- Planta en la que se sitúa el inmueble
- Número de anexos (garaje, piscina, terraza, zona verde, etc.)

- Número de plantas del edificio

3.5.3. SELECCIÓN DE MUESTRAS

Una vez analizados los atributos a incluir, se recoge la información de las muestras y se crea la base de datos. En la base de datos, cada muestra queda definida por información precisa relativa a cada variable o atributo seleccionado.

El método de selección de muestras basado en inmuebles residenciales ofertados obedece a las limitaciones de tiempo y acceso a información más precisa. Las restricciones de acceso a información del catastro relativa a transacciones inmobiliarias reales ha sido determinante a la hora de escoger este método de selección de muestras alternativo. Los resultados obtenidos a partir de transacciones reales hubiesen sido con toda seguridad más precisos.

La selección de inmuebles se ha realizado en 2009 y ha sido de tipo transversal. El corto período de tiempo en el que las muestras ofertadas son recogidas hace más fácil la comparación entre inmuebles y es por tanto un factor beneficioso para este estudio en particular. Simplificando elementos innecesarios como las series de tiempo, se consigue que la información extraída del análisis sea más clara y precisa.

La selección de inmuebles ha sido realizada tomando como base los anuncios existentes en la página de idealista.com y discriminando únicamente aquellos anuncios sin fotos para la primera selección, ya que éstas resultan imprescindibles para valorar la calidad constructiva del inmueble.

Los datos inicialmente recogidos a través de los anuncios existentes para cada muestra son los siguientes:

- Dirección
- Barrio
- Superficie construida

- Antigüedad del edificio
- Número de dormitorios
- Planta en la que se sitúa el inmueble
- Zonas comunes y anexos vinculados
- Calidades constructivas
- Precio

Imagen 3.6. Muestras seleccionadas

Fuente: Elaboración propia

Tabla 3.1. Base de datos con información inicial

dirección	barrio	m²	antig.	nº dormit.	planta	zonas comun	calidades	€
CORAZÓN DE MARIA, 62	Prosperidad	82	50	2	9	1	2.5	360,000
CANILLAS, 91	Prosperidad	90	50	3	2	0	3	391,000
PADRE JESÚS ORDOÑEZ, 5	Prosperidad	107	40	3	6	2	2	529,000
RODRIGUEZ MARÍN, 3	El Viso	223	20	4	2	1	3	1,112,000
PASEO DE LA HABANA, 168	Nueva España	151	50	4	1	3	3	900,000
CARDENAL SILICEO, 30	Prosperidad	110	50	3	3	0	3	399,000
SERRANO, 220	Hispanoamérica	240	50	5	1	1	4	900,000
COSTA RICA, 5	Hispanoamérica	130	50	3	6	1	3	540,000
AVDA. BURGOS, 28	Castilla	102	30	3	14	5	3	450,000
JOSEP PLA, 16	Ciudad Jardín	112	15	2	9	6	2.5	680,000
POTOSI, 9	Hispanoamérica	81	30	2	11	7	3	550,000
COMANDANTE FRANCO, 4	Nueva España	170	30	4	3	5	3	1,150,000
LUIS CABRERA, 52	Prosperidad	70	60	1	4	2	4	230,000
PASEO DE LOS CEREZOS, 29	Hispanoamérica	500	20	3	2	1	1	2,100,000
SANTA HORTENSIA, 33	Prosperidad	115	50	4	3	1	4	490,000
PASTORA IMPERIO, 5	Castilla	242	30	5	12	4	3	1,050,000
LOPEZ DE HOYOS, 200	Ciudad Jardín	70	3	1	1	3	3	312,000
CARDENAL SILICEO, 2	Prosperidad	70	50	3	5	0	5	317,400
PADRE CLARET, 2	Prosperidad	140	20	3	13	4	3	644,000
MATEO INURRIA, 52	Castilla	200	30	5	3	2	4	970,000
PS. CASTELLANA, 168	Hispanoamérica	167	50	4	7	2	3	810,000
SANTA HORTENSIA, 3	Prosperidad	75	50	2	4	1	3	366,240
LUIS CABRERA, 1	Prosperidad	68	50	2	2	0	3	334,595
RODRIGUEZ MARÍN, 88	Hispanoamérica	134	30	2	5	2	3	660,000
PASTORA IMPERIO, 1	Castilla	260	5	3	17	6	1	1,290,000
VÍCTOR DE LA SERNA, 58	Hispanoamérica	235	40	5	2	3	3	1,200,000
AVDA. BURGOS, 18	Castilla	230	20	5	3	5	4	1,200,000
PRÍNCIPE DE VERGARA, 212	Hispanoamérica	200	30	3	6	4	3	1,050,000
SERRANO, 229	Hispanoamérica	65	20	1	2	0	3	355,000
PS. CASTELLANA, 136	El Viso	241	40	4	2	4	4	1,400,000
GUATEMALA, 10	Nueva España	550	5	4	3	8	1	3,300,000
MARÍA FRANCISCA, 10	El Viso	150	50	4	1	4	4	900,000
PASEO DE LA HABANA, 67	Hispanoamérica	126	20	2	10	3	2	775,000
PRÍNCIPE DE VERGARA, 258	Hispanoamérica	72	30	2	5	4	3	444,000
PLATERÍAS, 8	Nueva España	469	30	3	3	5	2	2,900,000
FELIX BOIX, 4	Nueva España	125	5	3	1	0	2	800,200
ALFONSO RGUEZ. SANTAMARÍA, 23	El Viso	200	40	4	2	4	2	1,400,000
ALUSTANTE, 12	Prosperidad	130	30	3	1	2	1	600,000
PEREZ AYUSO, 8	Ciudad Jardín	51	5	1	1	1	2	239,000
CONCHA ESPINA, 29	Hispanoamérica	76	50	2	3	2	3.5	450,000
PINILLA DEL VALLE, 1	Prosperidad	105	50	3	5	0	3	468,000
PIO XII, 92	Castilla	55	30	1	10	4	2	320,000
MANTUANO, 25	Ciudad Jardín	47	30	1	0	2	2.5	278,000
DOCTOR FLEMING, 44	Nueva España	83	50	3	9	3	3	500,000
PADRE CLARET, 7	Prosperidad	126	50	3	1	1	3.5	560,000
CONSTANCIA, 40	Prosperidad	90	50	3	4	0	3	330,000
CONSTANCIA, 29	Prosperidad	50	20	1	0	0	3.5	184,000
FRANCISCO DE GOYA, 14	Nueva España	165	30	4	1	4	3	980,000
AZULINAS, 7	Nueva España	48	50	1	3	3	3	312,000
PS. CASTELLANA, 118	El Viso	111	50	2	3	2	2	725,000
PS. CASTELLANA, 144	Hispanoamérica	200	50	5	5	2	3	1,375,000
PS. CASTELLANA, 164	Nueva España	280	50	4	5	2	3	1,450,000
CLARA DEL REY, 35	Prosperidad	85	50	2	6	1	3	360,000
DOCTOR FLEMING, 16	Hispanoamérica	260	30	4	3	3	3	1,800,000
CONCHA ESPINA, 29	Hispanoamérica	76	50	2	3	2	3.5	450,000
EUCALIPTOS, 11	Hispanoamérica	170	10	3	5	3	3	950,000
JAVIER FERRERO, 8	Ciudad Jardín	50	3	1	3	1	2	294,496
MAESTRO LASALLE, 35	Nueva España	325	30	4	3	3	3	1,450,000
NIEREMBERG, 27	Ciudad Jardín	67	50	2	3	1	3	250,000
PRIETO UREÑA,5	Nueva España	115	40	3	2	3	2.5	650,000
PADRE DAMIAN, 37	Nueva España	125	20	2	3	4	3.5	780,000
PIO XII, 26	Nueva España	150	40	5	1	4	3.5	894,000
VICTOR ANDRES BELAUNDE, 36	Hispanoamérica	58	40	1	3	1	2	350,000
QUINTILIANO, 8	Prosperidad	70	50	2	1	1	2.5	297,000
CLARA DEL REY, 31	Prosperidad	325	50	4	3	3	3	1,670,000
PS. CASTELLANA, 158	Hispanoamérica	358	50	5	7	1	3	2,300,000
TORPEDERO TUCUMÁN,12	Nueva España	125	40	3	2	2	2	750,000
SANTIAGO BERNABEU, 4	El Viso	350	50	3	7	3	2	3,000,000
MARÍA DE MOLINA, 27	El Viso	350	15	4	5	2	2	3,500,000
PADRE XIFRE, 3	Prosperidad	100	40	1	9	1	2	479,000

MARÍA DE MOLINA, 153	Prosperidad	102	50	3	0	1	3	508,700
CORAZÓN DE MARÍA, 5	Prosperidad	120	40	4	5	0	2.5	599,000
VINAROZ, 8	Ciudad Jardín	60	30	1	2	0	3	310,000
VALPARAISO, 4	Hispanoamérica	100	10	2	2	0	3	559,000
ORURO, 3	Hispanoamérica	80	40	3	1	0	3	450,000
ORURO, 8	Hispanoamérica	120	20	3	2	1	3	670,000

Fuente: elaboración propia

3.5.4. AJUSTE DE PRECIOS SEGÚN LA EVOLUCIÓN DE PRECIOS DE MERCADO

En este caso, al haberse recogido las muestras en un mismo momento temporal no se ha realizado un ajuste por índices específico para cada muestra, sino que simplemente ha sido necesario calcular un solo porcentaje de ajuste que se aplica a la totalidad de las muestras. En la realidad, se establecerán actualizaciones de la base de datos de forma regular, por lo que es importante consignar la fecha de inclusión de cada muestra en la base de datos de cara a la futura actualización de precios según la evolución del mercado.

Las muestras del presente caso práctico expuesto fueron recogidas en 2009 y este trabajo de investigación se presenta en 2015, por lo que el tiempo transcurrido desde la selección de variables resultaría excesivo para una aplicación práctica real. No obstante y a modo meramente explicativo se incluye a continuación un ajuste por índices según la evolución de los precios en el mercado para ajustarlos al momento actual.

Tabla 3.2. Evolución de precios en el distrito de Chamartín

madrid
(euros/m²)

distrito	3q 2009	3q 2010	3q 2011	3q 2012	3q 2013	3q 2014	3q 2015
Chamartín	4.790	4.694	4.423	4.028	3.758	3.652	3.732

Fuente: idealista.com

La variación del tercer trimestre de 2009, momento de recogida de datos comparables, hasta el tercer trimestre de 2015 ha sido del -22,09%.

Tabla 3.3. Base de datos con precios actualizados a Q3 2015

dirección	barrio	m²	antig. años	nº dormit.	planta	zonas comun	calidades	€ Q3 2015
CORAZÓN DE MARIA, 62	Prosperidad	82	50	2	9	1	2.5	280.484
CANILLAS, 91	Prosperidad	90	50	3	2	0	3	304.637
PADRE JESÚS ORDOÑEZ, 5	Prosperidad	107	40	3	6	2	2	412.156
RODRIGUEZ MARÍN, 3	El Viso	223	20	4	2	1	3	866.385
PASEO DE LA HABANA, 168	Nueva España	151	50	4	1	3	3	701.211
CARDENAL SILICEO, 30	Prosperidad	110	50	3	3	0	3	310.870
SERRANO, 220	Hispanoamérica	240	50	5	1	1	4	701.211
COSTA RICA, 5	Hispanoamérica	130	50	3	6	1	3	420.727
AVDA. BURGOS, 28	Castilla	102	30	3	14	5	3	350.605
JOSEP PLA, 16	Ciudad Jardín	112	15	2	9	6	2.5	529.804
POTOSI, 9	Hispanoamérica	81	30	2	11	7	3	428.518
COMANDANTE FRANCO, 4	Nueva España	170	30	4	3	5	3	895.992
LUIS CABRERA, 52	Prosperidad	70	60	1	4	2	4	179.198
PASEO DE LOS CEREZOS, 29	Hispanoamérica	500	20	3	2	1	1	1.636.159
SANTA HORTENSIA, 33	Prosperidad	115	50	4	3	1	4	381.770
PASTORA IMPERIO, 5	Castilla	242	30	5	12	4	3	818.079
LOPEZ DE HOYOS, 200	Ciudad Jardín	70	3	1	1	3	3	243.086
CARDENAL SILICEO, 2	Prosperidad	70	50	3	5	0	5	247.294
PADRE CLARET, 2	Prosperidad	140	20	3	13	4	3	501.755
MATEO INURRIA, 52	Castilla	200	30	5	3	2	4	755.749
PS. CASTELLANA, 168	Hispanoamérica	167	50	4	7	2	3	631.090
SANTA HORTENSIA, 3	Prosperidad	75	50	2	4	1	3	285.346
LUIS CABRERA, 1	Prosperidad	68	50	2	2	0	3	260.691
RODRIGUEZ MARÍN, 88	Hispanoamérica	134	30	2	5	2	3	514.221
PASTORA IMPERIO, 1	Castilla	260	5	3	17	6	1	1.005.069
VÍCTOR DE LA SERNA, 58	Hispanoamérica	235	40	5	2	3	3	934.948
AVDA. BURGOS, 18	Castilla	230	20	5	3	5	4	934.948
PRÍNCIPE DE VERGARA, 212	Hispanoamérica	200	30	3	6	4	3	818.079
SERRANO, 229	Hispanoamérica	65	20	1	2	0	3	276.589
PS. CASTELLANA, 136	El Viso	241	40	4	2	4	4	1.090.772
GUATEMALA, 10	Nueva España	550	5	4	3	8	1	2.571.106
MARÍA FRANCISCA, 10	El Viso	150	50	4	1	4	4	701.211
PASEO DE LA HABANA, 67	Hispanoamérica	126	20	2	10	3	2	603.820
PRÍNCIPE DE VERGARA, 258	Hispanoamérica	72	30	2	5	4	3	345.931
PLATERÍAS, 8	Nueva España	469	30	3	3	5	2	2.259.457
FELIX BOIX, 4	Nueva España	125	5	3	1	0	2	623.454
ALFONSO RGUEZ. SANTAMARÍA, 23	El Viso	200	40	4	2	4	2	1.090.772
ALUSTANTE, 12	Prosperidad	130	30	3	1	2	1	467.474
PEREZ AYUSO, 8	Ciudad Jardín	51	5	1	1	1	2	186.210
CONCHA ESPINA, 29	Hispanoamérica	76	50	2	3	2	3.5	350.605
PINILLA DEL VALLE, 1	Prosperidad	105	50	3	5	0	3	364.630
PIO XII, 92	Castilla	55	30	1	10	4	2	249.319
MANTUANO, 25	Ciudad Jardín	47	30	1	0	2	2.5	216.596
DOCTOR FLEMING, 44	Nueva España	83	50	3	9	3	3	389.562
PADRE CLARET, 7	Prosperidad	126	50	3	1	1	3.5	436.309
CONSTANCIA, 40	Prosperidad	90	50	3	4	0	3	257.111
CONSTANCIA, 29	Prosperidad	50	20	1	0	0	3.5	143.359
FRANCISCO DE GOYA, 14	Nueva España	165	30	4	1	4	3	763.541
AZULINAS, 7	Nueva España	48	50	1	3	3	3	243.086
PS. CASTELLANA, 118	El Viso	111	50	2	3	2	2	564.864
PS. CASTELLANA, 144	Hispanoamérica	200	50	5	5	2	3	1.071.294
PS. CASTELLANA, 164	Nueva España	280	50	4	5	2	3	1.129.729
CLARA DEL REY, 35	Prosperidad	85	50	2	6	1	3	280.484
DOCTOR FLEMING, 16	Hispanoamérica	260	30	4	3	3	3	1.402.422
CONCHA ESPINA, 29	Hispanoamérica	76	50	2	3	2	3.5	350.605
EUCALIPTOS, 11	Hispanoamérica	170	10	3	5	3	3	740.167
JAVIER FERRERO, 8	Ciudad Jardín	50	3	1	3	1	2	229.449
MAESTRO LASALLE, 35	Nueva España	325	30	4	3	3	3	1.129.729
NIEREMBERG, 27	Ciudad Jardín	67	50	2	3	1	3	194.781
PRIETO UREÑA,5	Nueva España	115	40	3	2	3	2.5	506.430
PADRE DAMIAN, 37	Nueva España	125	20	2	3	4	3.5	607.716
PIO XII, 26	Nueva España	150	40	5	1	4	3.5	696.536
VICTOR ANDRES BELAUNDE, 36	Hispanoamérica	58	40	1	3	1	2	272.693
QUINTILIANO, 8	Prosperidad	70	50	2	1	1	2.5	231.400
CLARA DEL REY, 31	Prosperidad	325	50	4	3	3	3	1.301.136
PS. CASTELLANA, 158	Hispanoamérica	358	50	5	7	1	3	1.791.983
TORPEDERO TUCAMÁN,12	Nueva España	125	40	3	2	2	2	584.342
SANTIAGO BERNABEU, 4	El Viso	350	50	3	7	3	2	2.337.370
MARÍA DE MOLINA, 27	El Viso	350	15	4	5	2	2	2.726.931
PADRE XIFRE, 3	Prosperidad	100	40	1	9	1	2	373.200

MARÍA DE MOLINA, 153	Prosperidad	102	50	3	0	1	3	396.340
CORAZÓN DE MARÍA, 5	Prosperidad	120	40	4	5	0	2,5	466.695
VINARÓZ, 8	Ciudad Jardín	60	30	1	2	0	3	241.528
VALPARAISO, 4	Hispanoamérica	100	10	2	2	0	3	435.530
ORURO, 3	Hispanoamérica	80	40	3	1	0	3	350.605
ORURO, 8	Hispanoamérica	120	20	3	2	1	3	522.013

Fuente: Elaboración propia

En España de 1985 a 2007 el incremento medio anual del precio de la vivienda ha sido de más de un 10% de media. Durante los años de la burbuja inmobiliaria este incremento anual superó incluso el 15%.

Imagen 3.8. Evolución de precios medios e incrementos anuales

Fuente: Sociedad de Tasación

En la práctica y como norma general no sería recomendable una profundidad histórica de la base de datos no superior a 3 años debido a que estadísticamente la oscilación de precios transcurrido este tiempo suele ser de más del 30%. No obstante esta profundidad histórica para el modelo deberá establecerse considerando el volumen de transacciones y la evolución de los precios en cada mercado local.

Cabe mencionar que en el caso de tener una referencia de precio de transacción real en otro momento temporal de la vivienda a valorar, esta puede tenerse en cuenta en la valoración aunque dicha referencia sea anterior a la profundidad histórica máxima establecida.

3.5.5. FILTRADO PRELIMINAR

En un filtrado preliminar se procede a eliminar aquellas muestras que pertenecer a otra tipología. En este caso práctico se han eliminado viviendas unifamiliares, así como áticos y dúplex por las importantes diferencias de precios entre estas tipologías y la vivienda piso en esta ubicación. En la práctica, se deberá establecer para cada modelo qué tipologías se engloban.

Tabla 3.4. Filtrado por tipología de vivienda

dirección	barrio	Tipología
CORAZÓN DE MARÍA, 62	Prosperidad	piso
CANILLAS, 91	Prosperidad	piso
PADRE JESÚS ORDOÑEZ, 5	Prosperidad	ático
RODRIGUEZ MARÍN, 3	El Viso	piso
PASEO DE LA HABANA, 168	Nueva España	piso
CARDENAL SILICEO, 30	Prosperidad	piso
SERRANO, 220	Hispanoamérica	piso
COSTA RICA, 5	Hispanoamérica	piso
AVDA, BURGOS, 28	Castilla	piso
JOSEP PLA, 16	Ciudad Jardín	ático
POTOSI, 9	Hispanoamérica	piso
COMANDANTE FRANCO, 4	Nueva España	piso
LUIS CABRERA, 52	Prosperidad	piso
PASEO DE LOS CEREZOS, 29	Hispanoamérica	adosado
SANTA HORTENSIA, 33	Prosperidad	piso
PASTORA IMPERIO, 5	Castilla	piso
LOPEZ DE HOYOS, 200	Ciudad Jardín	piso
CARDENAL SILICEO, 2	Prosperidad	piso
PADRE CLARET, 2	Prosperidad	piso
MATEO INURRIA, 52	Castilla	piso
PS. CASTELLANA, 168	Hispanoamérica	piso
SANTA HORTENSIA, 3	Prosperidad	piso
LUIS CABRERA, 1	Prosperidad	piso
RODRIGUEZ MARÍN, 88	Hispanoamérica	piso
PASTORA IMPERIO, 1	Castilla	dúplex
VÍCTOR DE LA SERNA, 58	Hispanoamérica	pareado
AVDA. BURGOS, 18	Castilla	piso
PRÍNCIPE DE VERGARA, 212	Hispanoamérica	ático
SERRANO, 229	Hispanoamérica	piso
PS. CASTELLANA, 136	El Viso	piso
GUATEMALA, 10	Nueva España	pareado
MARÍA FRANCISCA, 10	El Viso	piso
PASEO DE LA HABANA, 67	Hispanoamérica	piso
PRÍNCIPE DE VERGARA, 258	Hispanoamérica	piso
PLATERÍAS, 8	Nueva España	ático
FELIX BOIX, 4	Nueva España	piso
ALFONSO	El Viso	piso
SANTAMARÍA, 23	Prosperidad	dúplex
ALUSTANTE, 12	Ciudad Jardín	piso
PEREZ AYUSO, 8		

CONCHA ESPINA, 29	Hispanoamérica	piso
PINILLA DEL VALLE, 1	Prosperidad	piso
PIO XII, 92	Castilla	piso
MANTUANO, 25	Ciudad Jardín	piso
DOCTOR FLEMING, 44	Nueva España	piso
PADRE CLARET, 7	Prosperidad	piso
CONSTANCIA, 40	Prosperidad	piso
CONSTANCIA, 29	Prosperidad	piso
FRANCISCO DE GOYA, 14	Nueva España	piso
AZULINAS, 7	Nueva España	piso
PS. CASTELLANA, 118	El Viso	piso
PS. CASTELLANA, 144	Hispanoamérica	piso
PS. CASTELLANA, 164	Nueva España	piso
CLARA DEL REY, 35	Prosperidad	piso
DOCTOR FLEMING, 16	Hispanoamérica	piso
CONCHA ESPINA, 29	Hispanoamérica	piso
EUCALIPTOS, 11	Hispanoamérica	piso
JAVIER FERRERO, 8	Ciudad Jardín	piso
MAESTRO LASALLE, 35	Nueva España	chalet aislado
NIEREMBERG, 27	Ciudad Jardín	piso
PRIETO UREÑA, 5	Nueva España	piso
PADRE DAMIAN, 37	Nueva España	piso
PIO XII, 26	Nueva España	piso
VICTOR ANDRES BELAUNDE, 36	Hispanoamérica	piso
QUINTILIANO, 8	Prosperidad	piso
CLARA DEL REY, 31	Prosperidad	piso
PS. CASTELLANA, 158	Hispanoamérica	piso
TORPEDERO TUCAMÁN, 12	Nueva España	piso
SANTIAGO BERNABEU, 4	El Viso	piso
MARÍA DE MOLINA, 27	El Viso	piso
PADRE XIFRE, 3	Prosperidad	ático
MARÍA DE MOLINA, 153	Prosperidad	piso
CORAZÓN DE MARÍA, 5	Prosperidad	piso
VINAROZ, 8	Ciudad Jardín	piso
VALPARAISO, 4	Hispanoamérica	piso
ORURO, 3	Hispanoamérica	piso
ORURO, 8	Hispanoamérica	piso

Fuente: Elaboración propia

3.5.6. AJUSTE DE PRECIOS POR NEGOCIACIÓN Y ELIMINACIÓN DE ANEXOS VINCULADOS

Algunos de los inmuebles cuentan con anexos vinculados, que se han separado restándole al precio de la vivienda una estimación del precio medio del garaje o trastero en cada barrio.

Las plazas de garaje y los trasteros vinculados a la vivienda se han descontado estimando el valor medio de una plaza de garaje de unos 25 m² en cada barrio. En el caso de los trasteros se ha estimado el mismo valor unitario en euros/m² que para las plazas de garaje, estimándose un tamaño medio de 8 m² por trastero.

Además se ha estimado una negociación media para ajustar los precios de oferta a precios de mercado. En un caso real se deberá contar en la base de datos con comparables de transacciones reales o bien tasaciones realizadas anteriormente por la sociedad de tasación de manera que ofrezcan una referencia de los precios reales de mercado. Para establecer el porcentaje de ajuste se deberán separar las muestras con precios de venta reales de aquellas con precios de oferta y establecer la media para cada *cluster*. De esta forma al comparar la media de oferta con la media de precios reales se puede establecer un diferencia, que será el que se aplicará a cada muestra de oferta.

En este caso práctico, al no disponer de información de transacciones reales se ha estimado un ajuste del 5%.

Tabla 3.5. Negociación de precios de oferta

dirección			€ Q3 2015	€ sin anexos y negociado	€ negociado/m²
CORAZÓN DE MARÍA, 62	garaje		280.484	236.060	2.879
CANILLAS, 91			304.637	289.405	3.216
RODRIGUEZ MARÍN, 3	garaje	trastero	866.385	762.266	3.418
PASEO DE LA HABANA, 168	garaje		701.211	629.575	4.169
CARDENAL SILICEO, 30			310.870	295.327	2.685
SERRANO, 220		trastero	701.211	654.750	2.728
COSTA RICA, 5			420.727	399.690	3.075
AVDA, BURGOS, 28	garaje	trastero	350.605	285.100	2.795
POTOSI, 9	garaje	trastero	428.518	359.117	4.434
COMANDANTE FRANCO, 4	garaje	trastero	895.992	803.217	4.725
LUIS CABRERA, 52			179.198	170.238	2.432
SANTA HORTENSIA, 33		trastero	381.770	353.182	3.071
PASTORA IMPERIO, 5	garaje		818.079	740.600	3.060
LOPEZ DE HOYOS, 200			243.086	230.932	3.299
CARDENAL SILICEO, 2			247.294	234.929	3.356
PADRE CLARET, 2		trastero	501.755	467.168	3.337
MATEO INURRIA, 52			755.749	717.962	3.590
PS. CASTELLANA, 168	garaje	trastero	631.090	551.560	3.303
SANTA HORTENSIA, 3			285.346	271.079	3.614
LUIS CABRERA, 1			260.691	247.656	3.642
RODRIGUEZ MARÍN, 88	garaje	trastero	514.221	440.535	3.288
AVDA. BURGOS, 18	garaje		934.948	888.200	3.862
SERRANO, 229			276.589	262.759	4.042
PS. CASTELLANA, 136			1.090.772	1.036.234	4.300
MARÍA FRANCISCA, 10	garaje		701.211	619.600	4.131
PASEO DE LA HABANA, 67	garaje		603.820	537.054	4.262
PRÍNCIPE DE VERGARA, 258	garaje		345.931	292.059	4.056
FELIX BOIX, 4			623.454	592.282	4.738
ALFONSO RGUEZ. SANTAMARÍA, 23	2x garaje		1.090.772	943.134	4.716
PEREZ AYUSO, 8		trastero	186.210	167.400	3.282
CONCHA ESPINA, 29		trastero	350.605	321.675	4.233
PINILLA DEL VALLE, 1	garaje		364.630	315.998	3.010
PIO XII, 92		trastero	249.319	225.453	4.099
MANTUANO, 25		trastero	216.596	196.266	4.176
DOCTOR FLEMING, 44		trastero	389.562	358.684	4.321
PADRE CLARET, 7	garaje		436.309	384.094	3.048
CONSTANCIA, 40			257.111	244.255	2.714
CONSTANCIA, 29			143.359	136.191	2.724
FRANCISCO DE GOYA, 14			763.541	725.364	4.396

AZULINAS, 7		trastero	243.086	219.532	4.574
PS. CASTELLANA, 118	garaje	trastero	564.864	522.371	4.706
PS. CASTELLANA, 144		trastero	1.071.294	969.755	4.849
PS. CASTELLANA, 164		trastero	1.129.729	1.061.842	3.792
CLARA DEL REY, 35	garaje		280.484	266.460	3.135
DOCTOR FLEMING, 16		trastero	1.402.422	1.284.326	4.940
CONCHA ESPINA, 29	garaje	trastero	350.605	321.675	4.233
EUCALIPTOS, 11		trastero	740.167	655.184	3.854
JAVIER FERRERO, 8	garaje		229.449	187.576	3.752
NIEREMBERG, 27	garaje		194.781	185.042	2.762
PRIETO UREÑA, 5			506.430	444.534	3.866
PADRE DAMIAN, 37	garaje		607.716	540.755	4.326
PIO XII, 26	garaje		696.536	661.709	4.411
VICTOR ANDRES BELAUNDE, 36			272.693	259.058	4.467
QUINTILIANO, 8			231.400	189.430	2.706
CLARA DEL REY, 31	2x garaje	trastero	1.301.136	1.165.779	3.587
PS. CASTELLANA, 158	garaje		1.791.983	1.665.809	4.653
TORPEDERO TUCAMÁN, 12	garaje		584.342	518.550	4.148
SANTIAGO BERNABEU, 4	2x garaje		2.337.370	2.220.501	6.344
MARÍA DE MOLINA, 27		trastero	2.726.931	2.483.235	7.095
MARÍA DE MOLINA, 153			396.340	376.523	3.691
CORAZÓN DE MARÍA, 5	garaje		466.695	443.360	3.695
VINAROS, 8			241.528	229.452	3.824
VALPARAISO, 4			435.530	377.178	3.772
ORURO, 3	garaje		350.605	333.075	4.163
ORURO, 8			522.013	459.337	3.828

Fuente: elaboración propia

3.5.7. NORMALIZACIÓN DE DATOS

Una vez seleccionas las variables a emplear en el caso práctico, se procede a realizar un análisis de las mismas. En primer lugar, han de ser tratados los distintos valores que alcanzan cada una de las variables, considerando que, para optimizar el rendimiento de la red, las cifras con las que esta trabaja han de estar comprendidas dentro de un intervalo cuya amplitud debe ser lo más reducida posible; de no hacerlo así podemos tener problemas de no aprendizaje o establecimiento de bucles, ya que el coeficiente que multiplica a cada variable en la fórmula de la regresión depende directamente del valor de dicha variable. Utilizando este procedimiento se consiguen resultados porcentuales más realistas y fiables.

En este sentido se han intentado mantener los valores entre 0 y 6 para cada uno de los atributos de la mayoría de inmuebles, aunque en casos algunos los valores ya tratados han sobrepasado estos límites dado su carácter atípico ya sea por una superficie construida superior a la de los demás inmuebles de la base de datos o un valor de inmueble (*output*) más alto que para el resto de muestras consideradas. Sin embargo, se ha considerado beneficioso para el aprendizaje incluir estos inmuebles para obtener un modelo afín al mercado real.

Por otra parte se han agrupado las zonas comunes de la vivienda (piscina, zonas verdes, zonas deportivas, etc.) en una única variable para poder minimizar el número de variables a incluir en el modelo y mejorar así el aprendizaje y funcionamiento del mismo.

El número de dormitorios, planta y zonas comunes quedan intactos, sin necesidad de ser normalizados ya que en escasas ocasiones supera el valor máximo considerado (6)

La variable cualitativa de calidades constructivas se ha reflejado con base en un rango numérico del 1 al 5, representando el 1 las calidades más básicas y el 5 las calidades más altas.

Los años de antigüedad se dividen entre 10, la superficie del inmueble entre 50 y el precio por m2 entre 1.000.

No ha sido necesario realizar un autocompletado de datos, ya que se ha dispuesto de datos de partida completos para todos los comparables.

La variable no numérica de barrio será considerada en la fase posterior de zonificación.

Tabla 3.6. Base de datos con variables normalizadas

dirección	barrio	m ² normali z.	antig. norm aliz	nº dormit.	planta	zona s comu n	calidad es	€/m ² ajust ado norm aliz
CORAZÓN DE MARÍA, 62	Prosperidad	1,6	5	2	9	1	2,5	2,9
CANILLAS, 91	Prosperidad	1,8	5	3	2	0	3	3,2
RODRIGUEZ MARÍN, 3	El Viso	4,5	2	4	2	0	3	3,4
PASEO DE LA HABANA, 168	Nueva España	3,0	5	4	1	2	3	4,2
CARDENAL SILICEO, 30	Prosperidad	2,2	5	3	3	0	3	2,7
SERRANO, 220	Hispanoamérica	4,8	5	5	1	0	4	2,7
COSTA RICA, 5	Hispanoamérica	2,6	5	3	6	1	3	3,1
AVDA, BURGOS, 28	Castilla	2,0	3	3	14	3	3	2,8
POTOSI, 9	Hispanoamérica	1,6	3	2	11	5	3	4,4
COMANDANTE FRANCO, 4	Nueva España	3,4	3	4	3	3	3	4,7
LUIS CABRERA, 52	Prosperidad	1,4	6	1	4	2	4	2,4
SANTA HORTENSIA, 33	Prosperidad	2,3	5	4	3	0	4	3,1
PASTORA IMPERIO, 5	Castilla	4,8	3	5	12	3	3	3,1
LOPEZ DE HOYOS, 200	Ciudad Jardín	1,4	0	1	1	3	3	3,3
CARDENAL SILICEO, 2	Prosperidad	1,4	5	3	5	0	5	3,4

PADRE CLARET, 2	Prosperidad	2,8	2	3	13	3	3	3,3
MATEO INURRIA, 52	Castilla	4,0	3	5	3	2	4	3,6
PS. CASTELLANA, 168	Hispanoamérica	3,3	5	4	7	0	3	3,3
SANTA HORTENSIA, 3	Prosperidad	1,5	5	2	4	1	3	3,6
LUIS CABRERA, 1	Prosperidad	1,4	5	2	2	0	3	3,6
RODRIGUEZ MARÍN, 88	Hispanoamérica	2,7	3	2	5	0	3	3,3
AVDA. BURGOS, 18	Castilla	4,6	2	5	3	1	4	3,9
SERRANO, 229	Hispanoamérica	1,3	2	1	2	0	3	4,0
PS. CASTELLANA, 136	El Viso	4,8	4	4	2	4	4	4,3
MARÍA FRANCISCA, 10	El Viso	3,0	5	4	1	3	4	4,1
PASEO DE LA HABANA, 67	Hispanoamérica	2,5	2	2	10	2	2	4,3
PRÍNCIPE DE VERGARA, 258	Hispanoamérica	1,4	3	2	5	3	3	4,1
FELIX BOIX, 4	Nueva España	2,5	1	3	1	0	2	4,7
ALFONSO RGUEZ.	El Viso	4,0	4	4	2	2	2	4,7
SANTAMARÍA, 23	Ciudad Jardín	1,0	1	1	1	0	2	3,3
PEREZ AYUSO, 8								
CONCHA ESPINA, 29	Hispanoamérica	1,5	5	2	3	1	3,5	4,2
PINILLA DEL VALLE, 1	Prosperidad	2,1	5	3	5	0	3	3,0
PIO XII, 92	Castilla	1,1	3	1	10	3	2	4,1
MANTUANO, 25	Ciudad Jardín	0,9	3	1	0	1	2,5	4,2
DOCTOR FLEMING, 44	Nueva España	1,7	5	3	9	2	3	4,3
PADRE CLARET, 7	Prosperidad	2,5	5	3	1	1	3,5	3,0
CONSTANCIA, 40	Prosperidad	1,8	5	3	4	0	3	2,7
CONSTANCIA, 29	Prosperidad	1,0	2	1	0	0	3,5	2,7
FRANCISCO DE GOYA, 14	Nueva España	3,3	3	4	1	4	3	4,4
AZULINAS, 7	Nueva España	1,0	5	1	3	2	3	4,6
PS. CASTELLANA, 118	El Viso	2,2	5	2	3	1	2	4,7
PS. CASTELLANA, 144	Hispanoamérica	4,0	5	5	5	0	3	4,8
PS. CASTELLANA, 164	Nueva España	5,6	5	4	5	1	3	3,8
CLARA DEL REY, 35	Prosperidad	1,7	5	2	6	1	3	3,1
DOCTOR FLEMING, 16	Hispanoamérica	5,2	3	4	3	1	3	4,9
CONCHA ESPINA, 29	Hispanoamérica	1,5	5	2	3	1	3,5	4,2
EUCALIPTOS, 11	Hispanoamérica	3,4	1	3	5	1	3	3,9
JAVIER FERRERO, 8	Ciudad Jardín	1,0	0	1	3	0	2	3,8
NIEREMBERG, 27	Ciudad Jardín	1,3	5	2	3	1	3	2,8
PRIETO UREÑA, 5	Nueva España	2,3	4	3	2	2	2,5	3,9
PADRE DAMIAN, 37	Nueva España	2,5	2	2	3	3	3,5	4,3
PIO XII, 26	Nueva España	3,0	4	5	1	4	3,5	4,4
VICTOR ANDRES								
BELAUNDE, 36	Hispanoamérica	1,2	4	1	3	1	2	4,5
QUINTILIANO, 8	Prosperidad	1,4	5	2	1	0	2,5	2,7
CLARA DEL REY, 31	Prosperidad	6,5	5	4	3	0	3	3,6
PS. CASTELLANA, 158	Hispanoamérica	7,2	5	5	7	0	3	4,7
TORPEDERO TUCAMÁN, 12	Nueva España	2,5	4	3	2	2	2	4,1
SANTIAGO BERNABEU, 4	El Viso	7,0	5	3	7	3	2	6,3
MARÍA DE MOLINA, 27	El Viso	7,0	2	4	5	2	2	7,1
MARÍA DE MOLINA, 153	Prosperidad	2,0	5	3	0	1	3	3,7
CORAZÓN DE MARÍA, 5	Prosperidad	2,4	4	4	5	0	2,5	3,7
VINAROS, 8	Ciudad Jardín	1,2	3	1	2	0	3	3,8
VALPARAISO, 4	Hispanoamérica	2,0	1	2	2	0	3	3,8
ORURO, 3	Hispanoamérica	1,6	4	3	1	0	3	4,2
ORURO, 8	Hispanoamérica	2,4	2	3	2	1	3	3,8

Fuente: Elaboración propia

3.5.8. ELIMINACIÓN DE OUTLIERS

Mediante un filtrado de las muestras se han eliminado aquellas cuyo valor a priori es significativamente superior o inferior al de inmuebles similares (*outliers*) y por tanto dichas muestras no han sido incluidas en la base de datos. También se han eliminados aquellos con datos incongruentes, aunque el precio medio por metro cuadrado estuviese dentro de los rangos establecidos. Es esencial reconocer la sensibilidad de las valoraciones automatizadas a valores atípicos, ya que dichos valores pueden disminuir de forma significativa la eficacia predictiva del modelo; por

este motivo es importante identificar los valores atípicos y eliminarlos de la base de datos.

Tabla 3.7. Base de datos depurada

dirección	barrio	m ² normal izado	antig. Normal izado	nº dorm it.	plant a	zonas comun	calid ades	€/m ² ajusta do y norma lizado
RODRIGUEZ MARÍN, 3	El Viso	4,5	2	4	2	0	3	3,4
PASEO DE LA HABANA, 168	Nueva España	3,0	5	4	1	2	3	4,2
CARDENAL SILICEO, 30	Prosperidad	2,2	5	3	3	0	3	2,7
SERRANO, 220	Hispanoamérica	4,8	5	5	1	0	4	2,7
COSTA RICA, 5	Hispanoamérica	2,6	5	3	6	1	3	3,1
AVDA, BURGOS, 28	Castilla	2,0	3	3	14	3	3	2,8
POTOSI, 9	Hispanoamérica	1,6	3	2	11	5	3	4,4
COMANDANTE FRANCO, 4	Nueva España	3,4	3	4	3	3	3	4,7
LUIS CABRERA, 52	Prosperidad	1,4	6	1	4	2	4	2,4
SANTA HORTENSIA, 33	Prosperidad	2,3	5	4	3	0	4	3,1
PASTORA IMPERIO, 5	Castilla	4,8	3	5	12	3	3	3,1
LOPEZ DE HOYOS, 200	Ciudad Jardín	1,4	0	1	1	3	3	3,3
CARDENAL SILICEO, 2	Prosperidad	1,4	5	3	5	0	5	3,4
PADRE CLARET, 2	Prosperidad	2,8	2	3	13	3	3	3,3
MATEO INURRIA, 52	Castilla	4,0	3	5	3	2	4	3,6
PS. CASTELLANA, 168	Hispanoamérica	3,3	5	4	7	0	3	3,3
SANTA HORTENSIA, 3	Prosperidad	1,5	5	2	4	1	3	3,6
LUIS CABRERA, 1	Prosperidad	1,4	5	2	2	0	3	3,6
RODRIGUEZ MARÍN, 88	Hispanoamérica	2,7	3	2	5	0	3	3,3
AVDA. BURGOS, 18	Castilla	4,6	2	5	3	1	4	3,9
SERRANO, 229	Hispanoamérica	1,3	2	1	2	0	3	4,0
PS. CASTELLANA, 136	El Viso	4,8	4	4	2	4	4	4,3
MARÍA FRANCISCA, 10	El Viso	3,0	5	4	1	3	4	4,1
PASEO DE LA HABANA, 67	Hispanoamérica	2,5	2	2	10	2	2	4,3
PRÍNCIPE DE VERGARA, 258	Hispanoamérica	1,4	3	2	5	3	3	4,1
FELIX BOIX, 4	Nueva España	2,5	1	3	1	0	2	4,7
ALFONSO RGUEZ. SANTAMARÍA, 23	El Viso	4,0	4	4	2	2	2	4,7
PEREZ AYUSO, 8	Ciudad Jardín	1,0	1	1	1	0	2	3,3
CONCHA ESPINA, 29	Hispanoamérica	1,5	5	2	3	1	3,5	4,2
PIO XII, 92	Castilla	1,1	3	1	10	3	2	4,1
MANTUANO, 25	Ciudad Jardín	0,9	3	1	0	1	2,5	4,2
DOCTOR FLEMING, 44	Nueva España	1,7	5	3	9	2	3	4,3
CONSTANCIA, 40	Prosperidad	1,8	5	3	4	0	3	2,7
CONSTANCIA, 29	Prosperidad	1,0	2	1	0	0	3,5	2,7
FRANCISCO DE GOYA, 14	Nueva España	3,3	3	4	1	4	3	4,4
AZULINAS, 7	Nueva España	1,0	5	1	3	2	3	4,6
PS. CASTELLANA, 118	El Viso	2,2	5	2	3	1	2	4,7
PS. CASTELLANA, 144	Hispanoamérica	4,0	5	5	5	0	3	4,8
PS. CASTELLANA, 164	Nueva España	5,6	5	4	5	1	3	3,8
CLARA DEL REY, 35	Prosperidad	1,7	5	2	6	1	3	3,1
DOCTOR FLEMING, 16	Hispanoamérica	5,2	3	4	3	1	3	4,9
CONCHA ESPINA, 29	Hispanoamérica	1,5	5	2	3	1	3,5	4,2
EUCALIPTOS, 11	Hispanoamérica	3,4	1	3	5	1	3	3,9
JAVIER FERRERO, 8	Ciudad Jardín	1,0	0	1	3	0	2	3,8
NIEREMBERG, 27	Ciudad Jardín	1,3	5	2	3	1	3	2,8
PRIETO UREÑA, 5	Nueva España	2,3	4	3	2	2	2,5	3,9
PADRE DAMIAN, 37	Nueva España	2,5	2	2	3	3	3,5	4,3
PIO XII, 26	Nueva España	3,0	4	5	1	4	3,5	4,4
VICTOR ANDRES BELAUNDE, 36	Hispanoamérica	1,2	4	1	3	1	2	4,5
QUINTILIANO, 8	Prosperidad	1,4	5	2	1	0	2,5	2,7
CLARA DEL REY, 31	Prosperidad	6,5	5	4	3	0	3	3,6
PS. CASTELLANA, 158	Hispanoamérica	7,2	5	5	7	0	3	4,7
SANTIAGO BERNABEU, 4	El Viso	7,0	5	3	7	3	2	6,3
MARÍA DE MOLINA, 153	Prosperidad	2,0	5	3	0	1	3	3,7
CORAZÓN DE MARÍA, 5	Prosperidad	2,4	4	4	5	0	2,5	3,7

Fuente: Elaboración propia

3.6. ZONIFICACIÓN

3.6.1. INTRODUCCIÓN

Una de las características primordiales que diferencia al mercado inmobiliario de otros mercados es el hecho de que los bienes inmuebles son únicos e irrepetibles, principalmente debido al vínculo implícito con su emplazamiento específico en el ámbito metropolitano. La principal característica que permite distinguir al mercado inmobiliario de otro tipo de mercados es precisamente la localización única e irreemplazable de cada bien inmueble.

El factor localización tiene una influencia capital en el precio de la vivienda y los resultados demuestran que el impacto en el precio es incluso mayor de lo esperado, especialmente debido a la correlación entre la agrupación determinada de los barrios y el emplazamiento de las zonas prime en el distrito.

El precio del suelo en áreas metropolitanas llega a ser en muchos casos más del 50% del precio total de la vivienda. Por lo tanto, es vital comprender qué factores tienen una influencia más incisiva en el precio final de los inmuebles residenciales. Así, conocer de qué forma afectan determinados factores urbanos al precio de la vivienda puede facilitar un análisis más crítico de los valores de suelo en áreas metropolitanas.

Los factores que determinan el precio de la vivienda pueden clasificarse comúnmente en tres grupos: características estructurales, características de barrio y características de emplazamiento. Las características estructurales incluyen todos aquellos rasgos estructurales y constructivos inherentes a la unidad de vivienda en sí, tales como calidad constructiva, metros cuadrados construidos o número de dormitorios. Las otras dos categorías tratan sobre el valor económico de la localización, en términos de barrio y emplazamiento específico respectivamente.

Los atributos de emplazamiento pueden tener un impacto de hasta el 30% sobre el precio de la vivienda, mientras que los atributos de emplazamiento pueden tener un impacto de hasta el 20% sobre el precio de la vivienda.

Comprobamos por tanto que las características de emplazamiento tienen un papel importante en la determinación de precios de viviendas, mientras que los atributos correspondientes al barrio tienen un peso algo menos significativo en el valor de la vivienda. No obstante, ambos grupos de características tienden a estar a menudo correlacionados y por lo tanto es complicado medir de forma exacta los impactos específicos en relación al precio de los inmuebles.

Dada la importancia que tiene la ubicación en el precio de la vivienda, especialmente en ubicaciones metropolitanas, es necesario establecer zonificaciones que permitan diferenciar zonas homogéneas en las que las viviendas tienen similares características.

Se define la segmentación o estratificación como el proceso mediante el cual se agrupan inmuebles para su posterior modelado y análisis. Los inmuebles residenciales en zonas urbanas se suelen agrupar en zonas de mercado homogéneas, pudiéndose desarrollar un modelo para cada *cluster* o bien desarrollar un único modelo a nivel regional y posteriormente calibrar dicho modelo para cada *cluster* (IAAO, 2003).

En el caso de utilizar el método K-vecinos, mediante el cual se realiza una selección de los comparables más cercanos de forma previa a cada valoración, se debería restringir el campo de búsqueda, de forma que los comparables a utilizar estén localizados en el mismo cluster que el inmueble a valorar.

3.6.2. MÉTODOS DE ZONIFICACIÓN

Cano et al. (1999) combinan un análisis multivariante con un análisis geográfico, siendo clave la variable en la que se base la zonificación. Estos autores introducen en su investigación el concepto de “variable regionalizada” siendo esta una variable espacial de la cual dependen los valores de los bienes inmuebles. Además, los

autores citados establecieron como su variable regionalizada única la calidad global de la vivienda, siendo esta una combinación de las variables constructivas y de localización que tienen una mayor correlación con el precio de la vivienda.

Se considera que esta metodología basada tanto en el análisis de componentes principales como en la teoría de variables regionalizadas es acertada para establecer una zonificación sobre la cual establecer zonas homogéneas para la selección de comparables de viviendas.

La importancia de la localización en el valor final de los inmuebles es reconocido en todos los segmentos del mercado inmobiliario; sin embargo no hay un consenso acerca de qué factores o variables hacen que un barrio o un emplazamiento específico sea más valorado que otro. La falta de investigaciones unánimes y exhaustivas al respecto es consecuencia directa de la dificultad que entraña el análisis de factores blandos en los procedimientos de valoraciones masivas.

Las variables analizadas en la literatura existente varían según el autor, sin que se haya conseguido llegar a un consenso sobre cuáles son los factores relacionados con la ubicación de un inmueble que tienen un mayor peso en el precio de las viviendas. Por tanto no parece razonable que el marco regulatorio deba restringir en este sentido las variables de ubicación a tener en cuenta.

Lo que sí parece lógico es que con aquellas variables de localización que se consideren más estrechamente relacionadas con el valor de la vivienda, e incluyendo en su caso también variables de estructura o constructivas, se deberá generar un atributo único de zonificación (variable regionalizada) que resuma la información que incide en el precio de la vivienda distinguiendo una zona de otra.

La forma de ponderar cada variable no debe ser restringida, ya que además de la metodología utilizada por Cano et al. (1999), existen otras metodologías que darían resultados válidos, como por ejemplo ponderar las variables mediante un árbol de decisión.

También se puede utilizar un modelo de valoración automatizada para crear un mapa de valores. Gallego (2008) muestra en su artículo “Modelos de valoración automatizada” un mapa de valor de suelo realizado mediante una Red Neuronal Artificial para la ciudad de Madrid.

Imagen 3.9. Mapa de valores del suelo de Madrid

Fuente: Julio Gallego (2008)

Otra forma de incluir variables de ubicación en un modelo de valoración estadística es mediante LVRSA (Location Value Response Surface Analysis), que es un modelo basado en coordenadas x-y que forman una malla geográfica de forma que se pueden

analizar las distancias a hitos urbanos que tienen impacto en el valor de los inmuebles. Normalmente se utiliza esta técnica cuando se desarrolla un único modelo.

Se indican a continuación algunas variables que pueden ayudar a evaluar las características de localización. Esta lista es simplemente indicativa, por lo que no se trata de una lista cerrada:

- Hogares de clase alta
- Hogares de clase baja
- Promedio de paro
- Zona envejecida
- Gastos corrientes del hogar
- Gastos asociados a personal y equipamiento del hogar
- Gastos de conservación de las viviendas

Se recomienda asimismo que el conjunto de indicadores o variables utilizado para generar la variable regionalizada de zonificación sea revisada pudiéndose incluir nuevos indicadores en las fases de seguimiento y retroalimentación.

La identificación de zonas homogéneas es un procedimiento importante en el proceso de desarrollo de una herramienta de valoración automatizada. En Reino Unido se ha utilizado como punto de partida la clasificación por códigos postales complementada con información socio-económica aportada por proveedores independientes. Por ejemplo, el proveedor de servicios CACI dispone de la herramienta ACORN de segmentación demográfica, que es usada por algunos proveedores de servicios AVM para establecer sus zonificaciones (RICS, 2013)

3.7. CASO PRÁCTICO DE ZONIFICACIÓN

3.7.1. INTRODUCCIÓN

La metodología se centra en un enfoque deductivo y busca estudiar la problemática de los factores urbanos blandos mediante el estudio de datos.

Los inmuebles ofertados en el mercado inmobiliario ofrecen una fuente de información apta para este caso práctico y, con base en la literatura, se puede afirmar que una base de datos creada a partir de inmuebles existentes en el mercado es la mejor manera de abordar este tipo de análisis. La información recogida de la oferta en el mercado tiene la gran ventaja de proporcionar una amplia selección de inmuebles que han sido de antemano analizados y comparados con inmuebles similares a fin de ofrecer un precio de venta competitivo en el mercado local.

El conjunto de muestras refleja adecuadamente no sólo las condiciones de mercado existentes, sino también una amplia gama de emplazamientos dentro del área de estudio.

Debido a que los datos a recoger deben ser de tipo cuantitativo, es necesario obtener una cantidad de muestras razonablemente extensa para poder analizar el mayor número de variables posible. La base de datos desarrollada para este caso práctico consta de 76 muestras de vivienda situadas en el área de estudio. Dicha cantidad de muestras parece suficiente para este caso práctico, teniendo en cuenta el procedimiento manual para el tratamiento de datos y la limitación temporal.

El área de estudio es, al igual que en el caso práctico relacionado con las variables de estructura y su normalización, el distrito de Chamartín.

3.7.2. SELECCIÓN DE DATOS DE PARTIDA: VARIABLES

Es de vital importancia entender los múltiples atributos urbanos que pueden ser relevantes y conocer los diferentes enfoques a considerar en la fase de modelado.

La literatura relativa al tema ha sido una fuente de información vital acerca de las variables relevantes de emplazamiento y barrio a incluir en el caso práctico. Después de un análisis de variables de entrada relevantes, los siguientes atributos de barrio y emplazamiento fueron considerados como factores determinantes para este caso práctico:

Atributos de barrio:

- Precio medio por metro cuadrado
- Nivel económico de los residentes
- Nacionalidad de los residentes
- Densidad de población

Atributos de emplazamiento:

- Calidad del área local
- Nivel de ruido
- Distancia a un colegio de referencia
- Distancia al hospital
- Distancia a zona comercial
- Distancia a zona prime de oficinas
- Distancia a la M-30
- Distancia a zonas verdes
- Distancia a la estación de metro

3.7.3. SELECCIÓN DE MUESTRAS

Se ha utilizado la misma selección de muestras que en el caso práctico anterior para analizar el impacto de los atributos de barrio y emplazamiento y definir así una zonificación adecuada.

Imagen 3.10. Muestras seleccionadas

Fuente: Elaboración propia

Los datos iniciales de cada muestra fueron completados con información referente a los siguientes atributos:

- Precio medio por metro cuadrado en cada barrio
- Nivel económico de los residentes
- Nacionalidad de los residentes
- Densidad de población por sección censal
- Calidad del área local
- Nivel de ruido
- Calidad constructiva del inmueble

Los datos correspondientes a precio medio por metro cuadrado y barrio han sido obtenidos también en la página web www.idealista.com en base al precio medio de los inmuebles allí ofertados. Los precios medio por m² y barrio son los siguientes:

- 7,033 € por m² en El Viso
- 5,979 € por m² en Nueva España
- 5,565 € por m² en Hispanoamérica
- 4,841 € por m² en Ciudad Jardín
- 4,435 € por m² en Castilla
- 4,349 € por m² en Prosperidad

La información relativa al nivel económico ha sido obtenida por medio de la información de renta familiar disponible per cápita en el censo de población y viviendas.

Tabla 3.8. Renta familiar disponible per cápita

		51.	52.	53.	54.	55.	56.
Características	TOTAL	El Viso	Prosperidad	Ciudad Jardín	Hispanoamérica	Nueva España	Castilla
Renta familiar per cápita							
Euros	19,678	29,206	15,912	15,502	19,891	23,647	18,316
Índice	154.117	228.747	124.622	121.412	155.792	185.206	143.451

FUENTE: Instituto Nacional de Estadística. Censos de población y viviendas 2001.

Los datos incluidos en la base de datos referentes a la nacionalidad de los residentes corresponden al porcentaje de habitantes no europeos por barrio.

Tabla 3.9. Nacionalidad de los residentes en el distrito de Chamartín

Características	Total Distrito	5.1. El Viso	5.2. Prosperidad	5.3. Ciudad Jardín	5.4. Hispanoamérica	5.5. Nueva España	5.6. Castilla
Población a 1-1-2008							
Nacionalidad (Total)	145,088	16,233	38,159	18,828	31,409	23,447	17,012
España	128,037	14,007	33,619	15,954	28,586	20,518	15,353
Unión Europea (15)	3,542	562	709	421	793	715	342
Unión Europea (extensión 25)	313	41	88	56	43	53	32
Otros países OCDE	833	140	158	81	171	193	90
Otros países europeos	1,709	207	477	347	226	283	169
América Latina y Caribe	8,759	1,036	2,572	1,542	1,295	1,427	887
África	750	59	199	218	127	90	57
Asia y Oceanía	1,123	179	332	207	167	161	77
Apátridas	3	1	-	-	-	-	2
No consta	19	1	5	2	1	7	3
% Extranjeros	11.7	13.7	11.9	15.3	9	12.5	9.7

FUENTE: Área de Gobierno de Hacienda y Administración Pública

Los datos de nacionalidad resultantes son, por tanto, los siguientes:

- 10.3% residentes extracomunitarios en El Viso
- 9.4 % residentes extracomunitarios en Nueva España
- 7.3 % residentes extracomunitarios en Hispanoamérica
- 13 % residentes extracomunitarios en Ciudad Jardín
- 7.7 % residentes extracomunitarios en Castilla
- 10% residentes extracomunitarios en Prosperidad

La información correspondiente a la densidad de población se obtuvo a través del Padrón Municipal de habitantes.

Tabla 3.10. Población por distrito y secciones censales

Fecha:		01-03-2010			
Distrito	Barrio	Sección censal	Población total	m ²	Población por cada 100 m ²
Chamartin	El Viso	05001	1,111	295,826	0
Chamartin	El Viso	05002	1,518	418,433	0
Chamartin	El Viso	05003	1,844	190,140	1
Chamartin	El Viso	05004	1,379	69,403	2
Chamartin	El Viso	05005	1,340	46,212	3
Chamartin	El Viso	05006	1,298	67,870	2
Chamartin	El Viso	05007	1,807	127,669	1
Chamartin	El Viso	05008	1,502	224,075	1
Chamartin	El Viso	05009	1,142	103,223	1
Chamartin	El Viso	05010	871	53,819	2
Chamartin	El Viso	05011	816	17,494	5
Chamartin	El Viso	05012	1,157	39,993	3
Chamartin	El Viso	05013	826	34,128	2
Chamartin	Prosperidad	05014	1,104	45,265	2
Chamartin	Prosperidad	05015	1,077	25,573	4
Chamartin	Prosperidad	05016	856	14,164	6
Chamartin	Prosperidad	05017	1,432	23,220	6
Chamartin	Prosperidad	05018	1,335	21,619	6
Chamartin	Prosperidad	05019	950	23,583	4
Chamartin	Prosperidad	05020	1,357	24,107	6
Chamartin	Prosperidad	05021	957	22,365	4
Chamartin	Prosperidad	05022	1,483	21,192	7
Chamartin	Prosperidad	05023	1,394	31,077	4
Chamartin	Prosperidad	05024	981	13,319	7
Chamartin	Prosperidad	05025	1,163	21,798	5

Chamartin	Prosperidad	05026	1,624	25,278	6
Chamartin	Prosperidad	05027	1,884	26,287	7
Chamartin	Prosperidad	05028	2,120	31,858	7
Chamartin	Prosperidad	05029	1,930	38,797	5
Chamartin	Prosperidad	05030	1,372	32,398	4
Chamartin	Prosperidad	05031	1,022	28,289	4
Chamartin	Prosperidad	05032	1,454	44,280	3
Chamartin	Prosperidad	05033	2,047	75,743	3
Chamartin	Prosperidad	05034	1,649	47,898	3
Chamartin	Prosperidad	05035	1,296	15,709	8
Chamartin	Prosperidad	05036	1,750	152,591	1
Chamartin	Prosperidad	05037	1,309	51,196	3
Chamartin	Ciudad Jardin	05038	1,877	237,999	1
Chamartin	Ciudad Jardin	05039	1,389	109,946	1
Chamartin	Ciudad Jardin	05040	724	12,391	6
Chamartin	Ciudad Jardin	05041	1,263	22,394	6
Chamartin	Ciudad Jardin	05042	1,525	36,669	4
Chamartin	Ciudad Jardin	05043	1,317	74,441	2
Chamartin	Ciudad Jardin	05044	1,730	53,610	3
Chamartin	Ciudad Jardin	05045	1,798	30,688	6
Chamartin	Ciudad Jardin	05046	1,772	26,037	7
Chamartin	Ciudad Jardin	05047	1,771	26,839	7
Chamartin	Ciudad Jardin	05048	1,161	22,152	5
Chamartin	Ciudad Jardin	05049	1,220	84,658	1
Chamartin	Hispanoamerica	05051	1,471	60,559	2
Chamartin	Hispanoamerica	05052	1,463	93,147	2
Chamartin	Hispanoamerica	05053	2,087	296,424	1
Chamartin	Hispanoamerica	05054	1,627	64,122	3
Chamartin	Hispanoamerica	05055	1,526	75,961	2
Chamartin	Hispanoamerica	05056	1,491	45,099	3
Chamartin	Hispanoamerica	05057	1,161	47,371	2
Chamartin	Hispanoamerica	05058	1,804	140,769	1
Chamartin	Hispanoamerica	05059	755	15,082	5
Chamartin	Hispanoamerica	05060	1,453	28,530	5
Chamartin	Hispanoamerica	05061	2,179	62,259	3
Chamartin	Hispanoamerica	05062	1,719	60,491	3
Chamartin	Hispanoamerica	05063	826	34,227	2
Chamartin	Hispanoamerica	05064	2,047	103,457	2
Chamartin	Hispanoamerica	05065	1,187	21,904	5
Chamartin	Hispanoamerica	05066	1,352	86,560	2
Chamartin	Hispanoamerica	05067	2,342	213,266	1
Chamartin	Hispanoamerica	05068	1,777	140,040	1
Chamartin	Hispanoamerica	05069	877	67,529	1
Chamartin	Nueva España	05070	1,940	78,891	2
Chamartin	Nueva España	05072	1,607	62,421	3
Chamartin	Nueva España	05073	1,720	42,390	4
Chamartin	Nueva España	05074	1,507	174,076	1
Chamartin	Nueva España	05075	1,666	97,216	2
Chamartin	Nueva España	05076	2,371	269,785	1
Chamartin	Nueva España	05077	1,484	158,123	1
Chamartin	Nueva España	05078	898	92,674	1
Chamartin	Nueva España	05079	1,793	75,163	2
Chamartin	Nueva España	05080	1,054	18,238	6
Chamartin	Nueva España	05081	1,631	107,587	2
Chamartin	Nueva España	05082	2,407	262,821	1
Chamartin	Nueva España	05083	2,075	86,241	2
Chamartin	Nueva España	05084	1,744	261,333	1
Chamartin	Castilla	05085	1,877	330,755	1
Chamartin	Castilla	05086	1,275	42,274	3
Chamartin	Castilla	05087	1,848	69,276	3

Chamartin	Castilla	05088	798	36,884	2
Chamartin	Castilla	05089	775	22,280	3
Chamartin	Castilla	05090	1,021	50,934	2
Chamartin	Castilla	05091	889	87,624	1
Chamartin	Castilla	05092	790	39,442	2
Chamartin	Castilla	05093	740	39,126	2
Chamartin	Castilla	05094	2,597	1,044,529	0
Chamartin	Prosperidad	05095	1,921	39,567	5
Chamartin	Ciudad Jardin	05096	1,455	28,900	5
Chamartin	Hispanoamerica	05097	1,488	43,490	3
Chamartin	Hispanoamerica	05098	1,143	19,204	6
Chamartin	Prosperidad	05099	1,839	119,060	2
Chamartin	Castilla	05100	2,022	252,074	1
Chamartin	Castilla	05101	2,558	150,273	2
Chamartin	Prosperidad	05102	960	21,557	4

FUENTE: Padrón Municipal de Habitantes

Imagen 3.11. Seccionado censal distrito de Chamartín

Fuente: Elaboración propia

Los resultados de densidad de población para cada sección censal son obtenidos situando cada muestra en su sección censal correspondiente. Atributos tales como la calidad del área local o la calidad constructiva han sido normalizados atribuyéndoles valores entre el 1 (mejor calidad) y el 6 (peor calidad)

Para la normalización de los valores correspondientes a la situación II (área local) se han tenido en cuenta factores de calidad dentro de cada barrio, como se detallan a continuación:

El Viso

En 1936 se levantó la colonia de El Viso, un conjunto de viviendas unifamiliares para funcionarios. Hoy día es uno de los barrios más caros y de mayor prestigio de Madrid, pues es muy difícil encontrar chalets independientes y con jardín en una zona tan céntrica y bien situada de la capital.

Es un barrio fundamentalmente residencial donde además de los chalets abundan las construcciones de pocas alturas con viviendas de calidad e incluso de lujo. Sin embargo, apenas hay locales comerciales de ningún tipo, bares o cafeterías.

Entre Joaquín Costa, Velázquez y María de Molina el aire residencial cambia y se convierte en una zona de paso, llena de oficinas y con mucho tráfico y ruido.

Hispanoamérica

Hispanoamérica ocupa una franja bastante ancha y en ella conviven edificios y zonas distintas. La parte más cercana al Paseo de la Castellana es una zona de oficinas, donde también se encuentra el Ministerio de Economía y el estadio Santiago Bernabéu.

Detrás del estadio se extiende un barrio tranquilo y residencial, donde escasean los comercios, con casas de los 50 y 60 y viviendas de calidad. Si se cruza la calle Príncipe de Vergara hacia la m-30 se ven viviendas de los años 50 que se alternan con edificios más modernos, algunos incluso de lujo.

Entre Alfonso XIII y la m-30 se encuentra una antigua colonia de chalets de principios de siglo que hoy día constituye una zona exclusiva y cara.

Nueva España

Nueva España es un barrio tranquilo y residencial con viviendas caras y de calidad, pero las viviendas pertenecen a épocas distintas y hay zonas diferenciadas.

En la parte cercana a Pío XII se encuentran los edificios más modernos de los 70 y 80, viviendas caras con jardines y piscina. También quedan varias colonias de principios de siglo con chalets que hoy son muy caros. Por aquí escasean las tiendas, que se concentran en Pío XII, Príncipe de Vergara y la Plaza del Perú.

El entorno del paseo de la Habana, Apolonio Morales y Francisco Suárez es una zona residencial con edificios caros de los 70 y 80 que tienen pocos vecinos y a menudo zonas comunes amplias con anexos varios. En las zonas cercanas al paseo de la Castellana, abundan en cambio los edificios de los años 50 de mayor altura, los comercios y las oficinas.

Ciudad Jardín

Ciudad Jardín fue uno de los barrios que nació en la década de 1920 con la construcción de viviendas unifamiliares modestas en lo que entonces eran las afueras de Madrid.

Hoy día, además de la colonia de unifamiliares, en el barrio hay viviendas de los años 50, 60 y 70. Las más antiguas se sitúan en las calles más estrechas del centro del barrio, que siguen un trazado desordenado. Las viviendas más amplias y de mayor calidad se encuentran en las cercanías del Parque de Berlín. Esta parte es residencial, bastante tranquila y con escasos comercios, mientras que López de Hoyos es muy bullicioso y está lleno de tiendas y tráfico.

En Príncipe de Vergara se concentran varios edificios modernos de oficinas en las cercanías del Auditorio Nacional de Música.

Castilla

Este barrio se desarrolló a principios de siglo con la creación de colonias residenciales, como la colonia rosales que fue construida para empleados municipales en 1929. En décadas posteriores se crearon edificios con viviendas de calidad y se mantuvo el aire tranquilo y residencial del barrio.

Esta situación se vio alterada con la construcción de la Estación de Chamartín en la década de los años 70. En los años 80 el barrio recobró un gran auge con la construcción en el área de la Avenida de Burgos de edificios de viviendas de gran calidad y precio alto. Hoy la estación tiene nuevos accesos y comunicaciones con el resto de líneas de cercanías.

Hoy conviven estas tres zonas (las colonias, las construcciones de los 50 y 60 y las viviendas más nuevas de la Avenida de Burgos) sin que haya un espíritu de barrio que las unifique. Donde más abundan los pequeños comercios es en el área comprendida entre el Paseo de la Castellana y las calles Mauricio Legendre y Agustín de Foxá

Prosperidad

Prosperidad tiene una fuerte identidad de barrio. Su arteria principal es la calle López de Hoyos, ruidosa y llena de tráfico, donde abundan multitud de pequeños comercios de todo tipo. En la zona cercana a López de Hoyos predominan los edificios de los años 50. Sin embargo, en la zona de Corazón de María los edificios son más modernos, de más alturas y las viviendas de mayor calidad.

El nivel de ruido en cada muestra ha sido establecido teniendo como base el mapa informativo sobre niveles acústicos publicado por el Ayuntamiento de Madrid. Los valores asignados también están basados en una escala del 0 al 6, coincidiendo con los siguientes valores reales:

- 0 para niveles de ruido inferiores a 50 dB (A)
- 1 para niveles de ruido entre 50 y 55 dB (A)
- 2 para niveles de ruido entre 55 y 60 dB (A)
- 3 para niveles de ruido entre 60 y 65 dB (A)
- 4 para niveles de ruido entre 65 y 70 dB (A)

- 5 para niveles de ruido entre 70 y 75 dB (A)
- 6 para niveles de ruido superiores a 75 dB (A)

Imagen 3.12. Nivel de ruido en el distrito

Fuente: Ayuntamiento de Madrid y elaboración propia

Tabla 3.11. Base de datos con información suplementaria

dirección	barrio	precio medio (€/m²)	nivel económico (€)	% países no UE	densidad población	calidad local	ruido	calidad constr activa
CORAZÓN DE MARÍA, 62	Prosperidad	4,349	15,912	10.0	5	2.5	5	2.5
CANILLAS, 91	Prosperidad	4,349	15,912	10.0	7	3.0	2	3.0
PADRE JESÚS ORDOÑEZ, 5	Prosperidad	4,349	15,912	10.0	4	4.0	4	2.0
RODRIGUEZ MARÍN, 3	El Viso	7,033	29,206	10.3	1	2.0	4	3.0
PASEO DE LA HABANA, 168	Nueva España	5,979	23,647	9.4	2	1.5	5	3.0
CARDENAL SILICEO, 30	Prosperidad	4,349	15,912	10.0	6	3.0	3	3.0
SERRANO, 220	Hispanoamérica	5,565	19,891	7.3	2	1.0	5	4.0
COSTA RICA, 5	Hispanoamérica	5,565	19,891	7.3	3	2.0	1	3.0
AVDA, BURGOS, 28	Castilla	4,435	18,316	7.7	2	3.0	5	3.0
JOSEP PLA, 16	Ciudad Jardín	4,841	15,502	13.0	2	2.0	3	2.5
POTOSI, 9	Hispanoamérica	5,565	19,891	7.3	3	2.0	4	3.0
COMANDANTE FRANCO, 4	Nueva España	5,979	23,647	9.4	1	3.0	3	3.0
LUIS CABRERA, 52	Prosperidad	4,349	15,912	10.0	6	5.0	4	4.0
PASEO DE LOS CEREZOS, 29	Hispanoamérica	5,565	19,891	7.3	1	2.0	4	1.0
SANTA HORTENSIA, 33	Prosperidad	4,349	15,912	10.0	0	3.0	4	4.0
PASTORA IMPERIO, 5	Castilla	4,435	18,316	7.7	0	3.0	6	3.0
LOPEZ DE HOYOS, 200	Ciudad Jardín	4,841	15,502	13.0	1	3.0	5	3.0
CARDENAL SILICEO, 2	Prosperidad	4,349	15,912	10.0	6	4.0	4	5.0
PADRE CLARET, 2	Prosperidad	4,349	15,912	10.0	1	3.0	3	3.0
MATEO INURRIA, 52	Castilla	4,435	18,316	7.7	1	3.0	5	4.0
PS. CASTELLANA, 168	Hispanoamérica	5,565	19,891	7.3	5	1.0	5	3.0
SANTA HORTENSIA, 3	Prosperidad	4,349	15,912	10.0	3	4.0	3	3.0
LUIS CABRERA, 1	Prosperidad	4,349	15,912	10.0	4	4.0	4	3.0
RODRIGUEZ MARÍN, 88	Hispanoamérica	5,565	19,891	7.3	4	3.0	4	3.0
PASTORA IMPERIO, 1	Castilla	4,435	18,316	7.7	6	1.0	6	1.0
VÍCTOR DE LA SERNA, 58	Hispanoamérica	5,565	19,891	7.3	4	2.0	4	3.0
AVDA. BURGOS, 18	Castilla	4,435	18,316	7.7	5	3.0	5	4.0
PRÍNCIPE DE VERGARA, 212	Hispanoamérica	5,565	19,891	7.3	5	2.0	5	3.0
SERRANO, 229	Hispanoamérica	5,565	19,891	7.3	2	2.0	2	3.0
PS. CASTELLANA, 136	El Viso	7,033	29,206	10.3	5	1.0	5	4.0
GUATEMALA, 10	Nueva España	5,979	23,647	9.4	3	2.0	3	1.0
MARÍA FRANCISCA, 10	El Viso	7,033	29,206	10.3	3	3.0	3	4.0
PASEO DE LA HABANA, 67	Hispanoamérica	5,565	19,891	7.3	5	2.0	5	2.0
PRÍNCIPE DE VERGARA, 258	Hispanoamérica	5,565	19,891	7.3	5	2.0	5	3.0
PLATERÍAS, 8	Nueva España	5,979	23,647	9.4	3	3.0	3	2.0
FELIX BOIX, 4	Nueva España	5,979	23,647	9.4	2	1.0	3	2.0
ALFONSO RGUEZ SANTAMARÍA, 23	El Viso	7,033	29,206	10.3	2	1.0	4	2.0
ALUSTANTE, 12	Prosperidad	4,349	15,912	10.0	7	4.0	3	1.0
PEREZ AYUSO, 8	Ciudad Jardín	4,841	15,502	13.0	7	4.0	3	2.0
CONCHA ESPINA, 29	Hispanoamérica	5,565	19,891	7.3	2	1.0	4	3.5
PINILLA DEL VALLE, 1	Prosperidad	4,349	15,912	10.0	2	3.5	5	3.0
PIO XII, 92	Castilla	4,435	18,316	7.7	0	2.0	5	2.0
MANTUANO, 25	Ciudad Jardín	4,841	15,502	13.0	3	3.0	2	2.5
DOCTOR FLEMING, 44	Nueva España	5,979	23,647	9.4	3	2.0	4	3.0
PADRE CLARET, 7	Prosperidad	4,349	15,912	10.0	4	2.5	4	3.5
CONSTANCIA, 40	Prosperidad	4,349	15,912	10.0	7	4.0	2	3.0
CONSTANCIA, 29	Prosperidad	4,349	15,912	10.0	5	4.0	4	3.5
FRANCISCO DE GOYA, 14	Nueva España	5,979	23,647	9.4	1	2.5	2	3.0
AZULINAS, 7	Nueva España	5,979	23,647	9.4	2	3.0	3	3.0
PS. CASTELLANA, 118	El Viso	7,033	29,206	10.3	2	1.0	5	2.0
PS. CASTELLANA, 144	Hispanoamérica	5,565	19,891	10.3	1	1.0	5	3.0
PS. CASTELLANA, 164	Nueva España	5,979	23,647	9.4	2	1.0	5	3.0
CLARA DEL REY, 35	Prosperidad	4,349	15,912	10.0	3	3.0	3	3.0

DOCTOR FLEMING, 16	Hispanoamérica	5,565	19,891	7.3	1	2.0	4	3.0
CONCHA ESPINA, 29	Hispanoamérica	5,565	19,891	7.3	2	1.0	1	3.5
EUCALIPTOS, 11	Hispanoamérica	5,565	19,891	7.3	1	3.0	6	3.0
JAVIER FERRERO, 8	Ciudad Jardín	4,841	15,502	13.0	1	3.0	4	2.0
MAESTRO LASALLE, 35	Nueva España	5,979	23,647	9.4	1	2.5	3	3.0
NIEREMBERG, 27	Ciudad Jardín	4,841	15,502	13.0	7	2.5	4	3.0
PRIETO UREÑA, 5	Nueva España	5,979	23,647	9.4	1	3.0	2	2.5
PADRE DAMIAN, 37	Nueva España	5,979	23,647	9.4	4	2.0	5	3.5
PIO XII, 26	Nueva España	5,979	23,647	9.4	1	3.0	5	3.5
VICTOR ANDRES BELAUNDE, 36	Hispanoamérica	5,565	19,891	7.3	2	2.0	4	2.0
QUINTILIANO, 8	Prosperidad	4,349	15,912	10.0	4	3.5	3	2.5
CLARA DEL REY, 31	Prosperidad	4,349	15,912	10.0	3	3.0	4	3.0
PS. CASTELLANA, 158	Hispanoamérica	5,565	19,891	7.3	1	1.0	5	3.0
TORPEDERO TUCAMÁN, 12	Nueva España	5,979	23,647	9.4	2	3.0	2	2.0
SANTIAGO BERNABEU, 4	El Viso	7,033	29,206	10.3	2	1.0	5	2.0
MARÍA DE MOLINA, 27	El Viso	7,033	29,206	10.3	0	1.0	5	2.0
PADRE XIFRE, 3	Prosperidad	4,349	15,912	10.0	3	2.0	4	2.0
MARÍA DE MOLINA, 153	Prosperidad	4,349	15,912	10.0	2	1.0	5	3.0
CORAZÓN DE MARÍA, 5	Prosperidad	4,349	15,912	10.0	3	2.5	4	2.5
VINAROS, 8	Ciudad Jardín	4,841	15,502	13.0	5	3.5	3	3.0
VALPARAISO, 4	Hispanoamérica	5,565	19,891	7.3	3	3.0	3	3.0
ORURO, 3	Hispanoamérica	5,565	19,891	7.3	5	3.0	3	3.0
ORURO, 8	Hispanoamérica	5,565	19,891	7.3	3	3.0	3	3.0

Fuente: Elaboración propia

El último grupo de atributos corresponde a las distancias desde cada muestra a los siguientes servicios:

- Colegio de referencia
- Hospital
- Zona comercial
- Zona prime de oficinas
- M-30
- Zonas verdes
- Estación de metro

Las variables de distancia que corresponden a distancias han sido ponderadas según gradientes radiales. Distancias entre 0 y 200 metros tienen el valor de 1. El valor 2 equivale a distancias entre 200 y 500 metros. Asimismo el valor 3 se aplica a

distancias entre 550 y 1.000 metros y el 4 a distancias entre 1.000 y 2.000 metros. El valor máximo es 5 y equivale a distancias superiores a 2.000 metros.

La información referente a colegios de referencia en el área se obtuvo a través de un estudio que recoge anualmente los 100 mejores colegios en la Comunidad de Madrid. Cinco de estos colegios, situados dentro del distrito o en barrios colindantes, han sido tenidos en cuenta para el presente caso práctico:

Tabla 3.12. Colegios

COLEGIOS
1. Ramón y Cajal
2. San Agustín
3. Chamberí
4. Patrocinio de San José
5. Arturo Soria

Fuente: <http://www.elmundo.es/especiales/2009/03/aula/100colegios/madrid.html>

Los tres hospitales incluidos son hospitales generales situados en el distrito o bien en barrios colindantes.

Tabla 3.13. Hospitales

HOSPITALES
1. San Rafael
2. San José
3. La Paz

Fuente: Elaboración propia

Teniendo en cuenta que el distrito es principalmente residencial, sólo hay una zona comercial destacada dentro de su ámbito, situada en Hispanoamérica muy cerca de El Viso. Por este motivo, se ha incluido también la zona comercial de AZCA, próxima a El Viso y que comprende dos centros comerciales y varios locales comerciales aislados.

Tabla 3.14. Zonas comerciales

ZONAS COMERCIALES
1. “La esquina del Bernabéu”
2. AZCA (“El Corte Inglés”, “Moda Shopping”, etc.)

Fuente: Elaboración propia

Asimismo, las distancias a zonas prime de oficinas han sido calculadas respecto a AZCA y al Cuatro Torres Business Area CTBA, ambas fuera del distrito pero muy próximas a él.

Tabla 3.15. Zonas prime de oficinas

ZONAS PRIME OFICINAS
1. AZCA
2. Cuatro Torres Business Area

Fuente: Elaboración propia

La autopista o carretera más cercana al distrito es la M-30, que delimita el área de estudio por el norte y el este, lindando así con cuatro de los seis barrios que conforman el distrito.

Tabla 3.16. M-30

AUTOPISTA
1. M-30

Fuente: Elaboración propia

El siguiente atributo corresponde a zonas verdes públicas y parques con una superficie superior a 20.000 metros cuadrados. Además de los parques dentro del distrito, se han incluido el Parque Norte y el Parque de la Ventilla dada su extensión (más de 250.000 metros cuadrados entre los dos parques) y su cercanía al distrito de Chamartín.

Tabla 3.17. Zonas verdes

ZONAS VERDES Y PARQUES
1. Parque Plaza de Castilla
2. Parque Valparaíso
3. Parque Santas Martas
4. Parque de Berlín
5. Parque Félix Rodríguez de la Fuente
6. Zona verde en C/ Padre Claret
7. Parque La Ventilla
8. Parque Norte

Fuente: Elaboración propia

Por último, un total de 16 estaciones de metro están situadas o bien dentro del distrito o muy cerca de sus límites.

Tabla 3.18. Estaciones de metro

ESTACIONES DE METRO
1. Cruz del Rayo
2. Prosperidad
3. Alfonso XIII
4. Concha Espina
5. Avenida de América
6. Gregorio Marañón
7. Nuevos Ministerios
8. Santiago Bernabéu
9. Colombia
10. Cuzco
11. Avenida de la Paz
12. Pío XII
13. Plaza de Castilla
14. Chamartín
15. Bambú
16. Begoña

Fuente: Elaboración propia

El siguiente conjunto de planos ilustra las distancias relativas entre las muestras y los servicios públicos descritos:

Imagen 3.13. Distancia a colegio de referencia

Imagen 3.14. Distancia a hospital

Fuente: Elaboración propia

Imagen 3.15. Distancia a zona comercial

Imagen 3.16. Distancia a zona prime de oficinas

Fuente: Elaboración propia

Imagen 3.17. Distancia a M-30

Fuente: Elaboración propia

Imagen 3.18. Distancia a zonas verdes

Imagen 3.19. Distancia a estaciones de metro

Fuente: Elaboración propia

La siguiente tabla muestra las variables de distancia para cada muestra:

Tabla 3.19. Base de datos con variables de distancia

	dist. colegi o ref.	dist. hospita l	dist. zona comerci al	dist. oficina s prime	dist. zonas 30 s	dist. verde metr o
CORAZÓN DE MARÍA, 62	5	3	4	5	2	1
CANILLAS, 91	2	2	4	4	3	2
PADRE JESÚS ORDOÑEZ, 5	4	1	4	4	4	3
RODRIGUEZ MARÍN, 3	4	3	3	3	4	3
PASEO DE LA HABANA, 168	3	3	3	3	4	2
CARDENAL SILICEO, 30	4	2	4	4	3	2
SERRANO, 220	2	1	2	3	4	2
COSTA RICA, 5	4	3	4	4	3	1
AVDA, BURGOS, 28	5	3	5	3	1	3
JOSEP PLA, 16	4	2	4	4	4	3
POTOSI, 9	3	2	3	4	4	1
COMANDANTE FRANCO, 4	4	4	4	4	3	3
LUIS CABRERA, 52	4	1	4	4	4	3
PASEO DE LOS CEREZOS, 29	3	4	4	5	1	1
SANTA HORTENSIA, 33	4	3	4	5	3	1
PASTORA IMPERIO, 5	5	3	5	4	1	4
LOPEZ DE HOYOS, 200	4	3	4	5	2	2
CARDENAL SILICEO, 2	4	2	4	4	3	2
PADRE CLARET, 2	4	3	4	5	2	1
MATEO INURRIA, 52	4	4	4	3	3	2
PS. CASTELLANA, 168	3	3	3	3	4	2
SANTA HORTENSIA, 3	4	3	4	4	3	2
LUIS CABRERA, 1	4	1	4	4	4	3
RODRIGUEZ MARÍN, 88	3	1	3	4	4	1
PASTORA IMPERIO, 1	5	3	5	3	1	3
VÍCTOR DE LA SERNA, 58	3	3	4	4	3	1
AVDA. BURGOS, 18	5	3	5	3	1	3
PRÍNCIPE DE VERGARA, 212	3	2	3	4	4	1
SERRANO, 229	3	2	3	4	4	2
PS. CASTELLANA, 136	2	2	1	1	5	2
GUATEMALA, 10	4	3	4	4	3	2
MARÍA FRANCISCA, 10	4	3	3	4	4	3
PASEO DE LA HABANA, 67	2	3	2	3	4	1
PRÍNCIPE DE VERGARA, 258	3	2	3	4	4	2
PLATERÍAS, 8	4	4	4	4	3	3
FELIX BOIX, 4	3	4	4	4	4	1
ALFONSO RGUEZ. SANTAMARÍA, 23	2	2	2	3	4	2
ALUSTANTE, 12	4	2	4	4	3	2
PEREZ AYUSO, 8	4	2	4	4	4	2
CONCHA ESPINA, 29	2	1	1	3	4	1
PINILLA DEL VALLE, 1	5	1	4	4	4	3
PIO XII, 92	5	3	4	3	2	3
MANTUANO, 25	4	2	4	4	4	2
DOCTOR FLEMING, 44	3	3	3	4	4	2
PADRE CLARET, 7	4	3	4	5	2	1
CONSTANCIA, 40	4	2	4	4	3	3
CONSTANCIA, 29	4	2	4	4	4	3

FRANCISCO DE GOYA, 14	3	3	3	4	4	2	2
AZULINAS, 7	4	4	3	4	4	1	3
PS. CASTELLANA, 118	3	3	1	1	5	2	2
PS. CASTELLANA, 144	2	2	2	2	5	2	2
PS. CASTELLANA, 164	3	3	3	3	4	2	1
CLARA DEL REY, 35	4	2	4	4	3	2	2
DOCTOR FLEMING, 16	1	2	2	2	5	2	2
CONCHA ESPINA, 29	2	1	2	3	4	1	3
EUCALIPTOS, 11	3	4	4	5	1	2	2
JAVIER FERRERO, 8	4	3	4	4	3	2	2
MAESTRO LASALLE, 35	4	4	4	4	2	3	2
NIEREMBERG, 27	4	3	4	4	3	1	2
PRIETO UREÑA,5	3	4	4	4	1	3	3
PADRE DAMIAN, 37	3	3	3	3	4	2	2
PIO XII, 26	4	4	4	4	3	3	2
VICTOR ANDRES BELAUNDE, 36	3	2	3	3	4	1	2
QUINTILIANO, 8	4	1	4	4	4	3	2
CLARA DEL REY, 31	4	2	4	4	4	2	2
PS. CASTELLANA, 158	2	3	3	3	5	3	1
TORPEDERO TUCAMÁN,12	4	3	4	4	3	2	2
SANTIAGO BERNABEU, 4	2	2	1	1	5	2	1
MARÍA DE MOLINA, 27	4	3	3	3	5	4	3
PADRE XIFRE, 3	4	1	4	4	4	3	2
MARÍA DE MOLINA, 153	4	2	4	4	4	3	1
CORAZÓN DE MARÍA, 5	4	2	4	4	3	2	2
VINAROS, 8	3	2	4	4	4	2	1
VALPARAISO, 4	3	1	3	3	4	1	2
ORURO, 3	3	1	3	3	4	1	2
ORURO, 8	3	2	3	4	4	2	1

Fuente: Elaboración propia

3.7.4. NORMALIZACIÓN DE DATOS

Al igual que en las variables de estructura, se ha llevado a cabo una normalización de los datos de partida.

Para normalizar las variable de nivel económico y precio medio por metro cuadrado, se han tomado como referencia los valores medios de cada barrio, asignando el valor 1 al barrio de mayor nivel económico o precio respectivamente y el valor 6 al de menor nivel económico o precio. Los valores de los demás barrios se reparten de manera proporcional a la amplitud del intervalo.

La variable nacionalidad se normaliza sumando los porcentajes de residentes extracomunitarios en cada barrio. Dichos porcentajes varían entre el 13% y el 7.3%, por lo tanto la normalización simplemente traslada estas variaciones a la escala del 1 al 6 asignando el valor 6 al barrio con mayor porcentaje (13%). El valor mínimo es 0.3 y corresponde a un 7.3% de residentes extracomunitarios.

La densidad de población corresponde al número de residentes por cada 100 metros cuadrados de suelo y el valor de salida normalizado refleja un ratio de 1/200.000 respecto al precio real de venta en euros.

3.7.5. ANÁLISIS DE PRINCIPALES VARIABLES

Se ha realizado un análisis de los datos de barrio y emplazamiento mediante una regresión lineal.

Como resultado cabe concluir que en este caso, de los atributos de estructura y barrio analizados, aquellos que tienen un mayor peso en el valor de los inmuebles residenciales son los siguientes:

Atributos de barrio:

- Nivel económico de los residentes: 3,46%
- Nacionalidad de los residentes 2,38%
- Distancia a un colegio de referencia: 3,74%
- Distancia a M-30: 11,19%
- Distancia a zonas verdes: 3,73%
- Distancia a estación de metro: 5,31%

En la siguiente tabla se muestran los 10 factores que contribuyen de forma más notable al precio de la vivienda; Juntos suman más del 88,5% del precio final:

Tabla 3.20. Coeficientes de regresión de las principales variables

tipo de atributo	atributos de estructura				atributos de emplazamiento				atributos de barrio	
ESTIMACION LINEAL	plantas edificio	Nro. anexos	superf. construida	calidad constructiva	dist. metro	dist. zona verde	dist. M-30	dist. colegio	nacionalidad	nivel económ.
coeficientes	0,1219	0,1630	2,8379	-0,1574	0,2964	0,2081	0,6249	-0,2090	0,1330	-0,1935
error estándar	0,1108	0,0993	0,1865	0,2310	0,2577	0,2373	0,2027	0,3030	0,0986	0,4342
	0,9318	1,0820	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
	37,5772	55	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
	879,8322	64,3885	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
coeficientes porcentuales	2,18%	2,92%	50,81%	2,82%	5,31%	3,73%	11,19%	3,74%	2,38%	3,46%

En esta tabla, a cada una de las 10 variables se le asigna un coeficiente. La cuarta fila indica el margen de error estándar para cada coeficiente correspondiente. Las siguientes filas muestran las estadísticas propias de la regresión y la última fila revela el porcentaje de contribución al precio final de cada variable.

Las estadísticas correspondientes a esta regresión son:

- Coeficiente de determinación = $R^2 = 0.9318 = 93.18\%$
- Error estándar de la correlación = 0.0993
- Estadístico F = 37.5772
- Grados de libertad = $Df = 55 = v2$
- Suma del cuadrado de la regresión = 879.8322
- Suma del cuadrado del error = 64.3885

Ya que el coeficiente de determinación y el error estándar guardan correlación, para determinar la validez del modelo es posible centrar el análisis exclusivamente en el coeficiente de determinación.

Cuanto menor sea la suma del cuadrado del error en comparación con la suma del cuadrado total, mayor será el coeficiente de determinación r^2 , el cual indica con qué grado de coherencia entre la ecuación resultante y sus variables. (Microsoft Office, 2010)

El coeficiente de determinación establece el grado de analogía entre variables y el modelo. Por consiguiente, cuanto más próximo a 1 (100%), mejor será el modelo. En este caso, el coeficiente corresponde al 93,18%, lo cual verifica la validez del modelo. Por lo general, resultados superiores al 80% se consideran satisfactorios.

El indicador estadístico F y los grados de libertad (Df) se utilizan para evaluar en qué medida puede haberse obtenido un valor alto para F por casualidad. Asumiendo un valor de α igual a 0.05, hallamos el valor crítico para F. Si el F obtenido en la regresión es mucho mayor que el F crítico, es extremadamente improbable que el valor F y, por tanto el resultado de la regresión, se haya obtenido fortuitamente. (Microsoft Office, 2010)

En este caso, los valores de F, v1 y v2 serían:

- $F = 37.5772$
- $v1 = 76 \text{ (número de muestras)} - 55 \text{ (grados de libertad)} - 1 = 20$
- $v2 = \text{grados de libertad} = 55$

Siguiendo las indicaciones y asumiendo un valor normal de $\alpha = 0,05$, el desarrollo de la fórmula de distribución obtiene un F crítico de 4.97213E-25. Se puede deducir por lo tanto que es francamente improbable que los resultados se hayan obtenido de forma fortuita.

Estos resultados han sido obtenidos utilizando la fórmula de estimación lineal de Excel y la base de datos normalizada sin variables de distancia.

3.7.6. ANÁLISIS DE VARIABLES DE DISTANCIA

Además de la regresión lineal con Excel, se ha desarrollado una regresión paralela con la herramienta WEKA (Waikato Environment for Knowledge Analysis) a fin de

evaluar el efecto que tienen los distintos gradientes de distancia en el precio final. La regresión con Excel fue incapaz de calcular una fórmula para la base de datos que incluía estas variables por su formato y por este motivo se ha tenido que hacer esta parte del análisis con otro programa.

La regresión lineal obtenida con el programa WEKA se ha utilizado únicamente para dibujar los gráficos correspondientes a los 4 atributos de emplazamiento más relevantes (distancias a colegio de referencia, M-30, estación de metro y zonas verdes) Excepto estos gráficos, el resto de resultados e interpretación de los mismos están basados en la regresión desarrollada en Excel.

La fórmula de regresión lineal obtenida con la herramienta WEKA para la base de datos con variables de distancia es la siguiente:

=== Run information ===

Scheme: weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8

Relation: training diss

Instances: 70

Attributes: 45

Test mode: user supplied test set: 6 instances

=== Classifier model (full training set) ===

Linear Regression Model

Y =

-0.4127 * X2 +
3.4523 * X7 +
3.3984 * X8 +
0.6088 * X12 +
0.6406 * X13 +
-4.6916 * X16 +

$$\begin{aligned} & -4.1029 * X_{17} + \\ & 1.0666 * X_{19} + \\ & 4.2118 * X_{21} + \\ & 1.3626 * X_{22} + \\ & -0.9001 * X_{24} + \\ & -1.8496 * X_{26} + \\ & -1.3704 * X_{27} + \\ & 0.8921 * X_{30} + \\ & 2.8378 * X_{34} + \\ & 0.6689 * X_{37} + \\ & 2.3195 * X_{39} + \\ & -0.1699 * X_{40} + \\ & 0.3344 * X_{43} + \\ & 2.0466 \end{aligned}$$

En esta fórmula las variables corresponden a los siguientes atributos (aquellos atributos para los cuales se han elaborado gráficos de gradientes y que son considerados en la regresión lineal aparecen subrayados):

- Y = precio en euros
- X_2 = nivel económico
- X_7 = distancia inferior a 200 metros a colegio de referencia
- X_8 = distancia entre 200 y 500 metros a colegio de referencia
- X_{12} = distancia inferior a 200 metros a hospital
- X_{13} = distancia entre 200 y 500 metros a hospital
- X_{16} = distancia inferior a 200 metros a zona comercial
- X_{17} = distancia entre 200 y 500 metros a zona comercial
- X_{19} = distancia entre 1.000 y 2.000 metros a zona comercial
- X_{21} = distancia inferior a 200 metros a zona prime de oficinas
- X_{22} = distancia entre 200 y 500 metros a zona prime de oficinas
- X_{24} = distancia entre 1.000 y 2.000 metros a zona prime de oficinas
- X_{26} = distancia inferior a 200 metros a M-30
- X_{27} = distancia entre 200 y 500 metros a M-30
- X_{30} = distancia superior a 2.000 metros a M-30
- X_{34} = distancia entre 1.000 y 2.000 metros a zonas verdes
- X_{37} = distancia entre 500 y 1.000 metros a estación de metro
- X_{39} = superficie construida

- X40 = antigüedad del edificio
- X43 = número de anexos

Como resultado, de entre las variables analizadas en este caso práctico, los factores urbanos de barrio y emplazamiento que tienen un mayor impacto en el precio de la vivienda son nivel económico de los residentes, nacionalidad, distancia a colegio de referencia, distancia a M-30, distancia a zonas verdes y distancia a estación de metro.

En cuanto a las variables relacionadas con distancia, en el caso de distancia a M-30 y distancia a colegio de referencia, el análisis específico de distancias muestra específicamente el impacto para cada franja de valores, teniendo la cercanía a colegios de referencia un impacto positivo en las distancias más cortas y en el caso de la M-30 mostrando un impacto negativo en distancias cortas y positivo a partir de los 2.000 metros de distancia.

En el caso de las distancias a zonas verdes y metro, el análisis de distancias no muestra un impacto significativo en función del gradiente, ya que en la fórmula sólo interviene una franja de distancia. El resultado es muy poco relevante en el caso de la distancia al metro, cuyo peso para la única franja de distancia que interviene en la fórmula es inferior a 1.

A continuación se muestran los pesos ponderados de los principales atributos de emplazamiento y barrio:

- 3,46% nivel económico de los residentes
- 2,38% nacionalidad (porcentaje de residentes extracomunitarios en cada barrio)
- 3,74% distancia a colegio de referencia
- 11,19% distancia a M-30
- 3,73% distancia a zonas verdes
- 5,31% distancia a estación de metro

A continuación se analizan los resultados obtenidos para la variable que tiene un mayor impacto en el precio, que es la distancia a la M-30. Las distancias han sido convertidas a franjas de distancia para llevar a cabo la evaluación del impacto de las diferentes distancias a puntos de interés.

En la zona de estudio la distancia a la M-30 es el segundo atributo con mayor influencia en el precio, después de la superficie construida, con un porcentaje del 11,2% del precio. El signo del coeficiente correspondiente es positivo, lo cual se traduce en un incremento del precio a medida que la distancia a la M-30 aumenta. Aquellos inmuebles situados cerca de la M-30 son fuertemente penalizados en términos económicos y a su vez las viviendas situadas a mayor distancia de la M-30 son altamente valoradas.

El análisis de gradientes refleja que los precios de las propiedades situadas a menos de 500 metros de la M-30 sufren un importante descenso en los precios. En cambio, aquellas viviendas situadas a más de 2.000 metros de la M-30 tienen un valor añadido.

Imagen 3.20. Gradiente de distancia a la M-30 respecto al precio

Fuente: Elaboración propia

La distancia más amplia a la M-30 corresponde en este caso a las zonas cerca del Paseo de la Castellana y del centro de la ciudad. La proximidad al centro y a la Castellana son aspectos cruciales que ayudan a explicar el elevado impacto que tiene la distancia a la M-30 sobre el precio y el signo de su coeficiente.

Además, el área situada a más de 2.000 metros de la M-30 coincide casi en su totalidad con el barrio de El Viso. Los valores al alza en este barrio en comparación con el resto de barrios del distrito se basan en los orígenes tan diferentes de El Viso, el hecho de que aún mantiene su baja densidad de población y la escasez de viviendas unifamiliares en una zona tan céntrica y bien comunicada.

Los factores discutidos pueden razonablemente ayudar a explicar por qué la distancia a la M-30 tiene un impacto tan pronunciado en el precio de la vivienda en esta área de estudio.

3.7.7. RESULTADO ZONIFICACIÓN

Atendiendo a los resultados obtenidos del análisis de atributos de emplazamiento, se ha elaborado un mapa en el cual se analiza el impacto a diferentes hitos dentro del distrito. Estas localizaciones han sido halladas teniendo en cuenta las variables de emplazamiento incluidas en este caso práctico. El sombreado verde muestra las localizaciones óptimas, mientras que el rojo, muestra aquellas localizaciones que sufren el impacto negativo de la M-30 (menos de 500 m).

Imagen 3.21. Análisis de emplazamientos

Fuente: Elaboración propia

De la información mostrada en el mapa se concluye que El Viso, barrio más precio medio más elevado en el distrito, está en buena parte dentro de la zona sombreada en verde.

La parte oeste del barrio de Hispanoamérica presenta el sombreado verde más denso, hecho en parte atribuible a la inclusión en dicho barrio del único colegio de referencia y la amplia distancia que separa a esta zona de la M-30

El barrio de Castilla presenta un sombreado rojo en gran parte de su área.

Tomando como referencia el análisis de localizaciones, así como la estructura urbana de cada zona, la delimitación de secciones censales existentes y la delimitación entre barrios, se ha llevado a cabo la clusterización del distrito de Chamartín.

Sobre el plano de zonificación que se indica a continuación se han trasladado también los resultados del análisis de variables de distancia de forma aproximada.

Imagen 3.22. Clusters propuestos para el distrito de Chamartín

Fuente: Elaboración propia

3.8. CONSTRUCCIÓN DE MODELOS

La minería de datos nos proporciona toda una serie de técnicas para la selección, clasificación, evaluación e interpretación de la información, para poder conseguir que, dado un gran volumen de datos se consiga extraer un conocimiento y así poder predecir futuros comportamientos o acciones.

Mediante las diferentes etapas en que podríamos dividir el proceso de minería de datos (preprocesado, selección de características, extracción de conocimiento, interpretación y evaluación) conseguimos pasar de una secuencia de datos a un modelo de conocimiento.

Existen varios algoritmos que resultan adecuados para la valoración de inmuebles por métodos estadísticos. Es difícil determinar a priori cual será el modelo de valoración más adecuado en cada caso, por lo que lo más habitual es testear varios algoritmos y luego evaluar el grado de confianza que ofrece cada uno.

El grado de confianza dependerá del adecuado ajuste de cada fórmula al mercado local analizado, tomando como base para establecer la formulación los inmuebles que se hayan incluido en la base de datos, las características que se hayan analizado sobre los mismos y la agrupación de inmuebles que se haya establecido con base en el uso, localización u otros factores.

Algunos algoritmos presentan problemas en mercado complejos. Otros reducen su nivel de confianza si se cuenta con una base de inmuebles comparables limitada o ante información poco representativa o atípica. A continuación se describen los algoritmos más habituales en valoración automatizada así como sus aplicaciones más idóneas con base en los datos de partida, al tipo de mercado a analizar y al grado de conocimiento técnico de la persona que realiza el modelado.

3.8.1. REGRESIÓN LINEAL

La regresión lineal es un método de análisis hedónico. La regresión hedónica se utiliza comúnmente en econometría para determinar el valor estimado de las características individuales que definen la unidad a valorar. Consiguientemente, el bien a valorar puede ser reducido a la suma de las partes que lo constituyen. El análisis de regresión es uno de los métodos más comunes para desarrollar un análisis hedónico.

Imagen 3.23. Regresión lineal

Fuente: Revista Catastro

A través del análisis de regresión, el precio de un bien es calculado como el sumatorio de los valores individuales ponderados correspondientes a cada una de las características inherentes a dicho bien. De forma general, la función asigna un vector a cada atributo para definir así la contribución de cada uno de ellos al precio final.

Los modelos de regresión tienen ciertas ventajas sobre otros métodos, tales como la claridad y sencillez del procedimiento, el extenso uso de este método durante décadas y la familiaridad y entendimiento del método entre profesionales del sector.

Es un sistema sencillo y eficaz que relaciona el valor con las variables mediante una ecuación sencilla. Aunque se trata de un método sencillo, el método hedónico de regresión permite obtener resultados fiables y la precisión de los resultados puede llegar a ser alta si se cuenta con información fiable y un mercado no excesivamente complejo.

Una de sus desventajas es que es un sistema poco tolerante a errores en los datos de partida. También pierde fiabilidad en aquellos casos en los que la base de datos de comparables no es representativa de los inmuebles en el mercado para la zona geográfica sobre la que se realizan las valoraciones.

Por otra parte, presenta problemas de modelado en mercados complejos debido a la simplicidad de la fórmula. Debido a esto es posible que para ciertos mercados complejos la fórmula no se ajuste adecuadamente a la realidad del mercado.

Cabe destacar que una adecuada zonificación permite recurrir a modelos sencillos, como la regresión lineal, al delimitar zonas homogéneas en cuanto a precios y características de los inmuebles y eliminar el exceso de heterogeneidad en el mercado. Por este motivo es importante realizar una zonificación correcta que defina de forma apropiada los diferentes mercados locales homogéneos, especialmente en el caso de que se vaya a recurrir a algoritmos de valoración más sencillos, como la regresión lineal.

Además, es esencial reconocer la sensibilidad del método de la regresión a valores atípicos, ya que dichos valores pueden disminuir de forma significativa la eficacia predictiva del modelo. No obstante una de las ventajas de los modelos de regresión es que los resultados mejoran significativamente tras la correcta eliminación de *outliers* o valores atípicos.

Resumiendo, El método de regresión lineal estima el coeficiente de relación de las variables independientes con la variable dependiente (valor del inmueble). Constituye

una forma sencilla y adecuada de tratar aspectos relacionados con la valoración inmobiliaria de forma clara.

3.8.2. INTELIGENCIA ARTIFICIAL

La Inteligencia Artificial es un campo dentro de la informática que estudia los procesos de aprendizaje basados en la inteligencia y comportamiento humano. Dentro de la inteligencia artificial, las redes neuronales son el modelo más extendido en el campo de las valoraciones estadísticas.

Una red neuronal es un sistema de interconexión de neuronas que colaboran entre sí para dar solución a un problema. Una vez adaptada para un problema particular, la red neuronal usa una función numérica para ponderar cada conexión de la red. El usuario presenta nuevos ejemplos al sistema, causando que los pesos se alteren. El estado final de los nodos de salida determina el resultado.

Imagen 3.24. Red neuronal artificial

Fuente: Revista Catastro

Existen muchos tipos de red neuronal, pero uno de los más extendidos es el Perceptrón Multicapa con Retropropagación. Este algoritmo es una red jerárquica

consistente en capas de unidades de procesamiento completamente interconectadas que permite determinar los pesos para cada conexión en la red usando una función numérica, siendo el nodo de salida el que determina el resultado final.

El procedimiento de aprendizaje de retropropagación implementa una especie de gradiente descendente y la modificación de pesos en sucesivas etapas de aprendizaje fuerza a la red a permanecer tan pequeña como se pueda, previniendo el sobreajuste de los datos de entrenamiento.

Se trata de un sistema mucho más flexible y versátil que la regresión múltiple ya que su formulación no está tan restringida y permite un ajuste más exacto a las características del mercado inmobiliario. Esto ocurre especialmente en mercados complejos en los que la relación lineal entre las variables independientes y la dependiente puede no reflejar correctamente las complejidades de ciertos mercados y en los cuales las redes neuronales pueden alcanzar resultados con mucha mayor precisión.

La fórmula de valoración resultante suele ser mucho más compleja que la obtenida mediante una regresión lineal, pero a su vez es de fácil aplicación una vez obtenida. No obstante es necesario un alto conocimiento técnico para calibrar el modelo. También hay que subrayar que los cálculos internos que realiza la red son a menudo complejos y difíciles de exponer lo cual dificulta la explicación a terceros.

Una de las principales ventajas de las redes neuronales es que ofrecen una gran tolerancia a los errores en los datos de partida. Esto es especialmente relevante, sobre todo teniendo en cuenta que a menudo la información del mercado inmobiliario está incompleta o procede de fuentes poco fiables. Esto es especialmente problemático en España, donde no existen bases de datos de información inmobiliaria pública y, por lo tanto, gran parte de la información que recogen las sociedades de tasación es información de oferta, que no ofrece una imagen ajustada de los precios de cierre de las operaciones.

Existen otros modelos de inteligencia artificial que se pueden aplicar al campo de las valoraciones estadísticas o masivas, tales como los algoritmos genéticos o la lógica difusa.

La aplicación de métodos de lógica difusa en valoraciones automatizada comenzó en 2008 y es, por tanto, mucho más reciente que otros modelos de valoración. González C. (2011) define la lógica difusa de la siguiente manera: *“La Lógica Difusa proporciona un mecanismo de inferencia que permite simular los procedimientos de razonamiento humano en sistemas basados en el conocimiento. La teoría de la lógica difusa proporciona un marco matemático que permite modelar la incertidumbre de los procesos cognitivos humanos de forma que pueda ser tratable por un computador.”*

La lógica difusa se basa en la teoría de conjuntos difusos, que relaciona una función de pertenencia para cada variable dentro de un conjunto difuso. Mediante sistemas de inferencia (si (proposición difusa), entonces (proposición difusa)) obtenemos un conjunto de áreas solapadas, de las cuales el centroide nos dará el resultado final al problema a resolver.

Según la definición de Koza, J. (1992), un algoritmo genético es *“un algoritmo matemático altamente paralelo que transforma un conjunto de objetos matemáticos individuales con respecto al tiempo usando operaciones modeladas de acuerdo al principio Darwiniano de reproducción y supervivencia del más apto, y tras haberse presentado de forma natural una serie de operaciones genéticas de entre las que destaca la recombinación sexual. Cada uno de estos objetos matemáticos suele ser una cadena de caracteres (letras o números) de longitud fija que se ajusta al modelo de las cadenas de cromosomas, y se les asocia con una cierta función matemática que refleja su aptitud.”*

Los algoritmos genéticos se emplean frecuentemente en sistemas híbridos en combinación con otros algoritmos.

3.8.3. CLASIFICADORES K-VECINOS MÁS CERCANOS

Estos modelos se basan en el reconocimiento de patrones y en el cálculo de la distancia entre dos casos comparables para medir su grado de similitud. Clasifican un inmueble comparándolo con una base de datos de ejemplos pre-clasificados.

La principal suposición que se hace es que inmuebles similares tendrán clasificaciones similares. Estos modelos no generan una fórmula y realizan poco trabajo en la etapa de aprendizaje; en los casos más simples tan solo se almacenan los ejemplos en memoria, transfiriendo el esfuerzo al momento de clasificar un nuevo inmueble, cuando el sistema debe decidir cuáles de los ejemplos memorizados debe utilizar para hacer la clasificación.

El algoritmo determina cuáles son los comparables más parecidos utilizando la entropía, o contenido de información de los comparables, como medida de distancia entre ellas. Básicamente, este sistema replica la metodología de valoración por comparación que utilizan los valoradores.

Imagen 3.25. Clasificador K-vecinos

Fuente: scielo.org

Una de sus principales fortalezas es que mediante una metodología sencilla es capaz de alcanzar resultados muy precisos.

El principal inconveniente de estos modelos es que es necesario contar con un gran volumen de muestras de mercado y que estas muestras sean representativas del área local donde se va a utilizar el modelo. Sería por tanto un modelo muy adecuado para todos los casos si los valoradores dispusieran de acceso a bases de datos de información pública. En caso contrario, este modelo no ofrecerá resultados precisos en aquellas zonas donde el mercado sea menos activo y haya menos información de transacciones comparables.

El valorador debe decidir el número de “vecinos” que intervendrán en el cálculo.

3.8.4. MODELOS HÍBRIDOS

En los modelos híbridos intervienen varios algoritmos de forma que al complementarse uno pueda suplir las carencias del otro.

Estos modelos pueden estar integrados de forma libre (loosely-coupled) cuando se realizan varios procesos de forma independiente por cada uno de los algoritmos, o fuerte (tightly-coupled), cuando ambos algoritmos colaboran para realizar la misma tarea.

Imagen 3.26. Modelo híbrido loosely-coupled

Fuente: emerald-library.com

Imagen 3.27. Modelo híbrido tightly-coupled

Fuente: emerald-library.com

3.8.5. PARTICIÓN DE BASES DE DATOS

Se crearán bases de datos específicas para vivienda colectiva y garaje

Cada una de estas tipologías deberá contar con una base de datos de entrenamiento, validación y test

Es importante disponer de una metodología de modelización estadística contrastada y, a ser posible, implementada con éxito en otros sectores de negocio.

Imagen 3.28. Partición de datos

Fuente: WebMining consultores

3.8.6. SEGUIMIENTO Y CALIBRACIÓN

El modelo se debe actualizar, retroalimentar, calibrar y ajustar constantemente, permitiendo obtener valores adecuados y adaptados en el tiempo. Se trata de un proceso iterativo y continuo de mejora y aprendizaje

La construcción de un modelo analítico, es un proceso continuo que debe ser revisado en el tiempo para anticipar ajustes y cambios que se produzcan en el mercado inmobiliario.

Se deben establecer procedimientos para el seguimiento y la validación posterior de los modelos en el horizonte temporal, en forma de alertas cuando el modelo se está desajustando, bien por un cambio en el mercado o de la operativa del dato (en el futuro puede que sea necesario considerar variables adicionales con las que la valoración podría ganar en precisión)

Al tratarse de una valoración estadística, siempre llevará implícito un error, pero éste se sabrá con anticipación, gracias al conjunto de test y por tanto será un error controlado.

La modelización es un proceso continuo y vivo, ya que tanto los datos como las condiciones del mercado inmobiliario, son cambiantes en el tiempo. Mediante el proceso de retroalimentación y actualización continua de la base de datos y de los modelos, se garantiza que las valoraciones no queden obsoletas.

Con cada nueva carga de comparables en la base de datos, se deberá realizar un re-entrenamiento. El re-entrenamiento de los modelos garantiza el aprendizaje de los nuevos datos introducidos en cada actualización.

3.9. EVALUACIÓN DE MODELOS

3.9.1. INTRODUCCIÓN

Es necesario evaluar y comparar los diferentes modelos estadísticos, estableciendo cuál de ellos resuelve de manera más eficaz el problema planteado con el menor error. Este análisis se debería realizar con cada nueva carga de comparables y al menos una vez al año.

Además es necesario incluir en el propio informe de valoración el rango de confianza de la misma.

La fiabilidad de la valoración dependerá en gran medida de:

- La semejanza y proximidad entre el inmueble a valorar y los comparables utilizados.
- La homogeneidad en la zona en la que se ubica el inmueble
- La desviación que ofrece el algoritmo utilizado para la valoración respecto a la base de datos utilizada.
- La calidad de los datos de partida

No existe a nivel mundial un estándar que establezca la forma de calcular el rango de confianza. La consecuencia es que para los usuarios finales es complicado comparar valoraciones automatizadas de diferentes proveedores o incluso formarse una opinión clara de si los rangos de confianza reflejados en los informes son adecuados para la finalidad de la valoración.

Por ejemplo algunos proveedores mide el nivel de confianza de 1 a 7, otros en cambio utilizan un nivel de incertidumbre. Otros proveedores utilizan como indicador del nivel de fiabilidad la desviación estándar o un grado de confianza de 0 a 100. En este

sentido, debemos promover la unificación de criterios con el fin de ofrecer una mayor transparencia y facilitar el análisis de los resultados de la valoración a sus usuarios.

Se recomienda por este motivo seguir las instrucciones del International Association of Assessing Officers (I.A.A.O.) para el cálculo de unos rangos de confianza entendibles y reconocidos a nivel global.

Julio Gallego en su artículo “Modelos de valoración automatizada” también hace referencia a los estándares estadísticos sobre ratios de la IAAO, en los que se basan las recomendaciones que esta misma asociación hace en sus estándares AVM relativas a rangos de confianza. “Su mérito consiste en establecer unas condiciones mínimas que debe cumplir la valoración para poder considerarla correcta. Estas condiciones se refieren a medidas estadísticas sencillas, con límites numéricos. Otra ventaja es que tienen un alto grado de aceptación en muchos países.” (Gallego, 2008)

Los estándares de modelos de valoración automatizada de la I.A.A.O. se aprobaron en septiembre de 2003 y constituyen los principales estándares de valoración automatizada a nivel mundial.

Los siguientes apartados se basan completamente en el estándar de valoración AVM (Standard on Automated Valuation Models) y en los estándares estadísticos sobre ratios (Standard on Ratio Studies), ambos publicados por la International Association of Assessing Officers.

3.9.2. ANÁLISIS DE RATIOS

Los análisis de ratios son una forma fiable de medir la fiabilidad de un modelo. Aunque la I.A.A.O. recomienda que el análisis de ratios se adapte en cuanto a su nivel de detalle en función de cual sea su finalidad, todos los análisis de ratios deberían incluir al menos medidas de centralidad, de dispersión y de fiabilidad estadística.

Tomando la partición de la base de datos correspondiente a “Test”, debemos realizar una valoración de los inmuebles que contiene la muestra a través del modelo de valoración automatizada desarrollado. Al tratarse de inmuebles de nuestra base de datos, conocemos el precio de venta de dichos inmuebles.

Dividiendo el valor calculado entre el precio de venta real de cada inmueble obtenemos el ratio para cada uno de los inmuebles. Podemos realizar una representación gráfica de la distribución de todos los ratios calculados.

A continuación, tomando como datos de partida los ratios calculados para cada una de las muestras de la base de datos de test, se deberán calcular los siguientes parámetros:

- Medida de centralidad
- Medida de dispersión
- Diferencial de precios
- Intervalos de confianza

3.9.3. MEDIDAS DE CENTRALIDAD

La medida de centralidad se basa en la tendencia central. Los principales parámetros relativos a la tendencia central son la media, la mediana y la media ponderada. Se evalúa la tendencia central mediante el análisis de uno o varios de estos parámetros.

La mediana es uno de los parámetros más utilizados para evaluar la tendencia central de los modelos, ya que es menos sensible a valores extremos que otras medidas de centralidad (la media es muy sensible a valores extremos).

En el caso de la media aritmética tiene más peso el valor de los inmuebles valorados que el número de inmuebles, por ello en el cálculo de este parámetro tienen mayor peso los inmuebles de mayor valor.

De acuerdo al “Standard on Automated Valuation Modelos (IAAO, 2003), las medidas de centralidad deberán estar entre el 0,9 y el 1,1 para que se pueda considerar que el modelo ofrece un suficiente grado de fiabilidad.

3.9.4. INTERVALOS DE CONFIANZA

Los intervalos de confianza miden el grado de fiabilidad que podemos otorgar a la valoración de la muestra. Consiste en dos valores que delimitan un rango de ratios calculados con base en un parámetro de tendencia central.

El ratio entre valores calculados y precios para cada *cluster* (agrupaciones de inmuebles similares con base en su ubicación, tipología, etc.), debe estar dentro del $\pm 5\%$ del ratio global para la totalidad de *clusters*.

Para realizar este análisis utilizamos los intervalos de confianza. Este estándar se cumple si los intervalos de confianza de el parámetro escogido de tendencia central al 95% para cada *cluster* quedan solapados dentro del $\pm 5\%$ del nivel global.

3.9.5. MEDIDAS DE DISPERSIÓN

Las medidas de dispersión reflejan la variabilidad horizontal de los ratios. Estos parámetros miden cómo de uniformes son entre sí los ratios. Cuanto menor sea su valor, más uniforme es la fiabilidad entre distintas valoraciones. Sin embargo, valores extremadamente bajos pueden indicar problemas de cálculo o muestras poco representativas del conjunto.

Uno de los parámetros más utilizados en este sentido es el coeficiente de dispersión, que se calcula dividiendo la desviación absoluta media entre la mediana. Según el “Standard on Ratio Studies” (IAAO, 2013), el coeficiente de dispersión debería estar entre el 5 y el 10% en zonas nuevas y más homogéneas y entre 5 y 15% en zonas más antiguas y heterogéneas. En zonas rurales se permite un COD máximo de hasta el 20%.

Tabla 3.21. Rangos coeficiente de dispersión

Table 1-3. Ratio Study Uniformity Standards indicating acceptable general quality*

Type of property—General	Type of property—Specific	COD Range**
Single-family residential (including residential condominiums)	Newer or more homogeneous areas	5.0 to 10.0
Single-family residential	Older or more heterogeneous areas	5.0 to 15.0
Other residential	Rural, seasonal, recreational, manufactured housing, 2–4 unit family housing	5.0 to 20.0
Income-producing properties	Larger areas represented by large samples	5.0 to 15.0
Income-producing properties	Smaller areas represented by smaller samples	5.0 to 20.0
Vacant land		5.0 to 25.0
Other real and personal property		Varies with local conditions

These types of property are provided for guidance only and may not represent jurisdictional requirements.

** Appraisal level for each type of property shown should be between 0.90 and 1.10, unless stricter local standards are required.*

PRD's for each type of property should be between 0.98 and 1.03 to demonstrate vertical equity.

PRD standards are not absolute and may be less meaningful when samples are small or when wide variation in prices exist. In such cases, statistical tests of vertical equity hypotheses should be substituted (see table 1-2).

*** CODs lower than 5.0 may indicate sales chasing or non-representative samples.*

Fuente: Standard on Ratio Studies (IAAO, 2013)

Otros parámetros de medida de dispersión incluyen la desviación absoluta media, la desviación típica o estándar y el coeficiente de variación.

3.9.6. DIFERENCIAL DE PRECIOS

El diferencial de precios nos da una indicación de la variabilidad vertical de los ratios, es decir, si hay diferencias entre cómo se valoran inmuebles de valores altos y de valores bajos. Se calcula dividiendo la media entre la media ponderada de los ratios.

Si el resultado es menor que 0,98 la valoración es progresiva, lo cual indica que en general el modelo está valorando a niveles más bajos que mercado los inmuebles más baratos y a niveles superiores a mercado los inmuebles más caros.

Si el resultado es mayor que 1,03 la valoración es regresiva, lo cual indica que los pisos más baratos están siendo sobrevalorados por el modelo, mientras que para los inmuebles de valores altos el modelo está penalizando su valor por debajo del nivel de mercado real.

3.10. CASO PRÁCTICO DE EVALUACIÓN DE MODELOS

A continuación se presenta un caso práctico de evaluación de modelos, realizada sobre una muestra de 9 inmuebles de test:

Tabla 3.22. Ejemplo *evaluación* de modelos

muestras	Valor calculado(eur)	Precio de venta (eur)	Ratio (V/P)
1	360.000	367.000	0,981
2	255.000	211.000	1,209
3	225.000	235.000	0,957
4	1.200.000	1.195.000	1,004
5	267.000	270.000	0,989
6	225.000	257.000	0,875
7	300.000	274.000	1,095
8	330.000	351.000	0,940
9	960.000	990.000	0,970

Numero de muestras	9
Valor total estimado	4.122.000
Precio total	4.150.000
Valor estimado medio	458.000
Precio medio	461.111

Media de ratios	1,002
Mediana de ratios	0,981
Media ponderada de ratios	0,993

Desviación absoluta media	6,2%
Coeficiente de dispersión (COD)	6,3%

Diferencial de precios (PRD)	1,01
-------------------------------------	-------------

Desviación estándar	0,097
Coeficiente de variación (COV)	9,6%
intervalo de confianza media (95%)	0,935 1,069
intervalo confianza mediana (95%)	0,942 1,088
intervalo confianza media ponderada (95%)	0,930 1,057
Fecha de análisis	07/09/09
Uso de inmuebles analizados	Residencial plurifamiliar

Fuente: Elaboración propia

- Las tres medidas de centralidad se encuentran entre el 0,9 y el 1,1.
- El coeficiente de dispersión se encuentra entre el 5% y el 10%
- El diferencial de precios se encuentra entre el 0,98 y el 1,03.

Concluimos que, si la muestra analizada es representativa de la población a analizar, el modelo obtendrá resultados con una fiabilidad suficiente.

PARTE IV: CONCLUSIONES

4.1. CONCLUSIONES GENERALES

4.1.1. APLICABILIDAD

- Se recomienda la publicación de unas notas de guía de buenas prácticas, que no constituyan regulación, pero sí den pautas de cómo se deben realizar los modelos de valoración automatizada.

De este modo es posible trasladar al documento criterios mucho más específicos sin que por ello quede limitada la actuación del valorador, puesto que justificadamente podrá omitir aquellos criterios que no se consideren oportunos en casos particulares. Esta es también la tendencia en países con un uso más avanzado de modelos AVM.

- El acceso a la información de transacciones en el mercado inmobiliario es muy importante en el desarrollo de los modelos de valoración automatizada. Si la información de partida es abundante y representativa del mercado, la fiabilidad de las valoraciones estadísticas puede ser muy alta.

Es por tanto vital que la información de bases de datos públicas sea accesible para que se incremente el uso y la fiabilidad de las valoraciones por métodos estadísticos. Esto permitiría aumentar de forma muy significativa la fiabilidad de las valoraciones automatizadas al contar con información precisa de mercado basada en transacciones reales.

Otra posible solución es que las asociaciones profesionales de valoradores compartan entre sí sus bases de datos con el fin de disponer de un volumen mayor de datos fiables, como son las tasaciones o información de transacciones reales.

- Se recomienda que el uso de los modelos de valoración automatizada se limite a inmuebles residenciales de vivienda colectiva y sus anexos vinculados, y que

su utilización se limite a actualizaciones de valor en relación a las finalidades hipotecarias.

Si bien la circular 3/2008 del Banco de España establece que se permite el uso de estas metodologías para valorar todo tipo de viviendas y locales comerciales, sería más adecuado promover la utilización de métodos estadísticos principalmente en viviendas plurifamiliares que tienden más a cierta estandarización que en el caso de las viviendas plurifamiliares.

En el caso de los locales comerciales, la importancia de la ubicación específica hacen que no sea aconsejable la utilización de métodos estadísticos para su valoración. Si bien en otros países se utilizan estas metodologías para valorar inmuebles comerciales e incluso suelos, esto ocurre principalmente en mercados mucho más transparentes con gran disponibilidad de información de mercado fiable y en los que el uso de estas herramientas está mucho más extendido, por lo que los desarrolladores suelen tener más experiencia en cuanto al modelado y funcionamiento de los modelos, pudiendo entonces aplicar la metodología a otras tipologías de inmuebles.

Actualmente debido a las especificaciones de la orden ECO 805/2003 no es posible utilizar métodos estadísticos en primeras valoraciones para finalidad hipotecaria. No obstante, si en el futuro se produce un cambio normativo, sería posible proponer el uso de valoraciones automatizadas también para primeras valoraciones de viviendas con *Loan to Values* bajos y de escaso importe.

- Una de las cuestiones más importantes es la necesidad de incluir en el informe de valoración una indicación del rango de confianza de la valoración, siendo este aspecto algo en lo que todos los estándares coinciden.

El problema reside en la actual falta de estandarización en cuanto a la forma de calcular los rangos de confianza de las valoraciones, por lo cual este resulta el principal aspecto a mejorar en el corto plazo a nivel internacional.

Uno de los estándares con mayor influencia es el elaborado por la International Association of Assessing Officers, por lo que sería recomendable adoptar estos estándares como marco para el cálculo de los rangos de confianza, fomentado de este modo una estandarización que es posible instaurar a nivel internacional.

4.1.2. IDONEIDAD DE METODOLOGÍAS

- Se consideran adecuados los modelos de regresión, las redes neuronales artificiales y otros modelos de inteligencia artificial, así como los clasificadores basados en distancias (K-vecinos). También se consideran adecuados los modelos híbridos derivados de los anteriores.
- Los modelos de regresión se consideran adecuados únicamente para mercados sencillos y en los que se disponga de información de partida muy fiable.

Estos modelos pueden dar muy buenos resultados si se dispone de información pública de transacciones reales y se realiza una adecuada zonificación que permita segregar los comparables en grupos de zonas homogéneas.

- Los modelos de redes neuronales se consideran adecuados para desarrolladores avanzados, que sean capaces de explicar y calibrar adecuadamente el modelo pese a desconocer la formulación exacta del algoritmo de valoración.

El principal problema que presentan estos modelos es la denominada “caja negra”, es decir, el desarrollador desconoce la fórmula que se ha utilizado para realizar la valoración, por lo que la defensa de los valores ante terceros puede ser compleja, especialmente debido a la poca transparencia del mercado inmobiliario español y a la falta de conocimiento avanzado de estas metodologías de valoración por parte de muchos usuarios finales.

- Los clasificadores K-vecinos se consideran los modelos más completos, tanto por su fiabilidad como por su sencillez de uso y calibración.

Estos modelos simulan el proceso de valoración por comparación de los tasadores, por lo que resultan fácilmente entendibles a la vez que carecen de formulaciones complejas.

La mayor fortaleza de este método sobre otros análogos se produce en la fase de selección de comparables, en la cual el modelo mide las similitudes con sus vecinos más cercanos para escoger los mejores comparables. Esto permite que se pueda adaptar a mercados complejos e incluso combinar fácilmente con otros algoritmos de valoración que entrarían en la fase de asignación de pesos.

Debido a su capacidad para adaptarse a mercados complejos mediante la combinación de varios algoritmos, se espera que en los próximos años prolifere el uso de modelos híbridos.

- No se consideran adecuados los modelos basados exclusivamente en índices ya que pueden arrastrar errores de la valoración inicial y además no toman en cuenta las características del inmueble.

La actualización del valor por índices se realiza con base en la media de mercado, la cual puede no coincidir con la evolución del precio para un inmueble en particular, especialmente si el periodo de actualización es muy amplio.

4.1.3. SEGMENTACIÓN

- Es necesario realizar una segmentación los datos con base en su uso, distinguiendo entre residencial plurifamiliar y garajes.
- También es necesario segmentar con base en la zonificación, de forma que la selección de comparables esté restringida al mercado homogéneo o *cluster* en el que se ubique el inmueble a valorar.
- Se recomienda que la zonificación sea granular, con *clusters* más pequeños en núcleos urbanos densos y más grandes en zonas rurales.
- Para cada uso y *cluster* podrá existir un modelo diferente o bien un modelo general por uso con ajustes para cada cluster, permitiendo así un ajuste lo más

adecuado posible entre el algoritmo de valoración y cada mercado diferenciado.

4.1.4. SELECCIÓN DE VARIABLES

- Si bien en muchos estándares de valoración se citan variables que por norma general afectan al valor del inmueble, las variables a utilizar en cada valoración suelen ser opcionales, de modo que es el propio valorador el que puede decidir en cada caso qué variables pueden tener un mayor peso en el valor final.

No obstante, debido a las particularidades de nuestro mercado y a las solicitudes en muchos casos de valorar inmuebles de los cuales no se dispone de apenas información, considero necesario que se establezcan unos mínimos, de forma que si no se dispone de la información mínima no sea posible realizar la valoración por métodos estadísticos.

Las dos variables con mayor repercusión en el precio son, aparte del uso del inmueble, su superficie y ubicación. Estos son por tanto datos indispensables de partida para cualquier por métodos estadísticos y es necesario disponer de información fiable al respecto para poder garantizar la confianza en la valoración.

- Por otra parte, debido a la falta de información pública en muchos casos será necesario recurrir a otras fuentes de información. Debido a que la mayoría de modelos de valoración automatizada requieren un gran volumen de datos para garantizar su fiabilidad, si no se dispone de suficiente información de transacciones reales o tasaciones, puede que sea necesario recurrir a información de oferta.

En este sentido se debe tener en cuenta que la inclusión de demasiadas variables suelo ser perjudicial para la fiabilidad de los modelos, especialmente si se trata de variables no informadas en muchos de los comparables.

Por este motivo es recomendable reducir en la medida de lo posible el número de variables a aquellas de las que depende principalmente el valor. En caso de disponer de abundante información de todos los comparables, esta reducción en el número de variables se puede realizar mediante el uso de variables transformadas que engloban varias características del inmueble.

- Teniendo lo anterior en cuenta, además de los atributos excluyentes citados en el apartado de segmentación, se recomiendan las siguientes variables mínimas para cada uso:
 - Vivienda plurifamiliar:
 - Año de edificación
 - Tipología: ático, planta baja, dúplex, piso, etc.
 - Calidades constructivas
 - Superficie de la vivienda
 - Existencia de anexos: garajes o trasteros vinculados.
 - Garaje:
 - Superficie útil

4.1.5. PARÁMETROS MÍNIMOS DE DEPURACIÓN DE DATOS

La fase de depuración y normalización de datos es esencial para maximizar la fiabilidad de los datos de partida, adaptar los datos para la entrada en el modelo, eliminar comparables que distorsionan los valores, completar datos si es necesario y realizar los ajustes precisos en los precios de los inmuebles comparables.

Se deberá como mínimo realizar los siguientes procedimientos de depuración:

- Autocompletado de datos en caso necesario: Se deberá valorar si se trata de una variable que en la mayoría de los casos está informada o no. Si el dato falta en un porcentaje bajo de muestras y es una variable secundaria (no se deben autocompletar datos de ubicación ni superficie), se podrá autocompletar en base a los datos medios de inmuebles similares. Si es un dato no informado en la mayoría de las muestras hay que plantearse eliminar la variable.
- Asignación de zonificación correspondiente: cada muestra de la base de datos deberá tener asignada su correspondiente zona homogénea resultante del proceso de zonificación.
- Normalización, discretización y reducción de número de variables: Los modelos suelen mejorar sus resultados si se reduce el número de variables, bien sea eliminando aquellas que no afectan significativamente al precio de los inmuebles o bien agrupando varias variables relacionadas entre sí en una

variable transformada. En el caso de variables cualitativas y ciertas variables numéricas puede ser útil agruparlas por escalones de valor (discretización).

- Actualización de datos históricos a fecha actual: Se debe ajustar el precio de los inmuebles comparables en función de la evolución histórica de los precios en cada zona. Se recomienda que no se utilicen muestras con más de 1 año de antigüedad, puesto que su precio a fecha de valoración puede quedar muy distorsionado si no se ajusta a la evolución media en la zona.
- Ajuste de precios de oferta a la realidad de cada mercado: Si debido a la falta de información de transacciones reales resulta imprescindible incluir muestras de precios de oferta, se deberá siempre contrastar con precios reales para estimar un porcentaje de negociación en cada zona y ajustar así los precios de oferta para asimilarlos a los de mercado.
- Eliminación de *outliers*: Es muy importante eliminar aquellas muestras con valores atípicos que pueden producir problemas de aprendizaje en el modelo y sesgar los resultados de las valoraciones. Este es uno de los pasos más importantes, que se deberá repetir en la fase de calibración del modelo para conseguir mejorar los resultados y márgenes de confianza.
- Dar una solución adecuada a las viviendas con anexos vinculados. Se recomienda separar viviendas de garajes y valorar por separado. En ciertos cluster muy homogéneos donde no existen viviendas sin garaje se puede realizar un modelo que permita valorarlos conjuntamente. No obstante, se recomienda exclusivamente en zonas donde no haya ningún comparable de viviendas sin anexos vinculados.

4.1.6. CONSTRUCCIÓN Y EVALUACIÓN DE MODELOS

- Se recomienda seguir las instrucciones del International Association of Assessing Officers (I.A.A.O.) para el cálculo de unos rangos de confianza entendibles y reconocidos a nivel global.

Esto constituye el principal aspecto a mejorar a nivel mundial en el campo de las valoraciones automatizadas, de forma que todos los proveedores de valoraciones AVM indiquen en sus informes rangos de confianza que puedan

ser comparados con los de otras empresas. Esto simplificaría enormemente el control de calidad por parte de las entidades financieras y aportaría transparencia y confianza al sector.

- Para determinar la fiabilidad de un modelo se debe realizar un análisis de ratios sobre la pase de datos de test, siendo el ratio igual al valor calculado por el modelo AVM dividido entre el precio real del inmueble. Tomando como datos de partida los ratios calculados para cada inmueble de la base de datos de test, se analizan varios indicadores que indican la fiabilidad del modelo.

Para realizar el análisis de ratios, los principales indicadores según el International Association of Assessing Officers son los siguientes:

- Medida de centralidad. Se utiliza habitualmente la mediana
 - Medida de Dispersión, que se puede evaluar mediante el coeficiente de dispersión.
 - Diferencial de precios
- La medida de centralidad se basa en la tendencia central, siendo la mediana, la media y la media ponderada los principales parámetros para evaluar la tendencia central de los modelos.

De acuerdo al “Standard on Automated Valuation Modelos (IAAO, 2003), las medidas de centralidad deberán estar entre el 0,9 y el 1,1 para que se pueda considerar que el modelo ofrece un suficiente grado de fiabilidad.

Además de este criterio, el IAAO establece que el ratio entre valores calculados y precios para cada grupo de inmuebles debe estar dentro del $\pm 5\%$ del ratio global para la totalidad de los datos. Para realizar esta comprobación utilizamos los intervalos de confianza. Este estándar se cumple si los intervalos de confianza de el parámetro escogido de tendencia central al 95% para cada cluster quedan solapados dentro del $\pm 5\%$ del nivel global.

- Las medidas de dispersión reflejan la variabilidad horizontal de los ratios, siendo comúnmente utilizado el coeficiente de dispersión.

El coeficiente de dispersión debería estar según el IAAO entre el 5 y el 10% en zonas nuevas y más homogéneas y entre 5 y 15% en zonas más antiguas y heterogéneas. En zonas rurales se permite un COD máximo de hasta el 20%.

- El diferencial de precios nos da una indicación de la variabilidad vertical de los ratios. El valor debe estar entre 0,98 y 1,03 para que la fiabilidad del modelo sea aceptable.

Si el resultado es menor que 0,98 la valoración es progresiva, mientras que si el resultado es mayor que 1,03 la valoración es regresiva.

4.1.7. PROPUESTA DE REGULACIÓN

Como conclusión última se incluye a continuación la propuesta de regulación con carácter de notas de guía de buenas prácticas, al considerarse más adecuado al propósito permitiendo una mayor profundidad y versatilidad.

La propuesta es la conclusión final de la tesis y responde a los objetivos principales de la misma, que incluyen:

- Adaptar las metodologías de valoración automatizada a las necesidades del Banco de España para actualización de valor de inmuebles hipotecados.
- Redactar un borrador o aproximación normativa a lo que podría considerarse una regulación adecuada en la materia, con base en la normativa existente en otros países y a la regulación internacional.
- Establecer un procedimiento que pueda ser utilizado por el propio Banco de España como control de calidad interno para evitar sobrevaloración de activos inmobiliarios.

4.2. PROPUESTA DE NOTAS DE GUÍA DE BUENAS PRÁCTICAS

TÍTULO I. DISPOSICIONES GENERALES

Capítulo I. Ámbito, principios y definiciones

Artículo 1. Objeto

La presente propuesta tiene como finalidad establecer unas notas de guía de buenas prácticas para la realización de valoraciones estadísticas cuya finalidad sea la revisión de valores de inmuebles hipotecados de conformidad con lo dispuesto en el artículo 4.2.b de la circular 3/2008 del Banco de España.

Artículo 2. Tipologías de aplicación

Las valoraciones por métodos estadísticos para las finalidades expuestas en el artículo 1 se podrán realizar exclusivamente para viviendas colectivas, preferiblemente de carácter urbano, y sus anexos vinculados tales como garajes.

Artículo 3. Requisitos de aplicación

Se excluyen del ámbito de aplicación las viviendas sujetas a algún tipo de protección que limite su precio de venta así como las viviendas sujetas a contratos de arrendamiento.

Artículo 4. Inmuebles singulares

Tampoco será de aplicación en el caso de inmuebles singulares o no susceptibles de producción repetida.

Se identificarán los *outliers* siguiendo el mismo procedimiento establecido por la sociedad de tasación para identificar *outliers*, esto es, comprobando que el inmueble a valorar está dentro de los rangos establecidos para cada una de las variables en el *cluster* en el que se ubique el inmueble.

Artículo 5. Mercados no representativos

Para que un inmueble pueda ser valorado por métodos estadísticos es imprescindible que exista un mercado de inmuebles representativo de inmuebles similares. Por lo tanto, no se podrá realizar una valoración por métodos estadísticos si no se dispone de al menos 6 inmuebles comparables en la base de datos que estén ubicados en el mismo *cluster* que el inmueble objeto de valoración.

Artículo 6. Definiciones

AVM: Modelo de valoración automatizada o estadística.

Cluster: Subconjuntos de zonas homogéneas obtenidos como resultado del proceso de zonificación.

Comparable: Son inmuebles que se consideran similares al inmueble objeto de valoración o adecuados para aplicar la homogeneización, teniendo en cuenta su localización, uso, tipología, superficie, antigüedad, estado de conservación, u otra característica física relevante a dicho fin. (Orden ECO 805/2003)

Inmueble singular: Inmueble que, debido a sus características singulares no es susceptible de producción repetida. No es posible valorar estos inmuebles mediante modelos de valoración automatizada.

Outlier: Valores extremos en una distribución.

Segmentación: Proceso mediante el cual se agrupan inmuebles comparables para su modelado y análisis de forma que conjunto de muestras queda subdividido en dos o más subconjuntos homogéneos de acuerdo a un criterio establecido, como por ejemplo su uso, características o ubicación.

Variable: característica del inmueble que interviene en el cálculo de su valor. Se clasifican comúnmente en variables de barrio, emplazamiento y estructura.

Variables de barrio: Variables relacionadas con el barrio en el que se ubica el inmueble, entendiendo como tal el área geográfica caracterizada por un mercado homogéneo de inmuebles comparables.

Variables de distancia: Variables relacionadas con la distancia del inmueble a puntos de especial interés o a hitos urbanos, tales como centro de la ciudad, distrito financiero, transporte público, zonas verdes, etc.

Variable de emplazamiento: Variables relacionadas con la ubicación específica de un inmueble, tales como vistas al mar o ubicación en una calle específica dentro del barrio.

Variable de estructura: Rasgos estructurales y constructivos inherentes al inmueble en sí.

Conjunto de realizaciones que se obtienen al variar la localización en el interior de la ciudad. Se caracteriza porque en su distribución espacial presenta una componente aleatoria y una componente estructural de autocorrelación espacial.

Zonificación: Segmentación de los datos realizada utilizando como criterio la ubicación de los inmuebles.

Artículo 7. Informe y su contenido mínimo

El informe de valoración por métodos estadísticos deberá incluir como mínimo la información de partida del inmueble a valorar, su referencia catastral, los comparables dentro del mismo *cluster* o zona homogénea, así como sus características, el modelo de valoración utilizado, un rango de confianza de la valoración y un comentario de la evolución del mercado local en el que se ubique el inmueble.

Capítulo II. Comprobaciones y documentación

Artículo 8. Verificación de información con catastro

Se deberá verificar la información de partida del inmueble, incluyendo la ubicación exacta, superficie del inmueble, existencia de anexos vinculados y año de edificación, contrastando la misma con la información catastral.

Artículo 9. Información mínima necesaria para valorar por métodos estadísticos

La información mínima necesaria para valorar por métodos estadísticos será la ubicación exacta del inmueble, su referencia catastral, superficie adoptada en el cálculo, uso y tipología del inmueble, existencia de anexos vinculados y confirmación de que el inmueble no está arrendado ni sujeto a ningún tipo de protección que limite su precio de venta.

Es requisito indispensable conocer la ubicación exacta del inmueble, pudiendo completarse ciertos datos de partida que no sean aportados por el solicitante a partir de la información catastral. En aquellos casos en los que no sea posible identificar inequívocamente el inmueble y su referencia catastral, será necesario realizar una inspección del inmueble o revisar la documentación de la tasación anterior.

Capítulo III. Advertencias

Artículo 10. Características del inmueble

Se realizará una advertencia en el informe de valoración cuando existan discrepancias relevantes entre la información de partida recibida del inmueble a valorar y su descripción catastral. En este caso se recomendará realizar una inspección al inmueble o revisar el informe de valoración anterior con el fin de determinar con mayor certeza las características del inmueble a valorar.

Artículo 11. Rangos de confianza mínimos

Se realizará una advertencia cuando los rangos de confianza obtenidos para la valoración no superen los estándares establecidos por el IAAO, los cuales se resumen en el título III del presente texto. Asimismo, se recomendará que un valorador revise el resultado obtenido por métodos estadísticos, ajustando en su caso el valor para mejorar su precisión o bien que se realice una valoración presencial del inmueble.

Artículo 12. Identificación de *outliers*

Si las comprobaciones indicadas en el artículo 4 sugiriesen que gran parte de las características del inmueble están en el límite que permite que se valore por métodos estadísticos, se deberá realizar una advertencia de dicha circunstancia en el informe de valoración.

Artículo 13. Mercados poco representativos

Si el *cluster* en el que se ubica el inmueble a valorar encontramos menos de 10 comparables para el inmueble a valorar, se deberá realizar una advertencia al respecto, indicando que el mercado es poco representativo y que por tanto, la fiabilidad de la valoración puede quedar comprometida.

TÍTULO II. CÁLCULO DEL VALOR

Capítulo I. Requisitos de partida

Artículo 14. Modelo de datos

Definición del modelo de datos único integrando toda la información. Se deberá establecer el número de modelos analíticos a desarrollar, siendo recomendable establecer al menos dos modelos: uno para vivienda colectiva y otro para garajes.

Artículo 15. Rangos para inmuebles singulares

Se deberán establecer rangos dentro de cada variable para determinar qué tipo de inmuebles se consideran singulares y, por tanto, no susceptibles de ser valorados por métodos estadísticos.

Artículo 16. Rangos para inmuebles comparables

Se establecerán horquillas o rangos relativas a cada variable para determinar qué inmuebles consideramos similares o comparables respecto al inmueble a valorar. Sólo se podrán valorar aquellos inmuebles para los que se encuentren al menos el número de comparables establecido (al menos 6) y cuyas características estén dentro de las horquillas establecidas para cada variable.

Artículo 17. Nivel geográfico

Se deberá definir el nivel geográfico más bajo disponible (para localizaciones urbanas), así como el más alto (para localizaciones rurales) para la determinación de *clusters*. El nivel más bajo recomendado corresponde a la sección censal.

Artículo 18. Profundidad histórica

Se deberá definir y garantizar la profundidad histórica de los datos para la realización de los modelos. Se recomienda que la profundidad no sea superior a un año en zonas con abundante información de comparables. En general se recomienda que la profundidad histórica sea lo más corta posible para que el ajuste de precio a la fecha de valoración sea mínimo.

Artículo 19. Actualizaciones

Se deberá definir la frecuencia o periodicidad de las actualizaciones, recomendándose que sean al menos cada 3 meses.

Asimismo, se deberán definir los procesos de carga y actualización periódica de la información

Capítulo II. Variables

Artículo 20. Variables mínimas

Las variables mínimas para llevar a cabo una valoración por métodos estadísticos para la finalidad contemplada en el artículo 1 son las siguientes:

- Identificación exacta del inmueble
- Uso del inmueble a valorar (vivienda colectiva o garaje)
- Superficie adoptada

Además en el caso de viviendas, se consideran variables mínimas también las siguientes:

- Tipología del inmueble a valorar (piso, ático, dúplex, bajo con jardín, etc.)
- Calidades constructivas
- Año de edificación
- Existencia de anexos vinculados

Artículo 21. Identificación del inmueble a valorar

Será necesario conocer la identificación exacta del inmueble a valorar. También será necesario conocer su referencia catastral con el fin de establecer una identificación inequívoca y de realizar las comprobaciones indicadas en el artículo 8.

Artículo 22. Variables de emplazamiento

Se recomienda incluir variables de emplazamiento en la valoración, escogiendo variables específicas a cada tipo de ubicación (núcleo urbano, zona costera, zona rural, etc.) para cada *cluster*. La elección de estas variables y de los *cluster* a los que se aplica cada una será decidido por cada sociedad de tasación. Se recomienda incluir variables de distancia a puntos de interés, debiéndose definir adecuadamente los puntos de interés para cada *cluster* según el tipo de ubicación (por ejemplo, distancia al distrito financiero para núcleo urbano, distancia a la playa para zona costera, etc.)

Capítulo III. Segmentación

Artículo 23. Segmentación por usos

Se realizará al menos una segmentación por usos, estableciendo un modelo para viviendas colectivas y otro para garajes. Cada sociedad de tasación deberá decidir si realiza o no segmentaciones adicionales por tipología de inmueble con el fin de separar en diferentes modelos los pisos de otras tipologías de vivienda colectiva.

Artículo 24. Zonificación o segmentación por localización

Se realizará obligatoriamente una segmentación con base en la ubicación de los inmuebles. Se recomienda que la zonificación se realice teniendo en cuenta el número de inmuebles comparables en cada zona, de forma que los *cluster* sean más pequeños en núcleos urbanos densos y abarquen mayores superficies en zonas rurales.

Si no se desea realizar un análisis específico de ubicaciones para establecer una zonificación, será aceptable recurrir a zonificaciones preestablecidas, tales como barrios, distritos, regiones u otras demarcaciones urbanísticas o administrativas.

Artículo 25. Variables de ubicación

En el caso de realizar una zonificación mediante el análisis de zonas homogéneas, se tomarán variables de ubicación de tipo socioeconómico, ambiental, geoespacial, de estructura urbana u otras relevantes para la delimitación de los clusters de forma que cada cluster esté formado por zonas homogéneas en relación al mercado inmobiliario residencial.

Artículo 26. Variable regionalizada

En el caso de realizar un único modelo que se ajuste a varios *clusters*, la variable de ubicación se podrá asignar a los inmuebles en forma de variable regionalizada, es decir, una variable espacial de la cual dependen los valores de los inmuebles y que puede agrupar diversas características relacionadas con la ubicación en una única variable, tales como datos sociodemográficos, de calidad del área local, etc.

Capítulo IV. Parámetros mínimos de depuración y normalización

Artículo 27. Ajuste de precios de oferta

En el caso de incorporar comparables de oferta en las bases de datos, habrá que realizar un ajuste de precios por negociación. Para ello se recomienda que se utilicen

como referencia datos de transacciones reales con el fin de detectar el intervalo medio de negociación en cada mercado local.

Artículo 28. Actualización de precios a fecha de valoración

Las bases de datos tendrán una profundidad histórica. Para aquellos inmuebles introducidos en la bases de datos de comparables con fecha anterior a la valoración, se realizará una actualización de precios para ajustarlos a la fecha de valoración utilizando para ello los datos de evolución de precios en cada zona.

Artículo 29. Eliminación de *outliers*

Se deberán desechar de la base de datos todos aquellos inmuebles que se consideren *outliers*, bien porque el precio de venta difiera excesivamente del de inmuebles similares o bien porque se trata de inmuebles singulares no susceptibles de producción repetida y, por tanto, no adecuados para la utilización del método de comparación.

Artículo 30. Autocompletado de datos

Se deberán autocompletar aquellos datos secundarios de los inmuebles comparables que no estén informados con base en los datos medios de testigos similares. En ningún caso se podrá completar por estos medios la superficie de la vivienda o su ubicación.

Se podrá utilizar la información de catastro para autocompletado de datos de antigüedad de la edificación y comprobación de superficies.

Artículo 31. Variables excluyentes según segmentación

Como parte del proceso de normalización, será necesario asignar a cada inmueble de la base a datos su uso y ubicación correspondientes a base a la segmentación de datos que la sociedad de tasación haya realizado. Estos atributos funcionarán como variables excluyentes, de forma que para cada inmueble a valorar sólo se podrán seleccionar comparables del mismo uso y ubicados en el mismo *cluster*.

Capítulo V. Construcción de modelos

Artículo 32. Métodos de valoración estadística

Los métodos de valoración estadística más utilizados incluyen modelos basados en análisis de regresión, modelos basados en inteligencia artificial (tales como redes neuronales, algoritmos genéticos o lógica difusa) y clasificadores basados en distancia, como los algoritmos K-vecinos y similares. También es posible utilizar modelo híbridos que utilizan una combinación entre varios de los algoritmos anteriores. En general, es viable la utilización de cualquier modelo de valoración en el que, mediante la utilización de unos procesos y metodología adecuada, se calcula el valor del inmueble por comparación con inmuebles similares.

Artículo 33. Métodos no válidos

No se podrán realizar valoraciones por actualización mediante índices para las finalidades indicadas en el artículo 1 de esta propuesta.

Artículo 34. Metodología

Cada sociedad de tasación deberá establecer una metodología para la construcción de sus modelos de valoración automatizada, estableciendo las bases relativas a requisitos de partida, variables a tener en cuenta en la valoración, procesos de segmentación, depuración y normalización y algoritmos utilizados.

La metodología deberá someterse a revisión constante, al igual que el modelo, y deberá recoger todos los cambios o mejoras que se realicen sobre el modelo de valoración.

Artículo 35. Fuentes de información

Con el fin de asegurar la integridad y consistencia del modelo con la realidad del mercado, es necesario que al menos un 10% de los inmuebles de la base de datos procedan de transacciones reales o bien de tasaciones tradicionales realizadas por la propia sociedad de tasación.

Este 10% de inmuebles comparables debe ser representativo de las zonas donde se pretenden realizar valoraciones por métodos estadísticos, por lo que no será posible realizar valoraciones de inmuebles ubicados en zonas geográficas donde no se disponga de un volumen suficiente de información de transacciones reales o de tasaciones anteriores.

Artículo 36. Contraste de resultados

Se deberán testear al menos dos algoritmos diferentes de valoración y seleccionar aquel que ofrezca un mayor grado de confianza en cada caso.

Artículo 37. Partición de los datos

Se realizará una partición de los datos en entrenamiento, validación y test, debiendo corresponder a la muestra de test al menos un 10% de las muestras.

Es posible utilizar métodos de validación cruzada, especialmente cuando el número de muestras de la base de datos es escaso.

TÍTULO III. EVALUACIÓN Y PRESENTACIÓN DE RESULTADOS

Capítulo I. Rangos de confianza en la valoración

Artículo 38. Ratios

Es necesario realizar un análisis de ratios sobre los resultados obtenidos para la base de datos de test. El ratio para cada inmueble se calcula dividiendo el valor calculado por el modelo AVM entre el precio real del inmueble.

Artículo 39. Medidas de centralidad

Se tomará como mínimo la mediana de los ratios, si bien es recomendable realizar el análisis también sobre la media y la media ponderada de los ratios.

Las medidas de centralidad analizadas deberán estar entre el 0,9 y el 1,1 para que la fiabilidad del modelo se considere aceptable.

Además, el ratio para cada *cluster* debe estar dentro del $\pm 5\%$ del ratio global para la totalidad de *clusters*. Este estándar se cumple si los intervalos de confianza de el parámetro escogido de tendencia central al 95% para cada *cluster* quedan solapados dentro del $\pm 5\%$ del nivel global.

Artículo 40. Medidas de dispersión

Se deberá calcular el coeficiente de dispersión de los ratios.

El coeficiente de dispersión deberá ser igual o menos al 10% en zonas nuevas y más homogéneas, e igual o menos al 15% en zonas más antiguas y heterogéneas. En zonas rurales se permite un COD máximo del 20%.

Artículo 41. Diferencial de precios

El diferencial de precios se calcula dividiendo la media entre la media ponderada de los ratios.

El diferencial de precios deberá estar comprendido entre 0,98 y 1,03 para que la fiabilidad del modelo se considere aceptable.

Artículo 42. Resultados de mercado

Se deberán contrastar los resultados del modelo con comparables de mercado de la zona procedentes de tasaciones o transacciones reales recientes en la misma zona, incluyendo estas referencias de mercado en el informe de valoración.

Capítulo II. Criterios para la elaboración de informes

Artículo 43. Contenido mínimo del informe

El informe de valoración contará como mínimo con la siguiente información:

- Ubicación del inmueble y referencia catastral
- Uso y *cluster* del inmueble a valorar.
- Información de partida aportada por el solicitante, especialmente información de aquellas variables que han intervenido en el cálculo del valor.
- Comparables ubicados dentro del mismo *cluster*, así como su procedencia (fuente de datos), y ajustes realizados a cada comparable, tales como actualización de valor a fecha de valoración, ajustes por negociaciones, autocompletado de datos, etc.
- Representación gráfica sobre el plano de los límites del cluster y la ubicación del inmueble a valorar e inmuebles comparables.
- Procedimiento de valoración utilizado especificando algoritmo de valoración.
- Resultado de la valoración expresado en euros y en euros por metro cuadrado.
- Rango de confianza de la valoración.
- Valoración anterior y fecha de dicha valoración si se dispone de esta información.
- Otras referencias de mercado de transacciones reales o tasaciones recientes en la zona que no hayan intervenido en el cálculo.
- Comentario de mercado analizando la evolución del mercado en la zona donde se ubica el inmueble.

4.3. FUTURAS LÍNEAS DE INVESTIGACIÓN

Entre las futuras líneas de investigación cabe destacar las siguientes:

- El análisis de las metodologías y prácticas profesionales en zonificación de áreas homogéneas en el desarrollo de modelos de valoración automatizada.
- Prácticas habituales y criterios normativos en el análisis de rangos de confianza en los modelos de valoración automatizada. Hacia una estandarización a nivel internacional.

4.4. BIBLIOGRAFÍA

Andersson, H., et al. (2009) Property prices and exposure to multiple noise sources: hedonic regression with road and railway noise. *Environmental and resource Economics*. 45 (1), 73 – 89.

Anselin, L. and Lozano-Gracia, N. (2008) Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics* 34, 5 - 34.

Armengot Paradinas, J. (2013) Depreciación diferencial y estructura urbana: análisis tipológico de la edificación residencial en la almendra central de Madrid. Tesis doctoral. Madrid: Universidad Politécnica de Madrid.

Asociación Española de Análisis del Valor (2013) Estándares profesionales de los miembros de la AEV dirigidos a la aplicación rigurosa y prudente de los métodos estadísticos de valoración masiva de inmuebles (AVM), AEV, Madrid

Banco de España (2008) Circular 3/2008, de 22 de mayo, del Banco de España, a entidades de crédito, sobre determinación y control de los recursos propios mínimos (BOE de 10 de junio de 2008)

Banco de España (2012) Proyecto de Circular del Banco de España por la que se modifican las circulares 7/2010, 3/1998, 4/2004 y 3/2008 (diciembre de 2012)

Bitter, C. et al. (2007) Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*. 9 (1), 7 – 27.

Bourassa, S. et al. (2008) Predicting House Prices with Spatial Dependence: Impacts of Alternative Submarket Definitions. Swiss Finance Institute. Research Paper Series. 8 (1).

Bourassa, S. et al. (2007) Spatial dependence, housing submarkets and house price prediction. *The Journal of Real Estate Finance and Economics*. 35, 143 – 160.

Brunsdon, C., Fotheringham, S. and Charlton, M. (1998) Geographically weighted regression: a method for exploring spatial non-stationarity. *The Statistician*, 47 (3), 431 - 443

Caldwell, B. J. (1982) *Beyond Positivism: Economic Methodology in the Twentieth Century*. London. Allen & Unwin.

Can, A. (1992) Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22, 453 - 474.

Cano Guervós, R et al. (1999) Metodología para la zonificación de una ciudad. *Estudios de economía aplicada*, 13, 23-49.

Cartwright, N. (1999) *The Dappled World*. Cambridge: Cambridge University Press.

Chica-Olmo, J. (1995) Spatial estimation of housing prices and locational rents. *Urban Studies*, 32 (8), 1331 – 1344

Chin, T.L. and Chau, K.W. (2003) A critical review of literature on the hedonic price model. *International Journal for Housing Science and Its Applications*, 27 (2), 145 - 165.

Cohen, J. and Coughlin, C. (2008) Spatial hedonic models of airport noise, proximity and housing prices. *Journal of Regional Science*. 48, 859 – 878.

Council of the European Union. Capital Requirements Directive. Document 12890/05 of 18th October 2005. Council of the European Union (2005)

Davis, G.C. (2000) A Semantic Conception of Haavelmo's Structure of Econometrics. *Economics and Philosophy*. 16 (2), 205 - 28.

Downward, P. and Mearman, A. (2005) Reorienting Economics Through Triangulation of Methods. *Post-autistic economics review*, issue. 31 (4).

Dubin R. (1992) Spatial autocorrelation and neighbourhood quality. *Regional Science and Urban Economics*. 22 (3), 433 - 452

Evans, A., Asabere, W. and Paul, K. (1994) The price determinants of foreclosed urban land: A discussion. *Urban Studies*. 31 (10), 1777 - 1782.

Fernandes, V. et al. (2012) Assessing the effects of the São Paulo's metropolitan beltway on residential land prices. *Universidade Presbiteriana Mackenzie*.

[online]. <http://www.scielo.br/scielo.php?pid=S2238-10312013000200020&script=sci_arttext>

Fernández Pirla, S. (1982) Valoraciones administrativas y de mercado del suelo y las construcciones. Madrid, Consejo Superior de los Colegios de Arquitectos.

Fernández Pirla, S., García Erviti, F. et al. (1986) Curso sobre formación, conservación y revisión del catastro de urbana. Madrid, Colegio Oficial de Arquitectos de Madrid.

Ferrando Corell, J. V. (2004) Valoración de inmuebles de naturaleza urbana. Valencia, Universidad Politécnica de Valencia.

Ferrando Corell, J. V. y Marques Romero, M. A. (1992) Valoraciones inmobiliarias. Valencia, Universidad Politécnica de Valencia.

Ferreira Vaz, A.J. (2013) La dimensión de la subjetividad en la formación del valor inmobiliario : aplicación del método de análisis de ecuaciones estructurales al mercado residencial de Lisboa. Tesis doctoral. Madrid: Universidad Politécnica de Madrid.

Fisher, I. (1933) Statistics in the Service of Economics. Journal of the American Statistical Association. 28 (181), 1 - 13.

Frisch, R. (1933) Editor's Note. Econometrica. 1 (1), 1 - 4.

Gallego Mora-Esperanza, J. (2004) La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid. CT/Catastro, 50, 51 – 67.

Gallego Mora-Esperanza, J. (2008) Modelos de valoración automatizada. CT/Catastro, 62, 7 – 26.

Gallego Mora-Esperanza, J. (2010) Valoración Automatizada. [online]. <http://www.catastro-latino.org/documentos/sanluis2010/Valoracion_JGallego.pdf>

García Erviti, F. (1990). La Distribución Espacial de los Valores del Suelo Urbano en Madrid- Tesis Doctoral. Madrid: Universidad Politécnica de Madrid.

García Erviti, F. (2001) Compendio de arquitectura legal. Madrid. Marea/Celeste.

García Erviti, F. (2012) El valor estable de la vivienda a efecto hipotecario no existe. El País. [online]. <http://oa.upm.es/14449/1/FGE_El_Pa%C3%ADs_Pro-piedades.pdf>

Garrido-Lestache Rodríguez, E. (2004). Determinación de la variación del valor de mercado de la vivienda durante su vida útil - Tesis Doctoral. Barcelona: Universitat Politècnica de Catalunya.

Garrido-Lestache Rodríguez, E. y Humero Martín, A. E. (2009) Tratado técnico-jurídico de la edificación y del urbanismo. Tomo V. Valoraciones inmobiliarias. Madrid. Aranzadi – Thomson Reuters.

González Morcillo, C. (2011) Lógica difusa, Una introducción práctica. Técnicas de Softcomputing. [online]. <http://www.esi.uclm.es/www/cglez/downloads/docencia/2011_Softcomputing/LogicaDifusa.pdf>

Griffith, D. (1988) Advanced spatial statistics: special topics in the exploration of quantitative data series. Kluwer, Dordrecht.

Haavelmo, T. (1943) The Statistical Implications of a System of Simultaneous Equations. *Econometrica*. 11 (1), 1 - 12.

Heckman, J.J. (2000) Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective. *Quarterly Journal of Economics*. 115 (1), 45 - 97.

Hoover, K. (2006) Chapter 2 The Methodology of Econometrics in *Palgrave Handbook of Econometrics Volume 1: Econometric Theory*, Edited by Terence C. Millsand and Kerry Patterson. Basingstoke. Palgrave Macmillan.

International Association of Assessing Officers (2003) Standard of Automated Valuation Models (AVMs), I.A.A.O, Chicago.

International Association of Assessing Officers (2013) Standard on Mass Appraisal of Real Property, I.A.A.O, Chicago.

International Association of Assessing Officers (2013) Standard on Ratio Studies, I.A.A.O, Chicago.

Kim, C., et al. (2003) Measuring the benefits of air quality improvement: a spatial hedonic approach. *Journal of Environmental Economics and Management*. 45, 24 – 39.

Koza, J. R. (1992) Genetic programming. On the programming of computers by means of natural selection. Cambridge, The MIT Press.

Lawson, T. (1997) Economics and reality. London. Routledge

Lenk, M. at al. (1997) High tech valuation: should the artificial neural networks bypass the human valuer? *Journal of Property valuation & Investment*. 15 (1), 8 – 26.

Lesage J. and Pace R. (2004) Models for spatially dependent missing data. *The Journal of Real Estate Finance and Economics*. 29 (2), 233 – 254.

Mc Cluskey, W. and Anand, S. (1999) The application of intelligent hybrid techniques for mass appraisal of residential properties. *Journal of Property Investment & Finance*, Vol. 17, No. 3

Meen, G. (1996) Spatial aggregation, spatial dependence and predictability in the UK housing market. *Housing Studies*. 11, 345 – 372.

Microsoft Office (2010). Linest Excel. [online]. <<http://office.microsoft.com/en-us/excel-help/linest-HP005209155.aspx>>

Montgomery, D. et al. (2001) Introduction to linear regression analysis. 3rd Ed. New York, Wiley.

Páez, A. (2009) Recent research in spatial real estate hedonic analysis. *Journal of Geographical Systems*. 11 (4), 311 - 316.

Ramírez Pacheco, G. M. (2012) Desarrollo de un modelo de formación del valor inmobiliario en áreas de crecimiento suburbano: aplicación a la comarca de la Vega Baja del Segura en el litoral alicantino. Tesis doctoral. Madrid: Universidad Politécnica de Madrid.

Richardson, H. et al. (1974) Determinants of urban house prices. *Urban Studies*. 11, 189 – 199.

RICS (2013) Automated Valuation Models AVMs 1st edition, RICS, Coventry.

Roca i Cladera, J. (1986). *Manual de Valoraciones Inmobiliarias*. Barcelona, Ariel Economía.

Small, K. and Steimetz, S. (2006) Spatial hedonics and the willingness to pay for residential amenities. *Economics working paper 05, 06 - 31*, University of California, Irvine.

Smith, D. (2009) Polycentric Cities and Sustainable Development. [online]. <<http://www.slideshare.net/DuncanSmith/polycentric-cities-and-sustainable-development>>

Suriatini, I. (2006) Spatial Autocorrelation and Real Estate Studies: A Literature Review. *Malaysian Journal of Real Estate*, 1 (1). 1 - 13

Waddell, P., et al. (1993) Residential property values in a multinodal urban area: New evidence on the implicit price of location. *Journal of Real Estate Finance and Economics*. 7, 117 – 141.

Wen, H. et al. (2005) Hedonic price analysis of urban housing: an empirical research on Hangzhou, China. *Journal of Zhejiang University Science*. 6A (8), 907 – 914.

Witten, I. and Frank, E. (2005) *Data Mining: Machine learning tools and techniques*. Amsterdam , Elsevier Morgan Kaufmann

Yoo, E. and Kyriakidis, P. (2009) Area-to-point Kriging in spatial hedonic pricing models. *Journal of Geographical Systems*. 11 (4).

PARTE V: ANEXOS