

UNIVERSITAT POLITÈCNICA DE CATALUNYA
ESCOLA TÈCNICA SUPERIOR D'ARQUITECTURA DE
BARCELONA

PROGRAMA DE DOCTORAT
DE GESTIÓ I VALORACIÓ URBANA



Departament de Construccions
Arquitectòniques I

TESI DOCTORAL

UN SISTEMA EXPERT DE TAXACIÓ PEL MÈTODE DE COMPARACIÓ
ELS MÈTODES D'ESTIMACIÓ ESPACIAL I LA SEVA APLICACIÓ EN
LA METODOLOGIA DE TAXACIÓ PER COMPARACIÓ

JOAQUIM BARRIS GARCIA

DIRECTORA/TUTORA: PILAR GARCIA ALMIRALL

BARCELONA

JULIOL 2008

Índex	Pàg.
1. Introducció.....	1
1.1. Context del desenvolupament de la tesi.....	1
1.2. Antecedents de la investigació.....	4
1.3. Objectius de la investigació.....	7
1.4. Justificació.....	9
1.4.1. Justificació empresarial.....	9
1.4.2. Justificació de la tesi.....	9
1.5. Metodologia.....	13
1.6. Esquema de treball i estructura.....	16
2. L'ECO805/2003. El marc teòric.....	23
2.1. Introducció.....	23
2.2. Apartats de l'ECO805/2003 relacionats amb el mètode de mercat.....	24
2.2.1. Identificació i localització.....	24
2.2.2. Localitat i entorn.....	25
2.2.3. Descripció i superfície del terreny.....	26
2.2.4. Descripció i superfície de l'edificació.....	27
2.2.4.1. Descripció i superfície de l'edificació – Edificis.....	27
2.2.4.2. Descripció i superfície de l'edificació – Elements d'edifici.....	28
2.2.5. Descripció urbanística.....	30
2.2.6. Règim de protecció, tinença i ocupació.....	30
2.2.7. Anàlisi de mercat.....	32
2.2.8. Dades i càlcul dels valors tècnics.....	33
2.2.9. Valors de taxació, condicionats i advertiments, observacions.....	34
2.2.10. Documentació annexa a l'informe.....	35
2.2.11. Conceptualització de la informació.....	36

	Pàg.
3. Eines necessàries per al desenvolupament de la tesi	39
3.1. Introducció.....	39
3.2. Eines del projecte de CATSA emprades en el desenvolupament de la tesi.....	41
3.3. Eines necessàries per al desenvolupament de la tesi.....	43
3.4. El remostreig.....	44
3.4.1. Introducció.....	44
3.4.2. Plantejament bàsic.....	44
3.4.3. El mètode de validació creuada.....	47
3.4.4. Mètodes per estudiar la qualitat de les estimacions.....	48
3.4.5. Mètode de les permutacions.....	49
3.5. Mètodes d'estimació i interpolació espacial.....	51
3.5.1. Introducció.....	51
3.5.2. Mètode de les mitjanes mòbils. La distància inversa.....	53
3.5.3. Mètode de la curvatura mínima (splines).....	55
3.5.4. Mètode dels radials.....	58
3.5.5. Metodologia kriging. L'ús de l'anàlisi estructural.....	59
3.5.5.1. Formalització.....	60
3.5.5.2. El variograma experimental.....	64
3.5.5.3. Mètode d'obtenció del variograma experimental.....	70
3.5.5.4. Selecció del model teòric.....	72
3.5.5.5. Isotropia i anisotropia.....	79
3.5.5.6. Modelització del variograma experimental.....	83
3.5.6. El Kriging.....	85
3.5.6.1. Kriging simple.....	87
3.5.6.2. Kriging ordinari.....	90
3.5.6.3. Kriging universal o kriging tendencial.....	93
3.5.7. Altres mètodes d'estimació espacial.....	96
3.5.7.1. Mètode dels veïnatsges.....	96
3.5.7.2. Mètode de triangulació.....	97
3.6. Mètodes d'estimació no lineals.....	99
3.6.1 Introducció.....	99

	Pàg.
3.6.2. Mètode d'obtenció del variograma teòric.....	99
3.6.3. Mètode dels mínims quadrats ordinaris i l'algoritme de Gauss-Newton.....	100
3.7. L'estimació de la tendència.....	104
4. El mètode de comparació de mercat basat en un sistema expert.....	105
4.1. Introducció.....	105
4.2. Estudi de necessitats i obligacions.....	106
4.3. Selecció dels camps participants en el mètode de comparació.....	108
4.3.1. Tipologia i subtipologia.....	109
4.3.2. Definició de l'habitatge.....	110
4.3.3. Caracterització de les parts específiques de l'edifici.....	113
4.3.4. La localització per geoposició.....	115
4.4. Establiment dels criteris de selecció de mostres comparables en base a filtres excloents.....	116
4.5. Creació d'un sistema d'homogeneïtzació de les característiques constructives de l'habitatge.....	119
4.5.1. Homogeneïtzació de la superfície.....	120
4.5.2. Homogeneïtzació de l'estat i la qualitat.....	120
4.5.3. Homogeneïtzació del nombre de banys.....	123
4.5.4. Homogeneïtzació del nombre de dormitoris i la seva relació amb la superfície.....	123
4.5.5. Homogeneïtzació de la planta i existència d'ascensor.....	123
4.5.6. Càlcul del valor homogeneïtzat.....	124
4.6. Estudi de comparabilitat i homogeneïtzació global.....	125
4.7. Creació d'una metodologia d'homogeneïtzació espacial.....	128
4.7.1. Primer procés – comparació de mètodes.....	132
4.7.2. Segon procés – procés iteratiu de recerca de solucions.....	133
4.8. Recerca de la funció de valor.....	136
5. Funcionament i resultats.....	139
5.1. Introducció.....	139
5.2. Definició de les bases.....	142
5.3. Estudi per a l'obtenció d'un mètode de selecció de la solució a partir d'una funció basada en la validació creuada.....	143
5.3.1. Objectiu de l'estudi.....	143
5.3.2. Definició del marc de treball.....	143

	Pàg.
5.3.3. Algoritme de treball.....	144
5.3.4. Proves sobre el mètode de selecció.....	146
5.4. Estudi de la millora de resultats assolida a partir de la utilització de l'algoritme iteratiu de treball plantejat.....	151
5.4.1. Objectius.....	151
5.4.2. Definició del marc de treball.....	151
5.4.3. Descripció de la mostra.....	151
5.4.4. Resultats de l'estudi.....	153
5.5. Valoracions realitzades a partir d'un objecte de característiques fixes.....	159
5.5.1. Introducció.....	159
5.5.2. Exemple: zona mercat de Sant Antoni.....	159
5.5.2.1. Estadístics bàsics sobre la mostra seleccionada.....	162
5.5.2.2 Resultats per comparació de mètodes.....	164
5.5.2.3 Resultats del procés iteratiu.....	166
5.6. Exemples d'informes i processos. L'obtenció de la funció de probabilitat empírica del valor objecte.....	169
5.6.1. Introducció.....	169
5.6.2. Exemple de valoració.....	169
5.6.3. Exemple d'informe.....	175
6. Conclusions i treballs futurs.....	179
7. Bibliografia.....	187

Índex de taules	Pàg.
Taula 2.1.: Identificació i localització.....	24
Taula 2.2.: Localització i entorn.....	25
Taula 2.3.: Descripció i superfície del terreny.....	26
Taula 2.4.: Descripció i superfície de l'edificació.....	27
Taula 2.5.: Descripció dels elements de l'edifici.....	28
Taula 2.6.: Descripció urbanística.....	30
Taula 2.7.: Règim de protecció, tinença i ocupació.....	30
Taula 2.8.: Anàlisi de mercat.....	32
Taula 2.9.: Dades i càlcul dels valors tècnics.....	33
Taula 2.10.: Valors, condicionats i advertiments, observacions, data i firmes.....	34
Taula 2.11.: Documentació annexa.....	35
Taula 3.1.: Mòduls de SAS.....	43
Taula 4.1.: Tipus de residencials plurifamiliars i percentatges de cost de construcció dels acabats.....	109
Taula 4.2.: Definició d'elements que conformen l'habitatge.....	110
Taula 4.3.: Exemple d'agrupacions de materials. Estudi de paviments.....	111
Taula 4.4.: Definició de reformes.....	112
Taula 4.5.: Elements d'edifici i reformes considerades.....	114
Taula 4.6.: Categorització de l'estat de l'habitatge.....	121
Taula 4.7.: Relació de nombre de banys i utilitat.....	123
Taula 4.8.: Diferencials per característiques.....	124
Taula 4.9.: Regla d'homogeneïtzació acceptable.....	127
Taula 5.1.: Definició de camps utilitzats en el procés d'homogeneïtzació.....	142
Taula 5.2.: Marc mostral.....	146
Taula 5.3.: Característiques de les mostres. Exemple 1.....	148
Taula 5.4.: Estudi dels diferents ECV per a la mostra de l'Hospitalet del Llobregat.....	149
Taula 5.5.: Estudi dels diferents ECV per a diferents mostres.....	149
Taula 5.6.: Relació percentual entre la selecció òptima i el ECV_4	149
Taula 5.7.: Relació monetària entre la selecció òptima i el ECV_4	150
Taula 5.8.: Característiques de les mostres.....	151

	Pàg.
Taula 5.9.: Solucions vàlides obtingudes.....	154
Taula 5.10.: Mètodes i solucions escollides.....	155
Taula 5.11.: Error absolut i percentils.Tots els casos possibles.....	155
Taula 5.12.: Error absolut i percentils. Casos solucionats a la primera iteració.....	155
Taula 5.13.: Relació lineal valor taxat – valor estimat. Primera iteració.....	156
Taula 5.14.: Relació lineal valor taxat – valor estimat. Resultat després de cinc iteracions.....	157
Taula 5.15.: Solucions categoritzades.....	157
Taula 5.16.: Objecte de treball.....	160
Taula 5.17.: Característiques de les mostres.....	162
Taula 5.18.: Estudi de tendència sobre el conjunt de dades homogeneïtzades.....	164
Taula 5.19.: Qualitat dels mètodes d'estimació.....	164
Taula 5.20.: Diferències entre els indicadors de qualitat a una volta i a cinc voltes sense aplicació de l'algoritme.....	166
Taula 5.21.: Valors obtinguts.....	167
Taula 5.22.: Taxació a valorar.....	169
Taula 5.23.: Selecció inicial sense transformació de l'espai.....	170
Taula 5.24.: Selecció inicial amb transformació de l'espai en funció del variograma anisotròpic.....	170
Taula 5.25.: Resultats inicials pels diferents mètodes.....	171
Taula 5.26.: Resultats finals pels diferents mètodes.....	171
Taula 5.27.: Selecció inicial amb transformació de l'espai en funció del variograma anisotròpic.....	172
Taula 5.28.: Resultat bootstrap sobre 15 mostres.....	173
Taula 5.29.: Resultats del procés de valoració.....	174

Índex de figures i gràfics	Pàg.
Figura 1.1.: Eines i usos del procés d'estimació.....	19
Figura 1.2.: Índex gràfic de les eines utilitzades durant el procés d'estimació.....	20
Figura 1.3.: Esquema detallat del procés d'estimació.....	21
Figura 2.1.: Detall de la conceptualització de la mostra sobre la base de l'ECO.....	36
Figura 2.2.: Detall sobre la conceptualització dels mètodes de taxació.....	37
Figura 2.3.: Detall de la conceptualització de l'objecte de valoració sobre la base de l'ECO/805/2003.....	37
Figura 3.1.: Detall del mòdul d'introducció d'objectes a valorar.....	41
Figura 3.2.: Mòdul d'introducció de mostres.....	41
Figura 3.3.: Mòdul d'introducció de mostres. Geoposició.....	42
Figura 3.4.: Representació de la distribució bootstrap.....	46
Figura 3.5.: Spline en 2D.....	56
Figura 3.6.: Representació d'una interpolació spline. Representació de punts influents..	56
Figura 3.7.: Representació d'una interpolació spline. Representació de punts influents..	58
Figura 3.8.: Distribució espacial.....	64
Figura 3.9.: Representació del càlcul del variograma.....	65
Figura 3.10.: Exemple de variograma.....	66
Figura 3.11.: Variograma Cloud.....	68
Figura 3.12.: Representació en diagrama de caixes del variograma Cloud.....	68
Figura 3.13.: Variograma direccional.....	69
Figura 3.14.: Exemple de dos variogrames realitzats en direccions diferents.....	69
Figura 3.15.: Variograma omnidireccional.....	70
Figura 3.16.: Representació de la forma de càlcul del variograma.....	71
Figura 3.17.: Variograma del model efecte llavor pur.....	73
Figura 3.18.: Variograma del model exponencial.....	74
Figura 3.19.: Variograma del model gaussià.....	75
Figura 3.20.: Variograma del model esfèric.....	76
Figura 3.21.: Convencions direccionals utilitzades en la geoestatística.....	79
Figura 3.22.: Representació gràfica de semivariogrames isotròpics.....	80
Figura 3.23.: Representació gràfica de anisotropia geomètrica.....	80

	Pàg.
Figura 3.24.: Representació gràfica de la anisotropia geomètrica en dues dimensions...	82
Figura 3.25.: Representació gràfica d'anisotropia combinada.....	83
Figura 3.26.: Representació estimació kriging en 2D.....	87
Figura 3.27.: Descomposició de la funció aleatòria en base al kriging universal.....	93
Figura 3.28.: Mostres en l'espai considerant els 7 veïns més propers.....	97
Figura 3.29.: Representació gràfica de les tres millors estimacions.....	100
Figura 4.1.: Comparabilitat en funció de la superfície.....	118
Figura 4.2.: Corbes de l'ARC, segons estat.....	121
Figura 4.3.: Corbes de l'ARC, segons qualitat.....	122
Figura 4.4.: Corbes de l'ARC, segons estat i qualitat.....	122
Figura 4.5.: Pàgina de recollida de dades.....	126
Figura 4.6.: Oferta fora de mercat que participa en el procés d'estimació.....	130
Figura 4.7.: Existència de dos mercats propers.....	131
Figura 4.8.: Representació de distribució a partir de l'estimació per bootstrap.....	136
Figura 5.1.: Valor taxat (x) – Valor estimat (y).....	150
Figura 5.2.: Localització de les mostres inicials amb preus no homogeneïtzats.....	160
Figura 5.3.: Localització de les mostres comparables a l'objecte creat.....	161
Figura 5.4.: Localització de les mostres comparables homogeneïtzades respecte l'objecte tipus.....	163
Figura 5.5.: Mapa de preus i espai interpolat mitjançant la tècnica dels triangles de Delaunay. Primera iteració.....	165
Figura 5.6.: Mapa d'influències de les mostres. Mostra de Comte Borrell.....	165
Figura 5.7.: Preus estimats després de cinc iteracions.....	166
Figura 5.8.: Mapa de preus i espai interpolat mitjançant la tècnica dels triangles de Delaunay. Cinquena iteració.....	167
Figura 5.9.: Gràfic bootstrap: resultat de les valoracions realitzades a partir del remostreig de 15 mostres properes utilitzades de 6 en 6.....	173
Figura 5.10.: Exemple d'informe de comprovació en fase de proves.....	175
Figura 5.11.: Exemple d'informe de comprovació en fase de proves.....	176
Figura 5.12.: Exemple d'informe de comprovació en fase de proves.....	176

1. Introducció

1.1. Context del desenvolupament de la tesi

L'any 2006 es van realitzar a l'estat espanyol prop de dos milions de taxacions hipotecàries, la major part de les quals van fixar el seu valor en funció del mètode de comparació de mercat. Les directrius per a la realització de les valoracions estan recollides a l'ordre ministerial ECO 805/2003; llei que defineix els objectius i les intencions d'una taxació, però no n'especifica, en el cas del mètode de comparació, el mètode per a arribar a obtenir el valor de taxació.

Aquesta manca de claredat en els procediments i la indefinició de les metodologies a utilitzar en el procés d'obtenció del valor per comparació de mercat han comportat que es produís una relaxació metodològica.

El contrast cada vegada més notori entre l'objectiu ambicionat per la llei, basat en la tecnificació típica d'una ciència com l'arquitectura, i la subjectivitat que impera en l'execució del procés, agreujat tot plegat per la interferència de factors externs com a agents de pressió, han col·locat el

sector en una posició poc còmoda i en certa manera precària, convertint l'informe de taxació en un tràmit a superar.

Actualment, el mètode de comparació de mercat és el mètode més utilitzat per a valorar béns immobles amb objectius hipotecaris, arribant a ser el mètode encarregat d'imputar el valor de taxació en quasi la totalitat dels casos de valoració d'habitatges plurifamiliars.

En el transcurs d'un procés de valoració pel mètode de comparació es duen a terme dos processos d'homogeneïtzació. El primer d'aquests processos tracta la informació relacionada amb les característiques constructives que enfronten la mostra i l'objecte a valorar. Mentre que el segon procés busca l'homogeneïtzació de la localització, és a dir, estudia les diferències de valor que pot originar un factor com la ubicació de l'habitatge a valorar.

L'homogeneïtzació que es deriva de la localització és un dels punts més complexos del procés de valoració donat que no es pot estudiar com una relació un a un, sinó que s'hauria d'obtenir de l'estudi de la totalitat del conjunt mostral.

Els plantejaments que es presenten en aquesta tesi i que tenen com a objectiu l'estimació del valor de mercat a partir d'un sistema expert capaç de realitzar ambdues homogeneïtzacions, han estat punts centrals del programari que actualment s'està desenvolupant i utilitzant a l'empresa CATSA.

Aquesta tesi s'ha dut a terme en el marc d'un projecte empresarial de caire privat iniciat per l'empresa de taxació CATSA en col·laboració amb el CPSV, i ha estat desenvolupat durant el període 2004-2008. L'objectiu del projecte del qual forma part aquesta tesi era aconseguir determinar un procediment que permetés trobar el valor de taxació a partir d'una metodologia plantejada per 'parts'. Aquest mètode havia de ser capaç de simular el comportament d'un expert en tres grans processos: la selecció de la mostra, l'homogeneïtzació i, per últim, la interpolació espacial.

En aquest sentit, es plantejava crear un procés automatitzat que donés com a solució el valor de taxació a partir d'un sistema basat en una sèrie de

procediments tecnificats, i que tingués prou capacitat per a donar explicacions del procediment per parts realitzat.

1.2. Antecedents de la investigació

Pel que fa als antecedents d'aquesta tesi, l'exposició es centrarà en els dos àmbits que més presència tenen en aquesta part del projecte: els models hedònics i l'aplicació dels mètodes d'estimació espacial.

El primer model matemàtic que va ser aplicat a la valoració de béns immobles data de l'any 1928 (Waugh), però no va ser fins l'any 1939 que Court va intentar modelitzar, en l'àmbit automobilístic, el valor dels automòbils a partir de les seves característiques específiques. En la modelització de Court ja s'acomplia la hipòtesi bàsica dels models hedònics, que consisteix en assumir que les mateixes característiques dels objectes són les que imputen valor als objectes a partir de les funcions d'utilitat que presenten aquestes característiques.

El desenvolupament teòric sobre les funcions hedòniques va ser obra de Rosen (1974): els models teoritzats identificaven la relació entre els preus de béns complexos com ara els habitatges amb l'avaluació individual dels atributs.

Els treballs posteriors com els de Witte, Sumka, Erekson (1979), Brown i Rosen (1982) o Palmquist (1984) van continuar desenvolupant el procés teòric tot asseverant que era possible descriure els béns a partir d'atributs o característiques que havien de ser tractats explícitament pels mercats, però que els preus implícits d'aquests atributs es podien aconseguir a partir de les regressions hedòniques.

En relació amb aquesta temàtica s'ha generat una extensa bibliografia de referència, de la qual val remarcar els treballs de P. Linneman (1980), R. A. Dubin i C. Hsing (1990), A. Can (1992); i en l'àmbit nacional, les obres de J. Roca (1988) i P. Garcia (2005).

La teoria de les variables regionalitzades (Matheron, 1965) va ser formalitzada a partir dels estudis empírics duts a terme per l'enginyer de mines D. G. Krige (1951); i els mètodes geoestadístics basats en aquesta teoria conformen una eina molt efectiva per a estudiar la distribució espacial. Les aportacions de Anselin (1988) i Cressie (1991) van generalitzar l'ús i

l'aplicació de les eines geoespacial a àmbits com el dels mercats immobiliaris [Anselin (1990), R. Dubin (1992), R. K. Pace, R. Barry, i C. F. Sirmans (1998)].

Alguns exemples de la seva aplicació a l'estat espanyol es recullen als articles de J. M. Chica, (1995), R. Cano (2003, 2004), i J.M. Montero (2004).

En tots aquests treballs el tret comú -i alhora punt fonamental del procés- és el càlcul basat en la localització; a la vegada que s'hi descarten completament -o se'n fa un tractament molt residual- les característiques i la recerca de comparables reals.

El primer document a l'estat espanyol que inclou una referència clara al tema del tractament de característiques en un procés de correlació espacial és el treball de J.M. Montero i B. Larraz (2006). Més concretament, és a la tesi doctoral de B. Larraz (2006) on l'autor l'esmenta tot afirmant que: *“Con el fin de aislar la componente espacial de los precios de viviendas y locales comerciales, se estimaron, mediante un análisis de la varianza, los efectos que tienen las características del bien inmueble sobre el precio. De esta manera, una vez eliminadas dichas características del precio, se consigue aislar el efecto localización geográfica”*.

A pesar que en el document no s'exposen directament els procediments d'homogeneïtzació de les característiques constructives i d'estat, es fa evident que el procés d'estimació realitzat per B. Larraz es va realitzar en dues parts; en una primera part es va tractar la homogeneïtzació de les característiques constructives i d'estat dels habitatges, en una segona part es va estudiar a l'estimació relacionada amb la localització. En el treball de B. Larraz els objectes matemàtics que representen els habitatges homogeneïtzats resultants dels processos de càlcul són anomenats 'immobles equivalents', nomenclatura que ja havia estat utilitzada anteriorment per J. Roca (1986) i que avui dia és d'ús generalitzat i està dins l'àmbit de la valoració immobiliària.

Aquesta tesi s'ha concebut com l'últim pas d'un procés per parts que necessita del correcte funcionament dels processos previs a l'homogeneïtzació espacial per arribar a solucions de qualitat.

Un dels punts de gran importància i pes en el desenvolupament de la tesi ha estat l'obligació i alhora necessitat d'adequar-se a l'ECO 805/2003 i als procediments que habitualment es duen a terme en l'àmbit de les valoracions.

1.3. Objectius de la investigació

Dins del projecte empresarial de CATSA, aquesta tesi es va concentrar en l'estudi dels mètodes d'estimació espacial amb l'objectiu de trobar una metodologia aplicable en el procés de valoració que complís tant amb les imposicions legals en funció de la llei vigent com amb els objectius de treball i empresarials de la companyia.

Partint d'aquest context, l'objectiu principal d'aquesta tesi era la creació d'una metodologia d'homogeneïtzació del valor de taxació en funció de la localització dels immobles que actués dins del marc legal vigent establert per l'ECO805/2003. Per aconseguir aquest objectiu es van plantejar i estudiar els diferents mètodes d'interpolació i d'estimació espacial i es van plantejar les diverses possibilitats de tractament de la informació.

El segon objectiu era establir un mètode de selecció entre els diferents estimadors del valor de taxació que permetés minimitzar l'error d'estimació i fos alhora un indicador de la qualitat de l'estimació.

El tercer objectiu era la presentació d'una metodologia que permetés calcular la funció de probabilitat empírica del valor de mercat per a un bé immoble determinat. És de tots conegut que el valor d'un bé no és únic (Roca, 1986), sinó que té una funció de distribució. Gràcies a les metodologies de computació existents en l'actualitat va ser possible realitzar una representació del valor del bé immoble en funció del conjunt de la informació mostral disponible.

La resta de procediments i mètodes presentats parteixen de la necessitat de tractar amb béns comparables i homogeneïtzats pel que fa a les seves característiques constructives. Tots ells són objectius del projecte de CATSA i els seus desenvolupaments seran explicats en la mesura que siguin necessaris per a fer entenedor el punt de partida d'aquesta tesi. Els procediments i els mètodes necessaris per a la realització del projecte de CATSA i emprats en aquesta tesi van ser els següents:

- L'optimització de les bases de dades i la integració de la localització com a criteri important no només per a les valoracions sinó també per

a les mostres. Sense resoldre aquest punt hagués estat impossible arribar a solucions de qualitat en el context de les valoracions.

- El desenvolupament d'un sistema d'homogeneïtzació en funció de l'estat de conservació, la qualitat, les característiques físiques i la utilitat d'aquestes en relació al conjunt de l'immoble. La solució va ser un model additiu en què els elements del model es plantejaven com a diferencials monetaris derivats de cada una de les característiques. En aquest aspecte aquesta part del procés acomplia perfectament el plantejament dels models hedònics.
- La creació d'un sistema de selecció de mostres comparables basat en les característiques, però també en els diferencials que provoca l'homogeneïtzació.
- La demostració com una metodologia plantejada per parts basada en un sistema expert pot generar resultats de qualitat suficient per a ser un ajut important en el procés de valoració.

En base a aquests objectius les eines que van ser necessàries per al desenvolupament d'aquesta tesi van ser: la definició de bases de dades, els funcionaments dels sistemes experts, la tecnificació de camps constructius, els mètodes estadístics de remostreig, correlació espacial i regressió, les funcions hedòniques obtingudes a partir dels treballs de camps realitzats a CATSA, les eines GIS i l'estudi dels procediments legals relacionats amb el procés de valoració.

1.4. Justificació

1.4.1. Justificació empresarial

L'empresa CATSA va decidir l'any 2004 iniciar un procés d'automatització del sistema de valoració, el primer pas del qual va ser crear i desenvolupar un procés d'ajuda a la taxació.

L'objectiu final en aquells inicis era crear un programari que permetés dur a terme l'estudi del valor de taxació de qualsevol bé immoble en base al mètode que es determinés com a més adient.

El propòsit era crear un conjunt de sistemes, mètodes i procediments que permetessin a CATSA presentar els seus informes d'una manera fiable, ràpida i contrastable, tot complint amb la lògica experta i amb la legalitat vigent.

El procés d'investigació i desenvolupament iniciat per CATSA tenia com a finalitat aconseguir un producte diferenciador amb un nivell de qualitat superiors a qualsevol producte existent en el mercat actual.

Els processos desenvolupats per CATSA durant aquests anys s'han definit en base a l'ECO805/2003, però també tenint en compte la realitat de l'àmbit, per tant les polítiques d'obtenció de procediments òptims han quedat restringides al marc que imposa la llei i al funcionament habitual del sector.

1.4.2. Justificació de la tesi

Les metodologies basades en processos matemàtics i estadístics que s'han desenvolupat recentment amb l'objectiu d'obtenir el valor de taxació de béns immobles han estat principalment els mètodes regressius (Garcia Almirall, 2005), les xarxes neuronals (J.M. Caridad i N. Ceular, 2001), els mètodes multicriteri (J. Aznar i F. Guijarro, 2005) i el mètode anomenat 'de les dues betes' (Ballesteró i Rodríguez, 1999).

Molts d'aquests procediments s'han basat en el principi de parsimònia per a la obtenció de les variables tot buscant un model únic i general capaç de

predir el valor de qualsevol tipus d'immoble. Aquest és el cas, per exemple, de les xarxes neuronals i els mètodes regressius.

Una de les principals desavantatges, no obstant, de les metodologies anteriorment esmentades radica en la seva incapacitat de facilitar una explicació de la solució adoptada, ja que la solució obtinguda no defineix quins elements han participat en el càlcul.

Moltes d'aquestes metodologies es basen en introduir tota mena d'habitatges en el procés de càlcul dels coeficients que cobreixin un hipercub en \mathfrak{R}^n prou gran per a fer interpolacions que permetin predir el valor de taxació d'un immoble. En aquest punt es produeix un distanciament amb el mètode de comparació que requereix la utilització únicament de béns equivalents. Una situació discordant amb el procés de comparació és la capacitat d'aquests mètodes per a generar sempre una solució, existeixi o no mercat de comparables a la zona a valorar. Per tant, el que s'obté amb els models tradicionals és un conjunt de dades indeterminat que amb més o menys encert sempre donarà com a resultat un valor de taxació.

La majoria d'aquest models no tenen la capacitat d'especificar quins béns han participat en el càlcul del valor de taxació i amb quina ponderació ho han fet, requeriments tanmateix exigits per l'ECO805/2003.

La localització és un dels punts fonamentals per a la valoració d'un habitatge. Els sistemes utilitzats habitualment en aquest context donen un valor mitjà del preu de l'habitatge en funció del polígon espacial en què es troba situat. En cas de tractar-se d'una regressió lineal aquest preu mitjà és la suma del coeficient β_0 i del resultat de la combinació lineal $\beta_{lloc} \cdot lloc$. Típicament, el β_0 es pot considerar com la mitjana global i β_{lloc} el multiplicador que acompanya la variable 'qualitat del polígon espacial de localització'. Aquesta variable de localització han estat en molts casos districtes, codis postals o districtes electorals, variables que en cap cas han estat plantejades per definir un indicador de qualitat de la zona i que per a moltes poblacions poden ser únics.

Aquests models són àmpliament criticables, ja que el que s'acumula a $\beta_0 + (\beta_{lloc} \cdot lloc)$ és un emmascarament del valor del sòl. Si aquest model

s'apliqués a zones on el valor del sòl suposés més d'un 60% del valor de taxació final –tot comptant que la variabilitat derivada de la qualitat possiblement no arribi a superar el 10% del valor final-, s'observaria que el marge d'error d'aquestes situacions és realment baix. No obstant això, si aquests procediments es traslladessin a zones on el sòl representés menys d'un 20% del cost final, les dificultats per a aconseguir models amb resultats acceptables s'accentuarien clarament. Cal tenir en compte a més que tots els models que tenen com a objectiu principal la minimització de l'error, com és el cas dels models regressius, poden acabar donant signes erronis a les variables utilitzades.

Una diferència important de la metodologia presentada en aquest treball respecte als mètodes estadístics clàssics és la capacitat per no presentar solució en cas de no existir un mercat de comparables.

Les característiques socioeconòmiques que habitualment s'utilitzen en els models de valoració i que se solen definir a partir de polígons espacials artificials com poden ser el codi o el districte postal (P. Garcia, 2005) es podran eliminar en cas de passar a treballar amb localitzacions exactes, ja que els vincles de veïnatge entre immobles es defineixen a partir d'aquestes localitzacions. Aquest fet suposarà treballar únicament amb variables intrínseques a l'habitatge.

Un canvi important en el plantejament envers els models referenciats és l'ús únicament de mostres comparables amb l'objecte per a la realització dels processos de càlcul. És evident que en treballar exclusivament amb una part de la base –és a dir, amb les mostres comparables- es perd capacitat predictiva, però es guanya fiabilitat i s'aconsegueix complir amb el marc legal establert per l'ECO805/2003.

A més cal destacar que amb la introducció dels models experts existeix la possibilitat de continuar millorant el sistema amb la introducció de noves regles que permetin definir la realitat del mercat d'una forma més precisa.

Els mètodes estadístics d'anàlisi de correlació espacials poden ser de gran utilitat en l'àmbit de la valoració immobiliària, però necessiten d'un procediment tècnic previ per tal de garantir-ne el funcionament acurat i d'una

base de coneixements que permeti que el sistema reconegui les irregularitats en l'espai continu que suposa la ciutat. Un procediment per parts com el plantejat en aquesta tesi pot ser una eina de gran ajuda tant per a entitats de valoració en els seus processos de taxació individual com per a institucions que tenen com a objectiu la valoració massiva.

1.5. Metodologia

Els àmbits de coneixement que han confluït en aquesta tesi han estat les matemàtiques, l'estadística, la informàtica, la valoració de bens immobles i les lleis que regeixen aquesta disciplina dins l'estat espanyol.

El conjunt de metodologies desenvolupades per a arribar a assolir els tres objectius plantejats va comportar la necessitat de desglossar l'ECO805/2003, d'utilitzar tècniques estadístiques d'estimació espacial, mètodes regressius lineals i no lineals, metodologia *bootstrap* i procediments basats en sistemes experts com els arbres de decisió o les bases de regles.

Per poder començar a treballar pròpiament va caler recopilar i assolir uns coneixements sòlids en l'àmbit de les valoracions. En aquest sentit, el primer pas va ser familiaritzar-se amb tots els procediments i les metodologies del procés expert de taxació. Durant aquesta fase del procés es van desenvolupar una sèrie de coeficients i metodologies amb l'objectiu de simular els procediments realitzats pels tècnics.

En primer lloc, es va dur a terme un desglossament exhaustiu de l'ECO805/2003 a partir del qual es va establir el procediment de taxació pel mètode de comparació de mercat. Aquesta especificació va establir la possibilitat de resoldre el procés de valoració d'una forma modular.

Donada la necessitat que el procés fos completament autònom, és a dir, que fos capaç d'arribar a una solució sense necessitat d'intervenció humana (tant en la selecció de mètodes d'estimació com en la selecció dels paràmetres) s'hi van haver d'introduir els mètodes d'estimació no lineals. La utilització d'aquests mètodes va permetre triar el model de definició del variograma; eina bàsica en el procés d'estimació espacial.

El processos d'estimació basats en el variograma i coneguts amb el nom de tècniques d'estimació *kriging*, van ser l'única eina d'estimació plantejada inicialment com a mètode d'estimació espacial, però es va fer evident que aquestes requerien de més informació de la que habitualment es troba durant un procés real de valoració, és per això que aquestes tècniques es van complementar amb les tècniques clàssiques d'interpolació basades en

distàncies. Donat que aquestes tècniques no són processos d'estimació no es possible calcular ni intervals de confiança ni cap altre estadístic que ens pugui ser d'utilitat. Per solucionar aquest problema es van introduir les tècniques *bootstrap*.

Les tècniques *bootstrap* (Efron i Tibshirani, 1993) acostumen a tenir com a objectius: l'estudi del biaix produït per distribucions irregulars, l'estimació de la funció de densitat i el càlcul d'estadístics com la mitjana o els intervals de confiança. En aquesta tesi s'han utilitzat les tècniques *bootstrap* per a la realització de la validació creuada i per a l'estimació de la funció empírica de probabilitat del valor estimat.

L'obtenció de la funció probabilista del valor de taxació a partir d'un sistema de remostreig dóna una imatge del comportament del mercat en l'entorn de l'immoble a valorar.

Tots els procediments presentats en aquesta tesi van dirigits a aconseguir:

- Una major implicació tècnica en el procés de taxació.
- Una depuració dels errors derivats de la subjectivitat.
- La creació d'un protocol de treball.

Les finalitats de tot aquest procés són:

- La generació d'una explicació per part del Sistema Expert de cada un dels passos realitzats.
- La identificació ràpida i eficaç d'errors comesos pel tècnic, gràcies a la presentació constant d'informació.

Els errors identificats s'hauran de poder solucionar a partir de la:

- Modificació dels pesos.
- Incorporació de noves variables.
- Definició de noves regles espacials, com per exemple les relacions de veïnatge.
- Definició de noves regles basades en les característiques físiques, l'estat i la qualitat de l'habitatge.

- Definició de noves regles legals.

Per poder desenvolupar aquesta tesi ha calgut assolir sòlids coneixements en les àrees següents:

- Bases de dades
- Sistemes experts
- Xarxes neuronals
- Teoria de la decisió
- Mètodes no lineals d'estimació
- Estadística espacial
- Metodologia *bootstrap*

1.6. Esquema de treball i estructura

Per a poder desenvolupar el Sistema Expert, ha calgut fer una anàlisi precisa sobre:

- La informació existent a les bases de dades de CATSA per establir les categories de les diferents variables.
- Totes aquelles variables que participen en el procés de taxació.
- La llei aplicable, tot desglossant-ne el contingut per tal d'extraure'n els camps d'ús obligatori.

A més, ha calgut:

- Estudiar i desglossar el concepte de la comparabilitat.
- Realitzar els estudis necessaris per a poder definir tots els processos d'homogeneïtzació sobre característiques físiques, estat i qualitat dels immobles.
- Definir les regles lògiques no modelitzables i els arbres de decisió necessaris per a simular el comportament humà.
- Realitzar un recull de mètodes i estudiar-ne les característiques i possibilitats.
- Plantejar les necessitats derivades del procés d'homogeneïtzació espacial: qualitat dels estimadors, selecció del mètode d'estimació, possible creació d'algoritmes iteratius.

En base a totes les necessitat plantejades anteriorment l'estructura de la tesi serà la següent:

En el capítol segon es presenta un breu apunt sobre l'anàlisi realitzat de l'ECO805/2003.

El capítol tercer aplega totes les necessitats teòriques per a la realització dels desenvolupaments plantejats a la tesi.

En l'apartat 3.2. s'exposen les eines que han estat utilitzades per a aquest projecte, com per exemple les bases de dades o els nous sistemes d'introducció d'informació.

En l'apartat 3.3. es presenta el conjunt d'eines estudiades i creades per a la realització d'aquesta tesi.

En l'apartat 3.4 s'enumeren i detallen les eines basades en els mètodes *bootstrap* desenvolupats per Efron i Tibshirani, que van possibilitar l'estudi de la qualitat de l'estimació basada en la validació creuada com ara els mètodes de remostreig que van permetre generar l'histograma del valor estimat.

En l'apartat 3.5. es desenvolupen els diferents plantejaments teòrics sobre els mètodes d'interpolació espacials directes, com el mètode de les mitjanes mòbils i el cas particular de la distància inversa, el mètode de la mínima curvatura i les funcions radials.

En els apartats que 3.5.5. i 3.5.6. es fa una descripció exhaustiva de l'anàlisi estructural i de la metodologia *kriging*.

En l'apartat 3.6. es presenta el mètode de Gauss-Newton, mètode escollit per realitzar el procés d'estimació del variograma a utilitzar en els processos *kriging* posteriors.

En el capítol quart s'explica pas a pas el mètode de comparació de mercat i el funcionament bàsic de treball del sistema expert. S'hi presenten també les necessitats del sistema i els camps d'introducció obligatòria. En l'apartat 4.4. s'estableixen els criteris de selecció de mostres comparables en base a filtres excloents. En l'apartat 4.5. s'inclou un breu resum dels mètodes d'homogeneïtzació de les característiques físiques, d'estat i qualitat de l'immoble. En l'apartat 4.6. es presenta l'estudi estadístic que es va realitzar amb l'objectiu de demostrar que el procés d'homogeneïtzació per parts realitzat pel sistema expert concordava amb el procés d'homogeneïtzació realitzat pels tècnics. En l'apartat 4.7. s'expliquen les diferents metodologies d'homogeneïtzació utilitzades. Finalment en l'apartat 4.8. es presenta un mètode per representar la funció del valor de taxació. Aquesta representació, a més d'encloure la definició de la funció de valor, és un indicador de la complexitat de la zona.

En el capítol cinquè es desglossen diversos exemples amb la finalitat d'il·lustrar tota la teoria exposada fins el moment. En aquests exemples es

recrea gràficament l'aplicació del sistema a la valoració massiva, però també s'hi presenta un procediment complet individual de valoració.

Finalment en el capítol sisè es presenten les conclusions de la tesi i es recorren els possibles desenvolupaments futurs.

El capítol setè conté la recopilació de les fonts bibliogràfiques consultades per a realitzar aquesta tesi.

Les figures 1.1., 1.2. i 1.3. d'aquest capítol mostren de manera esquematitzada l'estructura del procés d'estimació. A la figura 1.1. es presenten les diferents eines acompanyades respectivament d'una breu descripció de l'ús.

A la figura 1.2. es relacionen els diferents mètodes amb els capítols o apartats en què han estat definits.

Finalment a la figura 1.3. es presenta de manera detallada cada una de les eines utilitzades.

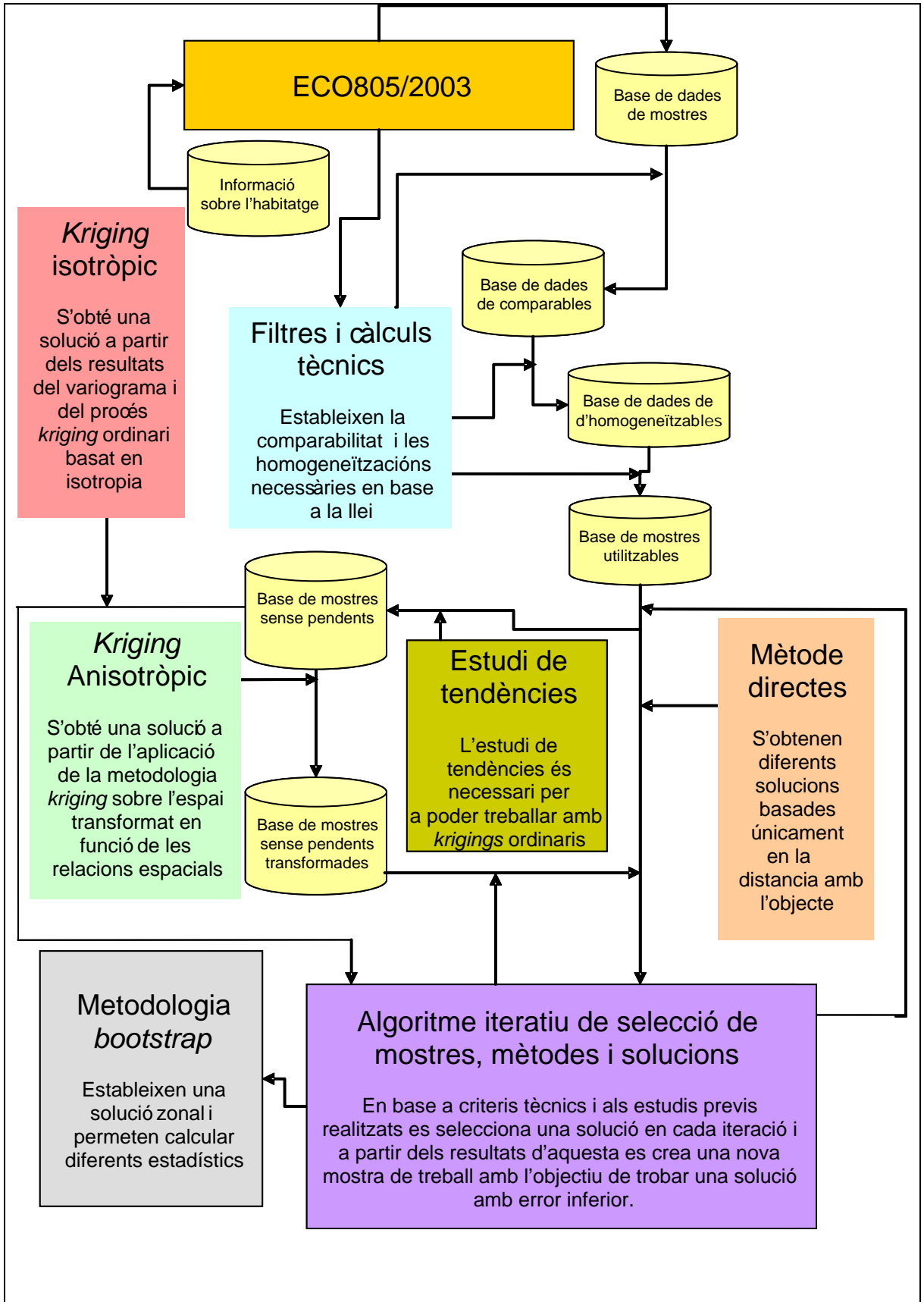


Figura 1.1.: Eines i usos del procés d'estimació

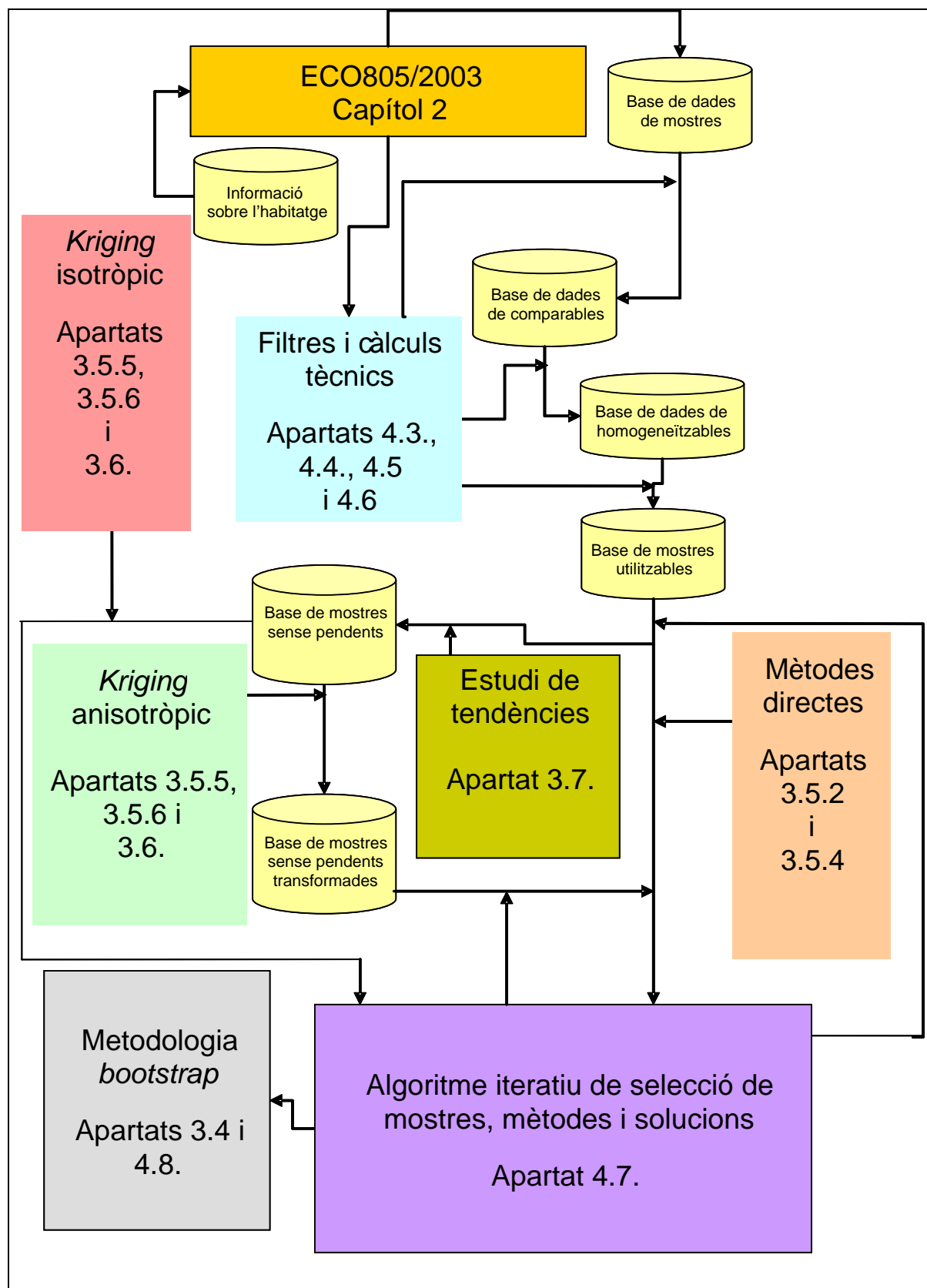


Figura 1.2.: Índex gràfic de les eines utilitzades durant el procés d'estimació

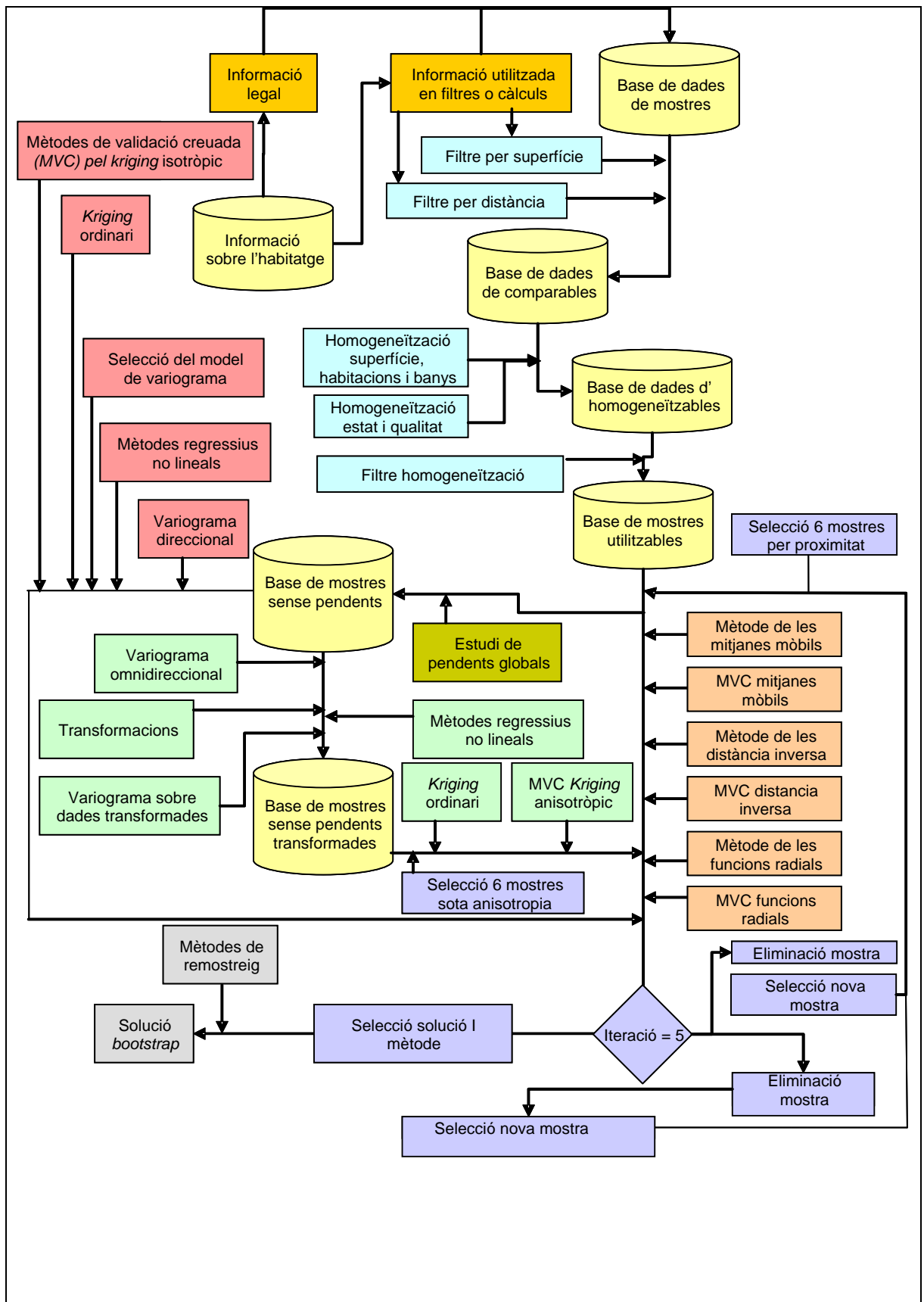


Figura 1.3.: Esquema detallat del procés d'estimació

2. L'ECO805/2003. El marc teòric

2.1. Introducció

En aquest capítol de l'estudi es desglossen aquells articles de l'ordre del ministeri d'economia ECO 805/2003 que tenen relació amb el mètode de valoració per comparació de mercat i que hauran de ser considerats pel sistema expert.

Per a poder dur a terme el mètode de comparació amb garantia caldrà efectuar una identificació adequada i acurada tant de la localització com de l'estat i les característiques de l'immoble i de la zona on està ubicat.

Així mateix, caldrà realitzar -per requeriments legals- la identificació de totes les parts participants en el procés.

2.2. Apartats de l'ECO805/2003 relacionats amb el mètode de mercat

Una de les primeres necessitats en la creació d'un programari és efectuar l'anàlisi i el desglossament dels elements participants fins aconseguir-ne la conversió en regles al més senzilles possibles. L'esquematització que es presenta seguidament ha estat realitzada pel departament d'I+D de CATSA. A mode sintètic, en aquesta tesi només se'n presenten aquells punts considerats de major interès per a aconseguir els objectius plantejats.

2.2.1. Identificació i localització

L'ordre ECO 805/2003 estableix en l'article 67 la informació requerida en aquest apartat.

Condicció	Objecte	Valors possibles
Per a totes les taxacions	Tipus d'immoble	Edifici. Element d'edifici . Terreny. Terreny no urbanitzable. Finca rústica.
	Estat de construcció	Acabat. En projecte. En construcció. En rehabilitació.
	Tipus de via	Carrer, avinguda, passeig, passatge,...
	Via	Obtingut del Cadastre
	Nombre	
	Escala	
	Pis	
	Porta	
	Codi INE	
	Codi postal	
	Població	
	País	Espanya
	Província	
	Plànol d'ubicació	
	Registre de la propietat	
	Unitat registral	
	Volum	
	Llibre	
	Secció	
	Foli	
Finca		
Inscripció		

Taula 2.1.: Identificació i localització.

Amb la informació obtinguda durant el procés d'identificació i localització és possible calcular la geoposició i la planta on es troba l'objecte.

Per obtenir la geoposició són necessaris els camps següents: població, tipus de via, via i nombre. Per aconseguir la informació relacionada amb la planta cal saber-ne el pis.

A partir de la localització geogràfica de la població, les seves característiques i les qualitats dels materials utilitzats és possible calcular el cost de construcció aproximat en base a les dades aportades pels col·legis d'arquitectes de l'estat.

La resta d'informació recollida en aquesta taula són de caire legal.

2.2.2. Localitat i entorn

L'ordre ECO 805/2003 estableix en l'article 69 la informació requerida en aquest apartat.

Condicció	Objecte	Valors possibles
Per a totes les taxacions	Tipus de nucli	
	Ocupació laboral predominant	
	Població de dret	
	Evolució recent de la població	
	Característiques bàsiques de la localitat de què depèn	
Si tipus immoble = edifici, element d'edifici , terreny	Nivell de renda	
	Característiques d'homogeneïtat arquitectònica	
	Usos dominants	
	Infraestructures	
	Equipaments	
	Serveis	
	Comunicacions	
	Aparcaments	
	Nivell de desenvolupament edificatori de l'entorn	
	Antiguitat característica	
	Renovació recent	

Taula 2.2.: Localització i entorn

Tota la informació de la taula 2.2 són conceptes que no s'utilitzaran en els càlculs. Cal remarcar que en els models anteriors (Garcia-Almirall, 2004) en què no era possible utilitzar la localització per geoposició, la informació de la taula 2.2 era de gran utilitat alhora d'establir *clusters* d'informació.

2.2.3. Descripció i superfície del terreny

L'ordre ECO 805/2003 estableix en l'article 70 la informació requerida en aquest apartat.

Condicció	Objecte	Valors possibles
Si tipus immoble = element d'edifici	La informació del terreny és molt rellevant sobre el valor?	
	Unitat de superfície	
	Superfície comprovada	
	Superfície registral	
	Superfície projecte	
	Superfície adoptada	Útils, Construïda, construïda amb elements comuns
Si tipus immoble = edifici, element d'edifici , terreny urbà (2)	Infraestructura exterior existent	
	Infraestructura exterior pendent	
	Percentatge d'obra realitzada	
	Inversió pendent	

Taula 2.3.: Descripció i superfície del terreny.

La informació de la taula 2.3. serà utilitzada com a filtre, donat que únicament es valoraran elements d'edifici acabats. Ara per ara el programari només avaluarà habitatges plurifamiliars pel mètode de comparació amb un percentatge d'obra realitzada del 100%.

Els camps relacionats amb les superfícies s'utilitzaran per a realitzar comprovacions legals.

2.2.4. Descripció i superfície de l'edificació

L'ordre ECO 805/2003 estableix en l'article 71 la informació requerida en aquest apartat. Aquesta informació es divideix en dos grans blocs: la informació referent a l'edifici i la descripció de la unitat.

2.2.4.1. Descripció i superfície de l'edificació - Edificis

Condicció	Objecte	Valors possibles
Si tipus immoble = edifici, element d'edifici	Plànol de distribució d'edificacions i serveis en parcel·la	
	Plànol de distribució d'edificacions i serveis en parcel·la	
Si tipus immoble = edifici, element d'edifici	Unitat d'edificació	
	Descripció d'edificació	
	Ús de l'edifici	Habitatge unifamiliar Habitatge plurifamiliar Comercial Industrial
Si estat de la construcció = en construcció o rehabilitació	En la part d'obra encara no executada, es refereix al projecte base de llicència municipal	
	Fase de les obres	
	Percentatge d'obra executada	
Si tipus immoble = edifici, element d'edifici	Tipologia d'edificació	Aïllada Entre mitgeres
	Nombre de plantes sobre rasant	
	Nombre de plantes sota rasant	
	Dependència	Habitació Sala Cuina Bany Galeria Despatx Sotacobert Terrassa Balcó Traster Espais comuns
	Quantitat	
	Paviment	
	Paret	
	Sostre	
	Fusteria interior	
	Usos (3)	Habitatge Terrassa Trasters Garatge Jardí

Taula 2.4.: Descripció i superfície de l'edificació (1/2).

Condicció	Objecte	Valors possibles
Si tipus immoble = edifici, element d'edifici	Superfície total	
	Superfície útil	
	Superfície total construïda sense parts comunes	
	Superfície total construïda amb parts comunes	
	Elements comuns	
	Serveis generals interns	
	Segons ús, circumstàncies específiques rellevants	
	Fonaments	
	Estructura	
	Coberta	
	Tancaments exteriors	
	Fusteria exterior	
	Lampisteria	
	Calefacció	
	Aire condicionat	
	Gas	
	Electricitat	
	Ascensors	
	Prevenició d'incendis	
	Altres	
Terminacions i acabats rellevants		

Taula 2.4.: Descripció i superfície de l'edificació (2/2).

Els camps de la taula 2.4. que s'utilitzaran en el procés de càlcul plantejat seran: les superfícies, l'existència d'ascensor i les qualitats dels acabats dels tancaments exteriors, fusteria exterior i lampisteria.

A partir de la informació sobre la tipologia, el tipus d'edificació i l'ús es seleccionaran els coeficients relacionats amb les qualitats, les quals són el resultat d'un estudi previ realitzat pel departament d'I+D de CATSA.

2.2.4.2. Descripció i superfície de l'edificació – Elements d'edifici

Condicció	Objecte	Valors possibles
Si tipus immoble = edifici, element d'edifici	Usos (3)	Habitatge Terrassa Traster Garatge Jardí Oficina Local

Taula 2.5.: Descripció dels elements de l'edifici (1/2).

Condicció	Objecte	Valors possibles
Si tipus immoble = edifici, element d'edifici	Superfície total	
	Superfície útil	
	Superfície total construïda sense parts comunes	
	Superfície total construïda amb parts comunes	
	Dependències	Habitació Sala Cuina Bany Distribuïdor Galeria Terrassa Balcó Despatx Sotacobert Traster Aparcament Espais comuns
	Quantitat de cada tipus de dependència	
	Paviments de cada dependència	
	Parets de cada dependència	
	Sostres de cada dependència	
	Fusteria interior	
	Situació relativa	
	Annexos	
	Superfície d'annexos	
	Elements d'edifici	Característiques bàsiques de l'edifici
Característiques específiques rellevants per a cada ús		
Característiques constructives		
Instal·lacions rellevants		

Taula 2.5.: Descripció dels elements de l'edifici (2/2).

El nombre d'habitacions i de banys seran necessaris per a realitzar els càlculs del procés d'homogeneïtzació.

La valoració de trasters, terrasses i altres elements es podrà realitzar per comparació o a cost, tot depenent si existeix mercat de comparables. Els mètodes de recerca de comparables amb elements semblants o la decisió de realitzar el càlcul a cost no es detallaran en aquesta tesi.

El nombre de cuines i de banys seran dades necessàries per a poder executar l'algoritme de càlcul, donat que no es podrà considerar un habitatge plurifamiliar com a tal si no en consten aquests elements.

2.2.5. Descripció urbanística

L'ordre ECO 805/2003 estableix en l'article 72 la informació requerida en aquest apartat.

Camp	Incloure si...	Ref.
Grau d'adequació de característiques físiques i ús a normativa urbanística	Si tipus immoble = edifici o element d'edifici	72.1.a
Indicar si està subjecte a protecció urbanística o històrica individualitzada	Si tipus immoble = edifici o element d'edifici	72.1.a
Indicar si és conforme o no amb el planejament	Si tipus immoble = edifici	72.1.b

Taula 2.6.: Descripció urbanística.

Si l'objecte no és conforme al planejament o té algun tipus de protecció urbanística no podrà ser utilitzat ni com a mostra ni com a objecte, ja que el seu preu en cas de ser una mostra o el seu valor en cas de ser l'objecte a valorar pot no adequar-se al mercat existent.

2.2.6. Règim de protecció, tinença i ocupació

L'ordre ECO 805/2003 estableix en l'article 73 la informació requerida en aquest apartat.

Camp	Opcions	Incloure si...	Ref.
Estat d'ocupació	No ocupat Ocupat Llogat		73.a
Títol d'ocupació		Si estat d'ocupació = ocupat	73.a
Característiques títol d'ocupació		Si estat d'ocupació = ocupat	73.a
Termini de vigència de títol d'ocupació		Si estat d'ocupació = ocupat	73.a
Indicar si està destinat a activitat econòmica	Si/No		73.b
Indicar si el seu ús o destí és el lloguer	Si/No		
Característiques de l'activitat o explotació albergada o prevista		Si activitat econòmica = Si	73.b

Taula 2.7.: Règim de protecció, tinença i ocupació (1/2).

Camp	Opcions	Incloure si...	Ref.
Indicar si té algun dret u obligació que afecti el contingut del dret real	No Dret de superfície Concessió administrativa Servitut Nuda propietat Usufructe Ús i habitació Limitacions de domini Opcions de compra Temps compartit sobre immobles Compromisos de compra a terminis		73.c
Característiques i efecte sobre el valor degut a dret u obligació que afecta el contingut del dret real		Si dret u obligació <> No	73.c
Indicar si té limitació legal	Si/No		73.d
Característiques de limitació legal		Si limitació legal = Si	73.d
Indicar si té limitació urbanística	Si/No		73.d
Característiques de limitació urbanística		Si limitació urbanística = Si	73.d
Indicar si té limitació protecció pública	Si/No		73.d
Característiques de limitació protecció pública		Si limitació protecció pública = Si	73.d

Taula 2.7.: Règim de protecció, tinença i ocupació (2/2).

L'ocupació, l'activitat econòmica i el tipus de tinença, limitacions i drets seran dades que obligaran a utilitzar, a part del mètode de valoració per comparació, altres mètodes de càlcul del valor de l'objecte. Tanmateix, caldrà mantenir a la base el valor obtingut per mercat -que podrà ser diferent del valor final del procés de valoració- si es vol poder utilitzar aquest tipus de béns en el procés de valoració pel mètode de comparació.

2.2.7. Anàlisi de mercat

L'ordre ECO 805/2003 requereix incloure en l'informe de valoració la informació següent:

Camp	Ref.
Característiques del segment del mercat de l'objecte de valoració	74
Indicar diferències apreciables entre el valor de mercat i el valor hipotecari del segment	74
Oferta	74
Demanda	74
Interval de preus de venda al comptat	74
Interval de preus de lloguer	74
Expectatives d'oferta	74
Expectatives de demanda	74
Expectatives de revaloració	74

Taula 2.8.: Anàlisi de mercat.

L'interval de preus de venda al comptat s'obtindrà a partir dels procediments de remostreig que es presentaran en el capítol 4 i s'exemplificarà en el capítol 5.

Una altra de les indicacions de l'ECO 805/2003 en relació amb les mostres és el nombre mínim de mostres de mercat necessàries per a poder realitzar el procés d'estimació. En aquest sentit, L'ECO 805/2003 fixa en sis el nombre de mostres necessàries.

Segons consideracions dels tècnics aquest nombre n'és suficient i fa innecessària la recopilació de més mostres quan es disposa de les sis exigides. En aquest context, caldrà establir mètodes d'estimació del valor a partir de sis mostres.

2.2.8. Dades i càlcul dels valors tècnics

L'ordre ECO 805/2003 requereix incloure-hi la informació següent:

Camp	Opcions	Incloure si...	Ref.
Hipòtesis d'edifici acabat	Si/No	SI...	
Obra parada?	Si/No		
Indicar mètodes de taxació (MÈTODE)	Comparació Actualització Residual estàtic Residual dinàmic Cost		75.1
Indicar criteris de taxació			75.1
Indicar els càlculs realitzats per determinar el valor			75.1
Dades de mínim 6 immobles comparables utilitzats		Si MÈTODE = comparació	75.2
Immobles comparables: província		Si MÈTODE = comparació	75.2
Immobles comparables: població		Si MÈTODE = comparació	75.2
Immobles comparables: codi postal		Si MÈTODE = comparació	75.2
Immobles comparables: carrer		Si MÈTODE = comparació	75.2
Immobles comparables: número		Si MÈTODE = comparació i ús edifici <> edifici unifamiliar	75.2
Per a cada immoble: paràmetres, coeficients i/o ponderacions i/o desenvolupament estadístic utilitzats per a l'homogeneïtzació		Si MÈTODE = comparació	75.2

Taula 2.9.: Dades i càlcul dels valors tècnics.

La informació d'aquesta taula estableix el nombre de mostres necessàries per a la realització del mètode de comparació, alhora que exposa la necessitat de l'existència d'una metodologia de càlcul, donat que es requereix la identificació dels criteris de taxació. En el mateix context, cal exposar la informació relacionada amb la localització de la mostra. Això permetrà una identificació i el geoposicionament complet de les mostres utilitzades. Amb aquesta informació s'aconseguirà establir-ne les coordenades de localització que permetran realitzar el procés d'estimació espacial.

2.2.9. Valors de taxació, condicionats i advertiments, observacions, dates d'emissió, caducitat i firmes

L'ordre ECO 805/2003 requereix incloure-hi la informació següent:

Camp	Ref.
Valors tècnics resultants per a cada immoble	76.a
Valor de taxació considerant objecte i finalitat	76.b
Condicionats	76.c
Advertiments	76.c
Limitacions	76.d
Valor mínim d'assegurança d'incendis	76.e
Data de l'última visita a l'immoble	77.a
Data d'emissió de l'informe	77.b
Data de caducitat de l'informe	77.c
Noms, firmes i titulació o càrrec dels qui subscriuen l'informe	77.d
Nombre i titulació d'altres tècnics especialitzats que intervenen directament en la taxació	77.d

Taula 2.10.: Valors, condicionats i advertiments, observacions, data i firmes.

Aquesta informació és de caràcter legal en tots els seus camps excepte en la referència 76.a. en què cal indicar el valor aconseguit per mercat i per qualsevol altre mètode utilitzat.

2.2.10. Documentació annexa a l'informe

L'ordre ECO 805/2003 requereix incloure-hi la informació següent:

Camp	Opcions	Incloure si...	Ref.
Plànols a escala o croquis de situació i emplaçament de l'immoble en el municipi			78.1.a
Plànols a escala o croquis acotats de l'immoble			78.1.b
Aquelles necessàries per a definir i identificar característiques	Seccions Fotografies Altres		78.1.c
Documents utilitzats per a la identificació registral			78.2.a
Documents utilitzats per a la identificació cadastral			78.2.a
Còpia de llicència d'obres		Si estat = en construcció o rehabilitació	78.2.b
Resum del pressupost d'execució material		Si estat = projecte o en construcció o rehabilitació	78.2.c
Còpia de cèdula de qualificació o document que determini o permeti determinar el valor màxim legal		Si estat = en construcció i LIM PROP PUB = Sí	78.2.d
Còpia del certificat de titularitat i càrregues			78.2.e
Còpia de contracte de lloguer o títol de l'ocupació		Si estat ocupació = llogat	78.2.f
Certificat del propietari de la situació d'ocupació, rendes vigents i el seu estat de pagament i de les despeses imputables a l'immoble en l'últim any		Si estat ocupació = llogat i finalitat <> hipotecària	78.2.f
Còpia del balanç i compte de resultats dels últims 3 anys		Si activitat econòmica = Sí	78.2.g

Taula 2.11.: Documentació annexa.

Tota la informació de la taula 2.11. és estrictament legal i no ha estat utilitzada per a la realització dels càlculs de valor.

2.2.11. Conceptualització de la informació

A partir d'aquest desglossament de la llei el cos de tècnics de CATSA va establir les variables necessàries tant per a l'avaluació de les mostres (figura 2.1.) com dels objectes (figura 2.2.) a taxar; les condicions per a la realització dels diferents mètodes (figura 2.3.); i els procediments i les necessitats per al desenvolupament de cada mètode de valoració.

MUESTRAS	Vivienda				Oficina	Local	Uso Industrial				Aparc.	Terreno			Otros	Art. ECO				
	Unifamiliar Adosada	Unifamiliar Aislada	En suelo rústico (no urbanizable)	En suelo urbano (no urbanizable)			Local en edificio	Adosada	Aislada	En suelo rústico (no urbanizable)		M - Resid Unif. Als	MII - Industrial	MII - Rústico				V	V	V
Venta / Alquiler	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Venta / Alquiler			
Tipo	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Para Rústicos			
Fecha	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Fecha de Alta			
Tipo de fuente	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Particular; Inte			
Fuente	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b				
Teléfono o web de fuente	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b				
Precio global bruto	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Numerico > 0			
Renta bruta mensual	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Numerico > 0			
% comercialización	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Numerico; def			
% Bruta->Meta	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Numerico def.			
Domicilio																				
Tipo de vía	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde callejer			
Nombre de la vía	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde callejer			
Nombre de la vía (descripción)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde callejer			
Número	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Numerico > 0			
Planta	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Bajos; Entres			
Población	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde callejer			
Código Postal	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde callejer			
Barrio / Zona	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a				
Paraje	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a				
Provincia	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde callejer			
País	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde callejer			
Proximidad a núcleos de población	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Hasta 1km; Ha			
Geoposición (mostrar mapa)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Desde Cercali			
Error de geoposición	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.a	Sin error; Hast			
Datos																				
Tipo de superficie	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Util; Construí			
Superficie	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	numérico > 0 (
% a Construída «Comunes	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	numérico > 0 (
Sup. del terreno	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	numérico > 0 (
Subtipo de adosada	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Testero o par			
Habitaciones	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	numérico >=1			
Baños / Servicios (4)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	numérico >= 0			
Ascensor	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	0; 1; 2; 3; 4; má			
Capacidad de plaza	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Pequeña; Mex			
Longitud de fachada	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	numérico > 0 (
Profundidad	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	numérico > 0 (
Facilidad acceso	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Muy buena; Br			
Vistas y orientación	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Muy buena; Br			
Pendiente del terreno	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75.2.b	Inferior 10%; Ei			

Figura 2.1: Detall de la conceptualització de la mostra sobre la base de l'ECO/805/2003.

3. Eines necessàries per al desenvolupament de la tesi

3.1. Introducció

La decisió de CATSA de crear un nou programari de valoració basat en un sistema expert va comportar la necessitat de millorar les eines existents en base a les recomanacions dels experts consultats (P. Garcia, 2004), a les necessitats generades al llarg dels anys dins de la mateixa organització i al potencial de les noves tecnologies com ara els sistemes d'obtenció de la geoposició o les eines GIS.

De manera que el primer pas a seguir va ser efectuar una recopilació exhaustiva d'idees i la conseqüent conceptualització d'aquestes idees tot establint les necessitats, les prioritats i les fases del projecte. Les necessitats i prioritats del d'aquesta tesi es van definir a partir de la conceptualització i anàlisi realitzat al projecte. Tot i que moltes de les necessitats van ser idèntiques no així ho van ser els objectius. Aquesta tesi serà la base del desenvolupament dels mètodes espacials, encara que és possible que

certes parts siguin eliminades o simplificades degut als temps d'execució que suposen.

En els apartats 3.2. i 3.3. s'exposen de manera resumida les eines comercials bàsiques i específiques utilitzades per aquesta tesi.

En els apartats 3.3. i 3.4. es presenta el conjunt d'eines metodològiques necessàries per al desenvolupament d'aquesta tesi. En aquest context, val a dir que aquests procediments, tot i haver estat implementats en molts programaris comercials, no podien ser d'utilitat per a CATSA pel fet que no se'n disposava de l'especificació corresponent, i la intenció de l'empresa era reprogramar-los dins del nou aplicatiu de CATSA.

3.2. Eines del projecte de CATSA emprades en el desenvolupament de la tesi

Per poder dur a terme aquesta part del projecte es van haver de redefinir les bases de dades per tal d'aconseguir-ne una protocol·lització i tecnificació acurada i correcta. La definició de la base de dades i relacions establertes són propietat de CATSA i no poden ser presentades en aquesta tesi.

Es va haver de crear una nova eina de recollida d'informació que permetés una introducció ràpida i eficient de la informació. Aquesta eina es va realitzar amb Visual Basic del paquet .net de Microsoft (VB.net) i va ser concebuda sobre una interfície basada en botons de selecció, desplegable i camps d'autoemplenament.

Figura 3.1.: Detall del mòdul d'introducció d'objectes a valorar.

Figura 3.2.: Mòdul d'introducció de mostres.

Tanmateix, es va haver de crear especialment un aplicatiu nou que permetés geoposicionar tant les mostres (Figura 3.3.) com les taxacions. Aquest aplicatiu es va realitzar amb VB.net i és emprat per a enviar l'adreça exacta a un servidor de dades GIS de l'empresa Cercalia.

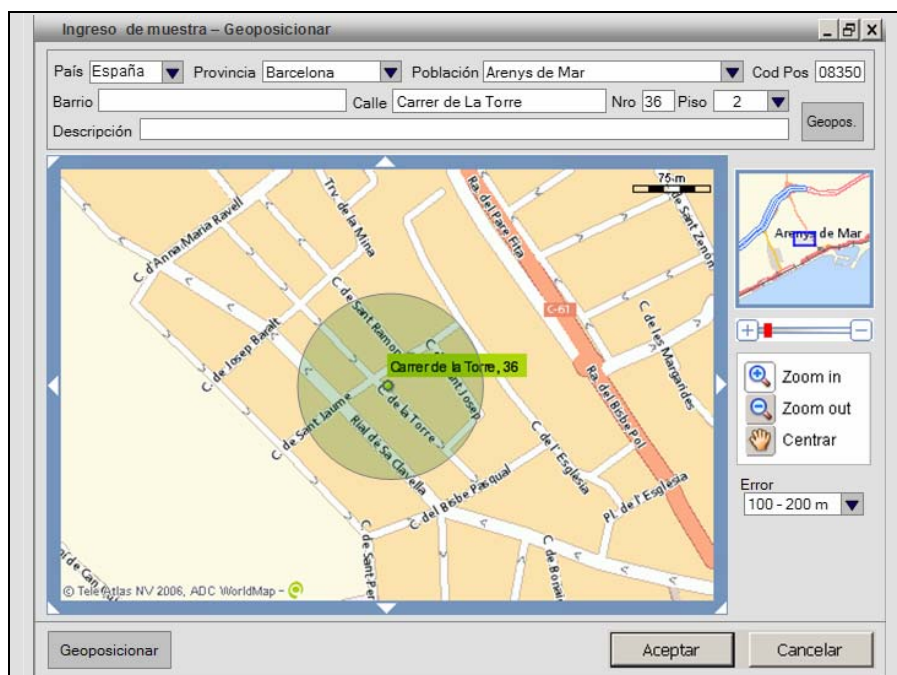


Figura 3.3.: Mòdul d'introducció de mostres. Geoposició.

3.3. Eines necessàries per al desenvolupament de la tesi

Per poder desenvolupar aquesta tesi va caler seleccionar un programari matemàtic-estadístic que permetés programar i incorporar mòduls nous sense interrompre el funcionament dels mòduls ja existents. En base a aquest requeriment, el programari triat va ser el SAS 9.1.3. Els mòduls més emprats per al desenvolupament d'aquesta tesi van ser:

Mòdul	Usos
BASE	Tractament de dades i estadística bàsica
MACRO	Programació
STAT	Càlculs estadístics i comprovacions dels mètodes programats
IML	Llenguatge matricial
GRAPH	Creació de gràfics
GIS	Visualització de mapes

Taula 3.1.: Mòduls de SAS.

El software utilitzat per a realitzar la implementació dels mètodes *kriging* han estat el SAS/BASE i el SAS/IML. El mòdul SAS/STAT va ser utilitzat per realitzar els càlculs de la modelització de la funció no lineal utilitzada per estudiar el variograma empíric.

Paral·lelament es va utilitzar el MS Excel per a realitzar processos matricials bàsics i poder comprovar el funcionament de manera visual.

A més dels ja esmentats, es va utilitzar el programari següent: el MAPinfo, l'ARCGIS d'ESRI i el Google Earth en la seva versió plus per a la generació de mapes i la creació de representacions gràfiques.

3.4. El remostreig

3.4.1. Introducció

Es denominen tècniques *bootstrap* al conjunt de tècniques elaborades per Bradley Efron (1979) i desenvolupades conjuntament amb Tibshirani (1993) que s'han revelat com una poderosa eina de gran utilitat en diversos camps de l'estadística. El *bootstrap* és un dels mètodes estadístics coneguts com a mètodes de computació intensiva.

Les tècniques de remostreig *bootstrap* permeten realitzar l'estimació de la distribució de mostreig sense la necessitat de realitzar supòsits sobre la distribució de les dades, donant com a resultats l'error estàndard, els intervals de confiança i la forma distributiva mostral. Per tant, la metodologia *bootstrap* permet efectuar inferències sense necessitat de postular prèviament que la distribució compleixi certes hipòtesis que, de vegades, són de difícil justificació o contrastació, tot permetent analitzar situacions complexes sense necessitat de deduccions analítiques.

La característica principal de tots els procediments *bootstrap* és l'ús de mostres extreptes de la mateixa mostra – que es designaran en endavant com a 'remostres' – i que permetran obtenir conclusions sobre la població si aquesta està correctament representada a la mostra inicial.

3.4.2. Plantejament bàsic

Es suposa una mostra aleatòria simple amb n d'elements $V=(v_1, v_2, \dots, v_n)$ amb funció de distribució $F(v)=Pr(V \leq v)$, per a totes les variables aleatòries V_i . Quan el valor del paràmetre θ poblacional és desconegut i, en conseqüència, cal utilitzar-ne l'estimador $\hat{\theta} = \hat{\theta}(v_1, v_2, \dots, v_n)$ és important conèixer la precisió d'aquest estimador. En moltes situacions pot resultar complicat definir aquesta precisió.

Si en una realització mostral del vector aleatori $V=(V_1, V_2, \dots, V_n)$ s'observa $V_1=v_1, V_2=v_2, \dots, V_n=v_n$ aleshores al vector de components (v_1, v_2, \dots, v_n) és la mostra original. Consegüentment, es pot afirmar que amb la metodologia

bootstrap les dades observades a la mostra original $v=(v_1, v_2, \dots, v_n)$ assumiran el paper de la vertadera distribució desconeguda de $F(v)$, quedant aquesta substituïda per la seva estimació $F_n(v)$, la qual acostuma a ser freqüentment la distribució empírica de (v_1, v_2, \dots, v_n) que assigna pesos $1/n$ a cada v_i .

En aquest sentit, la idea bàsica que defineix el *bootstrap* és que la mostra original obtinguda representa correctament la població; de manera que els estadístics extrets a partir de l'anàlisi de les remostres són equivalents als que s'aconseguirien si es fes un estudi a partir de mostres extretes de la població.

Si $V^*=(V^*_1, V^*_2, \dots, V^*_n)$ és una mostra aleatòria simple genèrica de F_n ; és a dir, cada V^*_i , $1 \leq i \leq n$ d'aquesta mostra s'obté independentment (amb reemplaçament) de la mostra original $v=(v_1, v_2, \dots, v_n)$ de forma que $V^*_i=v_j$, el conjunt $V^*=(V^*_1, V^*_2, \dots, V^*_n)$ se'l coneix com a mostra *bootstrap*. La notació $V^{*(b)}$ indica que ens referim a la b-èssima mostra *bootstrap* de la mostra natural que, de forma genèrica, es designa com:

$$V^{*(b)}=(V_1^{*(b)}, V_2^{*(b)}, \dots, V_n^{*(b)})$$

Habitualment el nombre total de remostres mínimes necessàries per a aconseguir la convergència és de 10.000.

L'algoritme de càlcul de l'estimador a partir de la metodologia *bootstrap* consisteix en els passos següents:

1. S'escollirà una mostra original $v=(v_1, v_2, \dots, v_n)$ per mostreig aleatori simple, sent aquestes les úniques dades reals a què es tindrà accés.
2. A partir de la mostra original s'obindrà una mostra *bootstrap* $V^*=(V^*_1, V^*_2, \dots, V^*_n)$ amb reposició.
3. La mateixa funció que defineix l'estadístic d'interès, $\hat{\theta}$, serà aplicada a la mostra construïda en el punt anterior, i amb això s'obindrà:

$$\hat{\theta}^{*(b)}=(V_1^{*(b)}, V_2^{*(b)}, \dots, V_n^{*(b)})$$

4. Es repetiran B vegades els punts 2 i 3.

5. Es construirà la distribució de probabilitat a partir dels B valors $\hat{\theta}^{*(b)}$ assignats amb freqüència relativa $1/B$ a cada punt $\hat{\theta}^{*(1)}, \hat{\theta}^{*(2)}, \dots, \hat{\theta}^{*(B)}$.

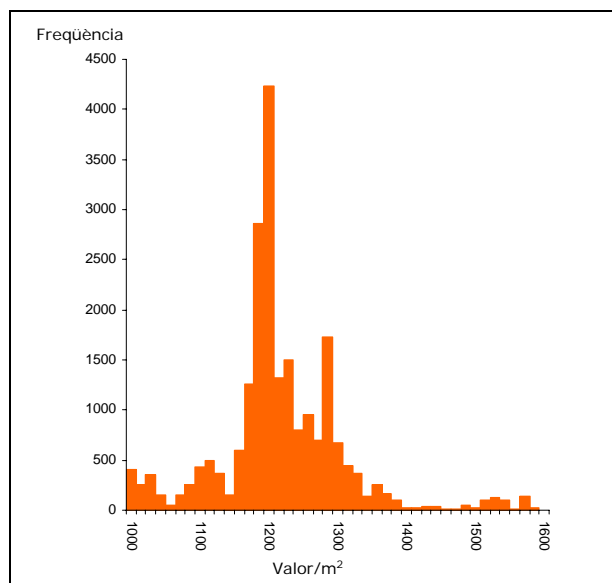


Figura 3.4.: Representació de la distribució *bootstrap*.

Les característiques principals d'aquest procés són les especificades a continuació:

- Cada remostra té exactament n elements igual que la mostra original.
- Les remostres poden tenir elements repetits de la mostra original, ja que per generar la remostra s'utilitza un mostreig amb reposició.
- De cada remostra es calcula l'estadístic desitjat, obtenint la distribució de l'estadístic en funció de les remostres.

La no verificació de les hipòtesis habituals (presència d'heterocedasticitat, no normalitat i/o correlació entre residus) comporta certes dificultats en la realització d'inferències; és per això que la metodologia *bootstrap* hi pot ser de convenient aplicació.

A partir del conjunt 'dades *bootstrap*', és a dir a partir dels N estadístics calculats a partir de N remostres, es podran calcular la mitjana *bootstrap* i els intervals de confiança *bootstrap* de la mateixa manera que es faria si es disposessin de dades en una mostra.

A tall d'exemple, si el que es vol calcular és la mitjana *bootstrap* i la seva desviació típica s'haurà d'aplicar:

$$\overline{X}_j^* = \sum_{k=1}^N x_{jk}^r$$

$$\overline{X}_{boot} = \frac{1}{B} \sum_{j=1}^B \overline{X}_j^*$$

$$SE_{boot} = \sqrt{\frac{1}{B-1} \sum_{j=1}^B (\overline{X}_j^* - \overline{X}_{boot})^2}$$

x_{jk}^r és la k -èssima dada generada amb repetició de la j -èssima remostra *bootstrap*.

\overline{X}_j^* és la mitjana de la remostra j .

3.4.3. El mètode de validació creuada

L'objectiu d'aquesta tècnica és l'estimació dels paràmetres i del model escollit.

El mètode de validació creuada (*Cross Validation*) es base en el supòsit d'homogeneïtat de la mateixa validació. Sota aquesta premissa es divideix la mostra inicial en dues submostres; amb la primera es realitza l'estimació i amb la segona es valida el model obtingut.

En aquest cas s'utilitzarà una variació d'aquest mètode denominada *one-to-leave*. El funcionament n'és el següent:

- Es realitza el procés d'estimació del model a partir de les n dades de la mostra inicial.
- S'elimina la dada i -èssima.
- S'efectua l'estimació del valor en la posició de la dada i -èssima a partir de les $n-1$ dades restants.
- Es calcula la diferència entre l'estimació i el valor real.
- Aquest procés es realitzarà per les n dades.

En completar aquest procés s'obté la variable aleatòria de l'error d'estimació per a la mostra original.

A partir dels resultats obtinguts es pot decidir modificar algun pas del procés d'estimació per intentar obtenir-ne millors solucions, o bé es pot considerar que el resultat és raonablement satisfactori.

Un exemple de validació creuada aplicat al mètode de veïns propers és el desenvolupat per Mullin i Sukthankar (2000).

3.4.4. Mètodes per estudiar la qualitat de les estimacions

Entre els estadístics utilitzats per a l'avaluació de la qualitat dels interpoladors es troben:

- l'error quadràtic mitjà, producte de la validació creuada (ECVc)

$$ECVc = \frac{\sum_{i=1}^N (Vh_i - Vhe_{-i})^2}{N-1}$$

Si s'hi afegeix una ponderació en funció del pes de la valoració o de la distància s'obtidran els estadístics següents:

$$EVC_1 = \sqrt{\sum_{i=1}^N (Vh_i - Vhe_{-i})^2 \cdot p_i}$$

sent-hi p_i el pes de la mostra i en la valoració del objecte. EVC_i són les inicials de estimador de validació creuada.

En aquest context pot ser interessant utilitzar estimadors amb comportament acotat entre 0 i 1.

En cas d'utilitzar una funció de la distancia en lloc de les ponderacions obtingudes a partir de l'estimació del valor l'EV podria ser de la següent forma:

$$EVC_3 = \sqrt{\sum_{i=1}^N (Vh_i - Vhe_{-i})^2 \cdot f(dist_i)}$$

$f(dist_i)$ representa una funció envers la distància.

Altres formes podrien ser:

$$EVC_3 = \sum_{i=1}^N abs(Vh_i - Vhe_{-i}) \cdot p_i$$

En aquest context pot ser interessant utilitzar estimadors amb comportament acotat entre 0 i 1.

$$EVC_4 = \sum_{i=1}^N abs\left(\frac{Vh_i - Vhe_{-i}}{Vh_i}\right) \cdot p_i$$

L'objectiu d'aquest procés és avaluar conjuntament les diferències entre els valors reals i els estimats, alhora que s'estudien els resultats que proporcionen les validacions creuades. A més de la selecció del mètode d'estimació, aquests procediments han de permetre establir la qualitat dels estimadors. Aquests processos seran estudiats i determinats en el punt 5.3.

3.4.5. Mètode de les permutacions

En el procés d'estimació presentat no tindria cap sentit que les mostres es repetissin dins d'un mateix remostrat *bootstrap* i per tant s'aplicarà una variació de la metodologia de remostreig clàssica anomenada 'mètode de les permutacions'.

El mètode de les permutacions (també coneguda com 'prova d'aleatorització' o 're-aleatorització') és una prova estadística en la qual s'obté una distribució de la referència calculant tots els valors possibles de l'estadístic a partir de permutar totes les possibles mostres observades. Aquest test va ser desenvolupat per R. A. Fisher l'any 1935.

El plantejament n'és el següent:

- Es parteix del supòsit que hi ha n mostres denominades: h_1, h_2, \dots, h_n .
- En lloc de realitzar l'estimació del valor en el punt (x_e, y_e) amb les n dades disponibles es farà amb només p dades, però aquesta estimació es durà a terme $V_{n,n-p}$ vegades.

La selecció de la p es farà en funció de dos objectius:

- obtenir un nombre de repeticions suficientment gran;
 - que el nombre de dades utilitzades en cada estimació sigui superior al nombre mínim establert per la legislació corresponent.
- Una vegada realitzades $V_{n,n-p}$ repeticions se n'hauran obtingut $V_{n,n-p}$ valors de taxació a partir dels quals s'estudiarà la funció de distribució i es decidirà quin és el valor presentat.

La utilització de les tècniques *bootstrap* serà de gran utilitat per a eliminar el biaix que pugui provocar una mostra desubicada, introduïda erròniament o senzillament amb unes característiques de localització diferenciades dintre de l'espai d'estudi mostral, però que donada la seva proximitat reben una ponderació important en aplicar qualsevol dels processos d'estimació espacials que es presenten.

3.5. Mètodes d'estimació i interpolació espacial

3.5.1. Introducció

El glossari estadístic del *U.S. National Institutes of Health (NIH)* dona la següent definició de correlació espacial.

Correlació espacial: Mesura de la tendència dels punts propers a tenir comportaments semblants entre ells (correlació positiva) o dissemblants (correlació negativa) en els seus estadístics.

En base a aquesta definició i seguint els plantejaments de Cressie (1991) es pot afirmar que per a totes aquelles dades que tenen associades un component espacial, s'hauria de suposar que les semblances entre dades properes són més evidents que entre aquelles dades que es troben més allunyades, i que en qualsevol d'aquestes situacions és important explotar i estudiar aquesta component espacial.

El nombre de situacions i àmbits en què es dur a terme un tractament de dades que tenen relació espacial és elevat. Alguns exemples en són: la meteorologia, la geologia, la contaminació atmosfèrica, la geografia, l'astronomia, el processat d'imatges, els experiments agrícoles, etc. Generalment, quan es considera que el component espacial és important en la modelització i l'anàlisi de la informació (per exemple quan es treballa amb dades recollides en diferents posicions espacials), caldrà una aproximació (estadística) diferent a les utilitzades tradicionalment.

Un dels primers estadístics que s'ha utilitzat per a aquest estudi i que ha permès estudiar la dependència espacial és el contrast de la I de Moran (Cressie, 1991). El problema de la I de Moran radica en què no permet obtenir una visió global dels processos d'interrelació espacial, ja que no ofereix l'estructura de l'autocorrelació espacial. A més, té un plantejament discret basat en regions o blocs.

L'estadístic de la I de Moran es calcula de la manera següent:

Si es parteix d'una variable en N regions

$$I = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad \text{i} \quad w_{ij} = \frac{1}{dist_{ij}^2}$$

i x representa la mesura d'una característica en N localitzacions diferents.

Aquest contrast té com a hipòtesi nul·la la no existència de dependència espacial, enfront a la hipòtesi alternativa de l'existència d'un esquema d'autocorrelació espacial.

Segons Cressie (1991) les característiques bàsiques que ha de complir un bon mètode de predicció han de ser les següents:

1. La funció de predicció s'ha d'ajustar a les dades en un nivell de precisió determinat.
2. Aquesta funció de predicció ha de ser contínua i suau en tots els punts.
3. Cada valor predit dependrà únicament d'un subconjunt de punts que estarà format per les dades més properes a la localització a predir. Aquesta restricció és necessària per a evitar que certes zones dominants emmascarin a d'altres i que es propagui algun error. D'aquesta manera, es garanteix que la funció generada serà estable, és a dir, un petit canvi en una dada no podria provocar una gran variació en la funció de predicció.
4. La funció de predicció no ha de presentar oscil·lacions fortes.
5. El mètode d'interpolació ha de ser independent de les mostres de dades, és a dir, el mètode d'interpolació s'ha de poder aplicar tant a disposicions de malla regular com a disposicions repartides de manera desordenada, i la qualitat dels estadístics obtinguts no ha de variar en funció de la disposició de les dades.

Existeix una àmplia varietat de tècniques que permeten la realització de prediccions amb components espacials. Aquestes tècniques poden ser classificades en dos grans grups: els mètodes directes i els mètodes basats en l'anàlisi estructural.

La principal característica dels mètodes directes és el fet de treballar directament sobre les distàncies al punt a predir o interpolar.

El principal avantatge de la majoria d'aquests mètodes es la facilitat de càlcul. El principal desavantatge d'aquests mètodes és que el radi de recerca no queda establert i, per tant, és fàcil d'introduir en el càlculs valors no relacionats.

A continuació s'exposen de manera exhaustiva i amb detall suficient per a la seva reproducció els mètodes de càlcul espacial i els conceptes utilitzats en el procés d'estimació del valor de taxació. Els mètodes escollits van ser:

- El mètode de les mitjanes mòbils.
- El mètode de la distància inversa.
- L'*Spline*.
- Les funcions radials.
- La metodologia *kriging* sota isotropia.
- La metodologia *kriging* sota anisotropia.

Al final d'aquest capítol es presenta una breu descripció d'altres mètodes de càlcul espacial existents.

3.5.2. Mètode de les mitjanes mòbils. La distància inversa.

El mètode d'interpolació de mitjanes mòbils ha estat amb tota probabilitat el mètode més utilitzat per a dur a terme interpolacions de valor amb relació espacial. Es basa en la idea intuïtiva -tret comú a la resta de mètodes directes- que les observacions més properes han de tenir més pes en la determinació del valor interpolat en un punt p_a . Es tracta d'un mètode exacte

i local que estima el valor de la variable V en la localització p_a a partir de l'expressió següent:

$$V(p_a) = \frac{\sum_{i=1}^n \lambda_i v(p_i)}{\sum_{i=1}^n \lambda_i} \quad \lambda_i = f(d_i) \quad \sum_{i=1}^n \lambda_i = 1$$

$V(p_a)$ és el valor interpolat de la variable V en el punt p_a . λ_i és el pes (o funció de ponderació) que s'assigna a la variable V en el punt de mostreig p_i . $V(p_i)$ és el valor de la variable en cada punt p_i de la mostra de tamany n . La restricció $\sum_{i=1}^n \lambda_i = 1$ assegura el no esbiaixament de l'estimació.

Les funcions de ponderació adopten la forma general:

$$\lambda_i = (d_i)^{-u}$$

Els valors més utilitzats per l'exponent són $u = 1$ i $u = 1.5$. i $u = 2$.

Un coeficient d'1 indica un decreixement menys accelerat de la contribució dels punts en funció de la distància d_{ij} .

Òbviament l'ús d'exponents elevats implica donar menys pes relatiu a punts més allunyats de p_a i viceversa.

El cas $u = 2$ es coneix com el mètode d'estimació per distància inversa.

Els valors per als diferents λ_i s'obtenen a partir de la ponderació següent:

$$\lambda_i = \frac{1}{\text{dist}_i^2} \bigg/ \sum_{i=1}^n \frac{1}{\text{dist}_i^2}$$

en què $\frac{1}{dist_i^2}$ és la inversa de la distància elevada al quadrat entre cada variable coneguda, i el valor a determinar en un punt i $\sum_{i=1}^n \frac{1}{dist_i^2}$ és la suma de l'invers de les distàncies elevades al quadrat de tots els punts existents.

Els principals desavantatges d'aquests mètodes són:

- El radi de recerca no queda establert. Això pot provocar la introducció de mostres no relacionades amb la localització a estimar.
- Les ponderacions poden ser extremadament influents sobre els punts propers, donat que el valor de la inversa de la distància és poc robust en aquest aspecte.
- Els valors interpolats són mesures ponderades que sempre prenen valors entre el màxim i el mínim dels punts mostrals, fet que en redueix l'eficàcia quan s'intenta modelitzar els casos de valor més alt o més baix.

3.5.3. Mètode de la curvatura mínima (*Splines*)

Com el seu nom indica, aquest mètode d'interpolació tracta de produir la superfície contínua de mínima curvatura sobre la qual es trobin localitzades totes les mostres utilitzades. Els mètodes *splines* són mètodes exactes, ja que el valor de les mostres es correspon amb el de les seves estimacions.

Aquest mètode va ser el primer que es va prendre en consideració inicialment per a buscar una solució al problema plantejat en aquesta tesi, però va ser descartat, ja que la forma d'obtenció del valor de taxació no aportava els pesos requerits, i per tant, podia plantejar problemes de compatibilitat amb l'ECO805/2003. No obstant això, es va decidir introduir-lo en aquesta tesi amb la intenció que quedés constància del seu estudi.

El mètode de la mínima curvatura està inspirat en el principi de les regles flexibles en què la pressió exercida en un punt donat produeix una reacció

(deformació) que afecta un ambient local determinat per certs punts fixos o fronterers.

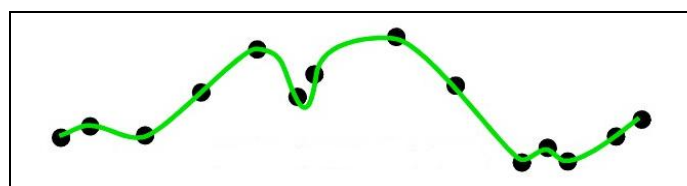


Figura 3.5: Spline en 2D.

La unió entre els diferents punts acabarà per definir una superfície contínua.

Matemàticament els *splines* són funcions polinomials de grau superior a tres derivables en els dos primers ordres (pendent i curvatura).

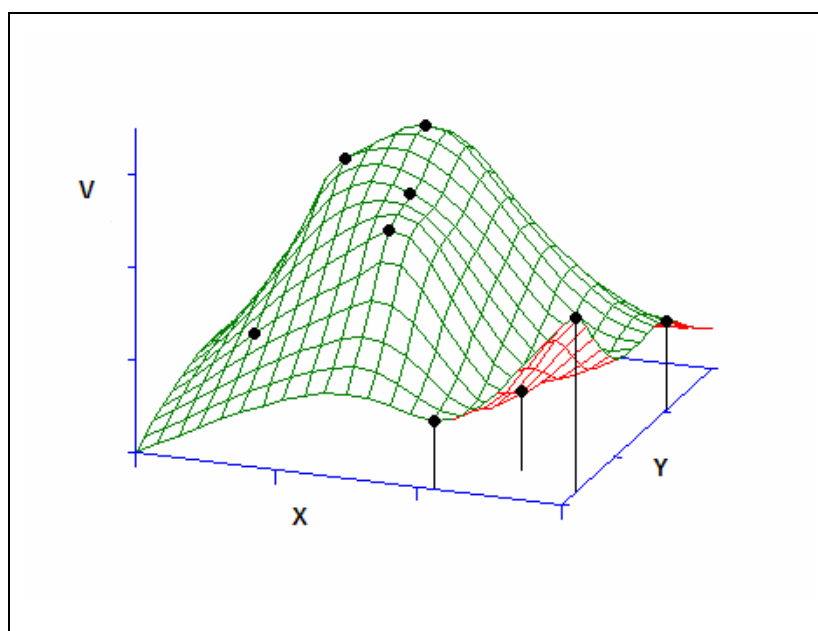


Figura 3.6.: Representació d'una interpolació spline.
Representació de punts influents.

Aquest mètode d'interpolació no obliga a conèixer de manera explícita la funció contínua en l'espai \mathbb{R}^3 que reproduceix la superfície real, ja que és possible obtenir les estimacions a partir de mètodes recursius.

Si p_i són les localitzacions on es troben valors mostrals i p_a és la localització on es pretén efectuar la predicció, la distància entre un punt mostral i qualsevol el punt a predir es definirà com:

$$\text{dist}(p_i, p_a) = \left((x_a - x_i)^2 + (y_a - y_i)^2 \right)^{1/2}$$

es defineix la funció:

$$V(p_a) = V(x_a, y_a) = \sum_{i=1}^n c_i \cdot E(t, t_i) + d_0 + d_1 x_a + d_2 y_a$$

en què $E(t, p_i) = \text{dist}(t, p_i) \cdot \log \text{dist}(t, p_i)$

Els coeficients c_1, c_2, \dots, c_n i d_1, d_2 i d_3 del polinomi es determinen a partir de les equacions:

$$(E + n\lambda I)c + Td = z \quad (\text{a})$$

sota la restricció $T'c = 0$ i

E és la matriu $n \times n$ d' $E(t_i, t_j)$.

I és la matriu identitat.

λ és el paràmetre de suavitzat.

c és el vector dels (c_1, \dots, c_n) .

d és el vector dels (d_1, d_2, d_3) .

i T és una matriu $n \times 3$ on a la i -èsima fila es troba $(1, x_i, y_i)$.

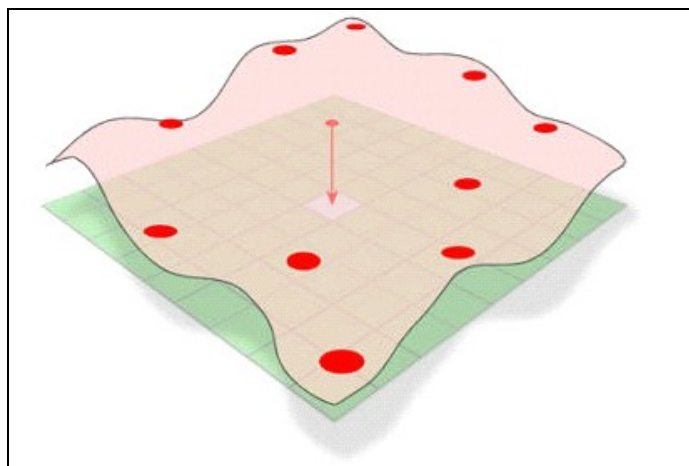


Figura 3.7.: Representació d'una interpolació *spline*.
Representació de punts influents.

3.5.4. Mètode dels radials

Les funcions de base radial comprenen un ampli grup de interpoladors exactes i locals que utilitzen una equació de base que depèn de la distància entre el punt interpolat i els punts mostrals propers. En general, el valor de la variable V en un punt no mostral p_a ve donada per l'expressió següent:

$$V(p_a) = \sum \lambda_i F(d_i)$$

en què $F(d_i)$ és una funció de base radial; d_i és la distància entre punts i λ_i els coeficients que hauran de ser calculats mitjançant la resolució d'un sistema lineal de n equacions, sent n el nombre de punts mostrals veïns que intervenen en la obtenció de $V(p_a)$.

A tall d'exemple, s'exposarà la funció de base radial de tipus multiquadràtic, desenvolupada per R. L. Hardy (1971).

$$F(d_i) = \sqrt{d_i^2 + r^2}$$

El paràmetre r és el factor de 'suavitzat'. L'obtenció d'aquest paràmetre s'hauria de basar en assaigs previs que permetessin trobar un resultat acceptable. Si el factor de suavitzat r és excessivament alt produirà una superfície massa suavitzada que probablement s'allunyi sensiblement de la geometria de la superfície generada pels valors reals.

3.5.5. Metodologia *kriging*. L'ús de l'anàlisi estructural.

Es pot definir l'estadística espacial com l'aplicació de la teoria de funcions aleatòries sobre variables numèriques distribuïdes en l'espai (Chauvet, 1994). Aquestes variables aleatòries presenten un comportament aleatori a escala local, però alhora un comportament estructural a gran escala. És a dir, si es treballa amb dues mesures realitzades en dos punts $p_i=(x_i,y_i)$ i $p_j=(x_j,y_j)$, s'obtidran dos valors $V(p_i)$ i $V(p_j)$ diferents però no independents i serà precisament el grau de correlació l'encarregat de reflectir la continuïtat espacial del fenomen.

L'anàlisi estructural consisteix en estimar i modelar el comportament a gran escala d'una funció a partir de l'adopció raonada de les hipòtesis més adequades en base a la seva variabilitat. En el cas que ens ocupa, s'intentarà modelar la correlació espacial de la variable d'interès valor total homogeneïtzat respecte la localització.

Totes les metodologies anteriorment presentades poden fer ús de les eines de l'anàlisi estructural per a resoldre el problema de la definició de l'espai de localització de l'espai mostral de treball, és a dir, que a partir de l'anàlisi estructural es pot arribar a definir la distància màxima d'interpolació. A part d'aquest ús potencial, cap de les metodologies de predicció del valor per mètodes directes utilitzen la informació que s'obté de l'anàlisi estructural. És per aquesta condició que els procediments de càlcul basats en les metodologies *kriging* s'anomenen 'mètodes basats en l'anàlisi estructural'.

El desavantatge principal de l'ús de l'anàlisi estructural és l'elevat nombre de dades que es necessiten per a aconseguir una anàlisi estructural vàlida.

L'anàlisi estructural té com base comuna l'estudi del comportament conjunt de les dades abans de l'anàlisi de les localitzacions puntuals.

Les metodologies basades en *kriging* han estat les tècniques de predicció espacial que més s'han popularitzat entre la comunitat científica en els últims anys.

3.5.5.1. Formalització

S'anomena variable regionalitzada a la variable distribuïda en l'espai de tal manera que presenta una estructura espacial de correlació.

Una variable espacial és una variable aleatòria v definida en un punt del espai (x,y) , és a dir $v=f(x,y)$.

Si a cada punt (x,y) que pertany a un domini en l'espai se li assigna una variable aleatòria $v(x,y)$, que en sentit general poden ser dependents, aleshores el conjunt de variables aleatòries espacialment distribuïdes $\{v(x,y), (x,y) \in \Omega\}$ serà una funció aleatòria $V(x,y)$.

Per tal de simplificar d'ara en endavant la nomenclatura s'escriurà p (de posició) en lloc de (x,y) .

▪ Funció de distribució i moments d'una funció aleatòria

Sigui $V(p)$ una funció aleatòria definida en \mathfrak{R}^3 , aleshores el vector aleatori $\{v(p_1), v(p_2), \dots, v(p_n)\}$ es caracteritzarà per la seva funció de distribució de probabilitat n-variada:

$$F_{v(p_1), v(p_2), \dots, v(p_n)}(v_1, v_2, \dots, v_n) = \Pr[V(p_1) \leq v_1, V(p_2) \leq v_2, \dots, V(p_n) \leq v_n]$$

El conjunt de totes les distribucions per a tot valor de p i per a qualsevol selecció de punts en \mathfrak{R}^3 constitueix la llei espacial de probabilitat de la funció aleatòria $V(p)$. Aquesta funció en la pràctica és impossible de determinar i només es pot esperar inferir els primers moments de la distribució de $V(p)$.

- Moments de la distribució de $V(p)$
- El moment de primer ordre de $V(p)$ és l'esperança matemàtica definida com:

$$m(p) = E[V(p)]$$

- Els moments de segon ordre considerats en geoestadística són:
 - La variància de $V(p)$

$$\sigma^2(p) = Var[V(p)] = E[\{V(p) - m(p)\}^2]$$

- La covariància de dos variables aleatòries $V(p_i)$ i $V(p_j)$ definida com:

$$C(p_i, p_j) = E[\{V(p_i) - m(p_i)\}\{V(p_j) - m(p_j)\}]$$

Aquesta funció es coneix també com funció de autocovariància

- El semivariograma $\gamma(p_i, p_j)$ que es defineix com:

$$2\gamma(p_i, p_j) = Var[V(p_i) - V(p_j)]$$

$$\gamma(p_i, p_j) = \frac{1}{2} Var[V(p_i) - V(p_j)]$$

També coneguda com funció de semivariància. A més, el variograma és $2\gamma(x_i, x_j)$ però amb freqüència s'utilitza el terme indistintament per a designar $\gamma(x_i, x_j)$.

▪ Funcions aleatòries estacionàries

Es diu que una funció aleatòria és estrictament estacionària si la seva funció de distribució *Ec.*(2.1) és invariant a qualsevol translació respecte a un vector h o el que és equivalent: la funció de distribució del vector aleatori $\{V(p_l)\}$,

$V(p_2), \dots, V(p_n)$ és idèntica a la del vector $\{V(p_1+h), V(p_2+h), \dots, V(p_n+h)\}$ per a qualsevol h .

Es diu que una funció aleatòria és estacionària de segon ordre si es compleix que:

- El seu valor esperat existeix i no depèn de p .
- Per a qualsevol parell de variables aleatòries $V(p)$ i $V(p+h)$, la seva covariància existeix i només depèn del vector de separació h

$$C(h) \equiv C(p+h, p) = E[V(p+h)V(p) - m^2]$$

- L'estacionarietat de la variància implica que la variància existeix, és finita i no depèn d' x , és a dir:

$$\sigma^2(x) = C(0) = Var[V(p)]$$

- Sota aquesta hipòtesi el semivariograma també és estacionari i es compleix que:

$$\gamma(h) \equiv \gamma(p+h, p) = \frac{1}{2} E[\{V(p+h) - V(p)\}^2]$$

A més, existeix una relació directa entre el semivariograma i la funció de covariància.

▪ Funcions aleatòries no estacionàries

Les funcions aleatòries no estacionàries són aquelles en què l'esperança depèn de la localització, és a dir del valor de p :

$$E[V(p)] = m(p)$$

$m(p)$ es coneix com funció de deriva o tendència.

Si la funció aleatòria $V(p)$ és considerada com la suma d'un component determinístic $m(p)$ i d'un residu $R(p)$ estacionari amb mitjana zero, és a dir:

$$V(p) = m(p) + R(p)$$

aleshores es pot apreciar que el semivariograma de $V(p)$ depèn de p

$$\gamma(p+h, p) = \gamma_R(h) + \frac{1}{2} \{m(p+h) - m(p)\}^2$$

En el cas en què la derivada sigui lineal $m(p) = m_0 + m_1 \cdot p$ el semivariograma no dependrà de p .

$$\gamma(p+h, p) = \gamma_R(h) + \frac{1}{2} \{m_1 \cdot h\}^2$$

No obstant això, creix amb el quadrat d' h , la qual cosa pot servir com a indicador per a la detecció de l'existència de tendència.

L'anàlisi prèvia de la informació amb l'objectiu d'obtenir un model que expliqui la continuïtat espacial s'aconsegueix a partir del càlcul del semivariograma experimental. Segons Chica (1987) aquesta és l'eina més important per tal de determinar les característiques de variabilitat i de correlació espacial.

3.5.5.2. El variograma experimental

L'anàlisi estructural es realitza en dues parts:

- El càlcul del semivariograma experimental.
- L'ajust del semivariograma experimental a un model teòric conegut.

Considerem dues mostres regionalitzades P_1 i P_2 de una mateixa variable en dues posicions diferents $P_1 = (x_1, y_1)$ i $P_2 = (x_2, y_2)$, i $V_1 = V(P_1)$ i $V_2 = V(P_2) = V(P_1+h)$ (per exemple el valor del sòl).

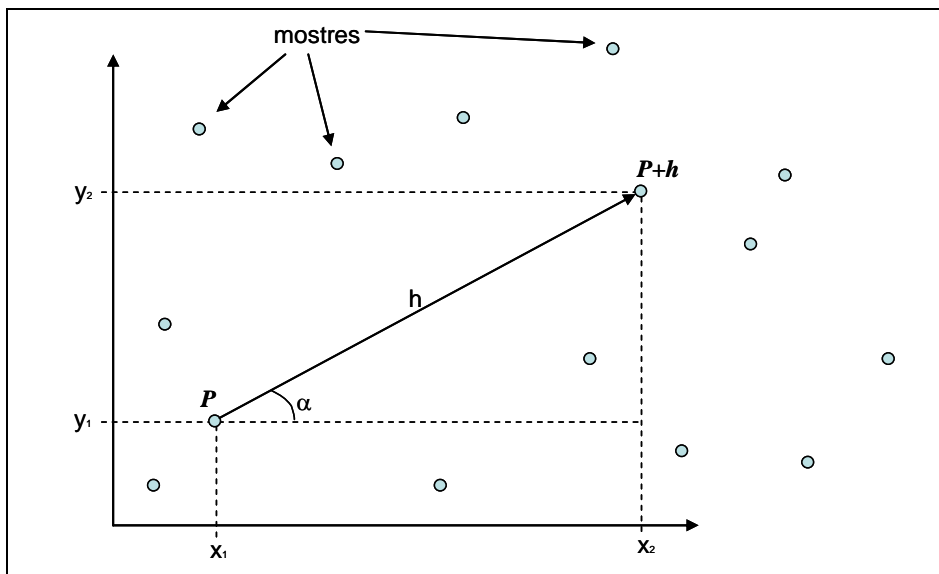


Figura 3.8.: Distribució espacial.

P_i denota una posició en dues dimensions, amb components (x_i, y_i) , i h un vector distància (mòdul i direcció) que separa els punts.

El nivell de dependència entre les dues variables regionalitzades, V_1 i V_2 , és representada pel variograma, $2\gamma(h)$, el qual es defineix com l'esperança matemàtica del quadrat de la diferència entre els valors de punts en l'espai, separats pel vector distància h , és a dir,

$$2\gamma(h) = Y\{[V(P_1)-V(P_1+h)]^2\} = Var[V(P_1)-V(P_1+h)] \quad ()$$

A través d'una mostra $v(P_i)$, $i=1, 2, \dots, n$, el variograma pot ser estimat per:

$$2\hat{\gamma}(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [v(p_i) - v(p_i + h)]^2$$

- $2\hat{\gamma}(h)$ és el variograma estimat;
- $N(h)$ - és el nombre de parells de valors mesurats, $v(\mathbf{P}_i)$ i $v(\mathbf{P}_i+h)$, separats per un vector distància \mathbf{h} .
- $v(\mathbf{P}_i)$ i $v(\mathbf{P}_i+h)$ - són valors de la i -èssima observació de la variable regionalitzada, recollits en els punts \mathbf{P}_i i \mathbf{P}_i+h ($i = 1, \dots, n$), separats pel vector \mathbf{h} .

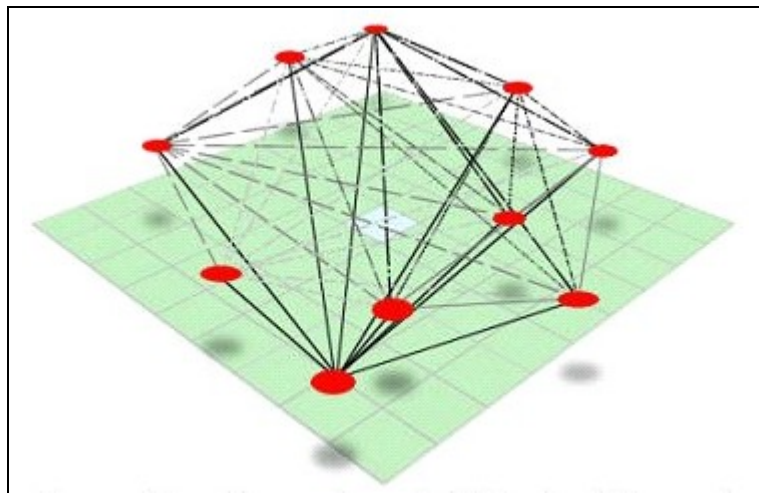


Figura 3.9.: Representació del càlcul del variograma.

Les seves principals característiques són:

- Com a mitjana mòbil que no és tenen problemes de robustesa.
- És un estimador no paramètric.
- Sota normalitat i regularitat en la distribució, l'estimador és de biaix mínim.

▪ Paràmetres del semivariograma

La figura 3.10. il·lustra un semivariograma experimental amb característiques molt pròximes al que seria una situació ideal. La seva forma representa la situació més típica, és a dir, que les diferències $\{V(\mathbf{P}_i) - V(\mathbf{P}_i + \mathbf{h})\}$ decreixin a mesura que ho fa \mathbf{h} i que alhora la distància que els separa també vagi decreixent, ja que és d'esperar que les observacions més pròximes geogràficament tinguin un comportament més semblant entre si que aquelles separades per més distància. Per tant, és d'esperar que si existeix correlació espacial $\gamma(\mathbf{h})$ augmenti a mesura que augmenta la distància \mathbf{h} .

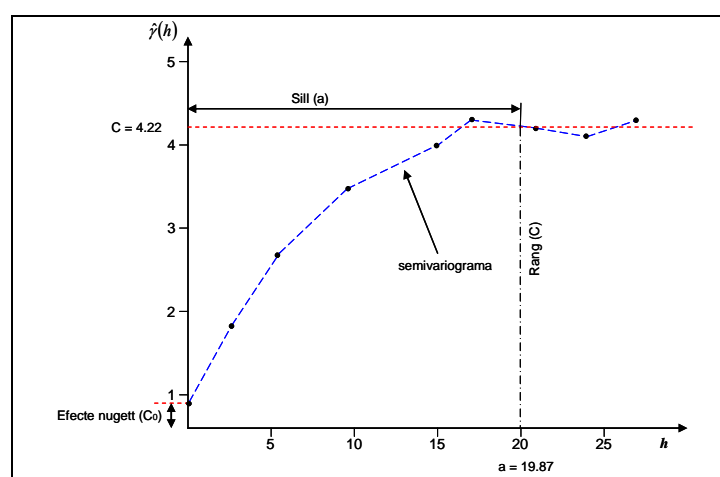


Figura 3.10.: Exemple de variograma.

Analitzant la figura 3.10. es poden apreciar els paràmetres que defineixen el variograma:

- *Sill (a)*: distància dintre de la qual les mostres presenten correlació espacial. Quan la corba del variograma deixa de tenir pendent, aleshores indica que les dades deixen de tenir correlació espacial i passen a ser independents.
- Rang (*C*): és el valor del semivariograma corresponent al *sill (a)*. A partir d'aquest punt en endavant es considerarà que ja no existeix dependència espacial entre les mostres, perquè la variància de la diferència entre parelles de mostres ($Var[V(\mathbf{P}) - V(\mathbf{P}+\mathbf{h})]$) es fa invariant amb la distància.

- Efecte *nugget* (C_0): Representa dues fonts de variabilitat que, amb freqüència, actuen alhora. La primera deu a la variabilitat a una distància més petita que el menor h considerat. La segona causa d'aquest efecte solen ser els errors experimentals. La interpretació que es faci d'un variograma dependrà de la grandària de l'efecte *nugget*, ja que la diferència entre el *sill* (si existeix) i l'efecte llavor representa la proporció de la variància mostral que pot modelitzar-se com variabilitat espacial. per definició, $\gamma(\mathbf{0})=0$ (en referència a l'equació 2.15). Tot i això, a la pràctica, a mesura que h tendeix cap a 0 (zero), $\gamma(h)$ s'aproxima cap a un valor positiu anomenat efecte *nugget* (C_0). El valor de C_0 revela la discontinuïtat del semivariograma per a distàncies menors que la menor distància entre les mostres. Part d'aquesta discontinuïtat pot ser també deguda a errors de mesura (Isaaks i Srivastava, 1989), però és impossible de quantificar si la major contribució prové dels errors de mesura o de la variabilitat a petita escala no captada pel mostreig.
- Contribució (C_1): és la diferència entre el nivell (C) i l'efecte *nugget* (C_0).

Els mètodes basats en *kriging* requereixen una especificació completa, tant pel que fa a la funció utilitzada com pel que fa als valors dels paràmetres de la dependència espacial del procés espacial en termes de covariància o de semivariograma.

Típicament, el model del semivariograma no és conegut amb anterioritat i caldrà estimar-lo, ja sigui visualment o mitjançant algun mètode d'estimació. En el nostre cas i donat que els càlculs es realitzaran dins d'un procés automatitzat serà imprescindible la utilització d'un mètode d'estimació no lineal per a obtenir els paràmetres necessaris. Per tant, caldrà efectuar una anàlisi preliminar per a obtenir el variograma empíric i els seus paràmetres abans de poder realitzar el procediment d'estimació *kriging*.

Tipus de variogrames

Variograma *Cloud*

El variograma *Cloud* ens permet escollir un valor inicial pel *lag*. S'aconsegueix a partir de representar el valor de γ per a totes les parelles de punts.

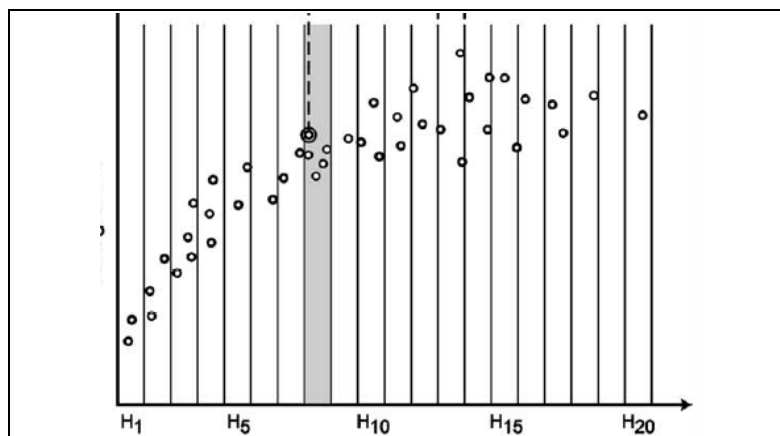


Figura 3.11: Variograma Cloud.

També permet observar la dispersió al voltant de γ^* .

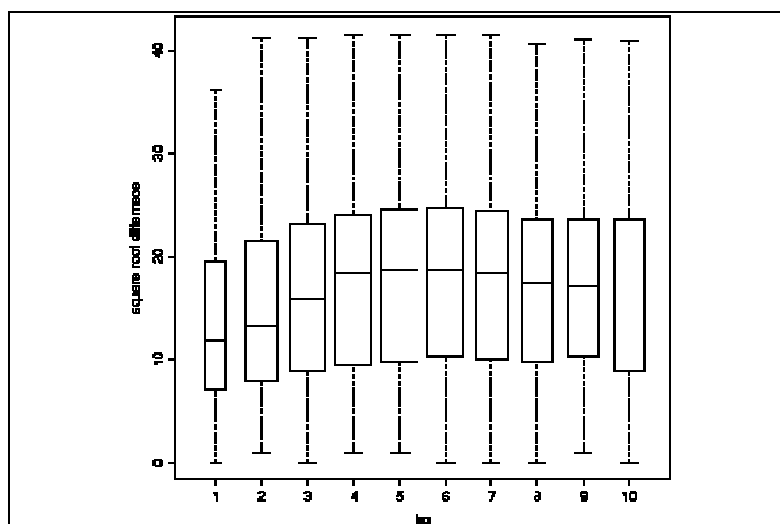


Figura 3.12.: Representació en diagrama de caixes del variograma Cloud.

Variograma direccional

- S'escull una direcció θ a partir de la qual es calcularà el variograma. Aquest tipus de variograma es fa servir quan es coneix la direccionalitat de les dades o quan es decideix fer una anàlisi per parts, recorrent tots els espais angulars a partir de la utilització de diferents variogrames.

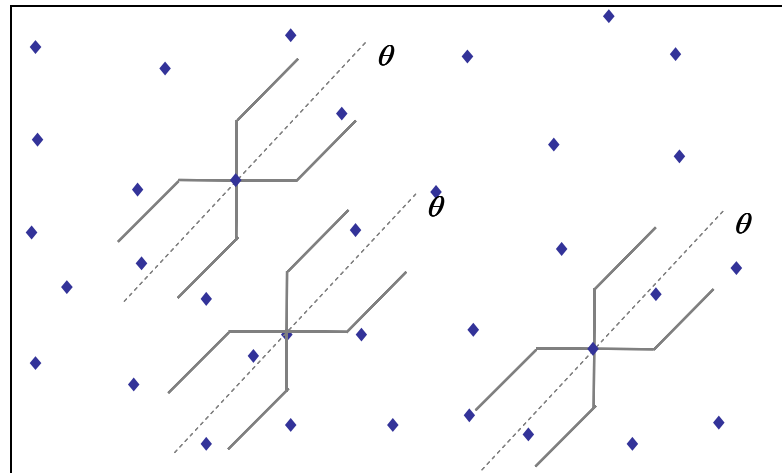


Figura 3.13.: Variograma direccional.

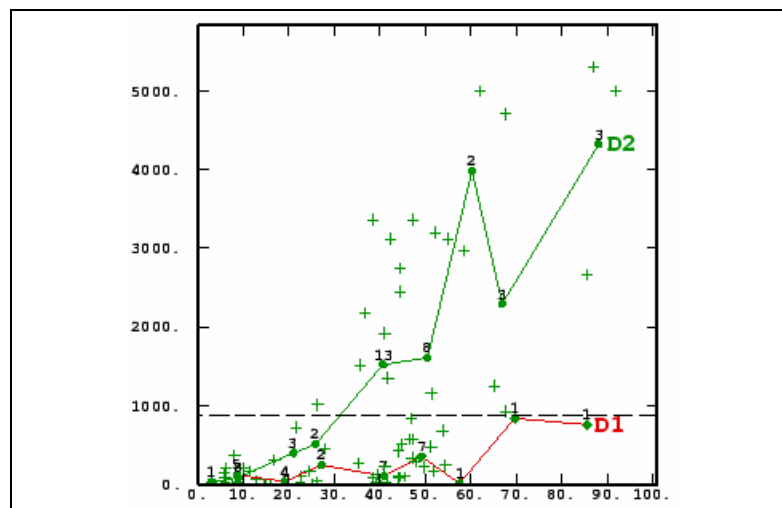


Figura 3.14.: Exemple de dos variogrames realitzats en direccions diferents.

Variograma omnidireccional

- No depèn de la direcció.
- S'obté en escollir la tolerància angular $\Delta\theta$ de tal manera que les direccions $\theta - \Delta\theta$ i $\theta + \Delta\theta$ siguin oposades i perpendiculars a la direcció θ .
- Es pot plantejar com la mitja de tots els variogrames experimentals en totes les direccions.

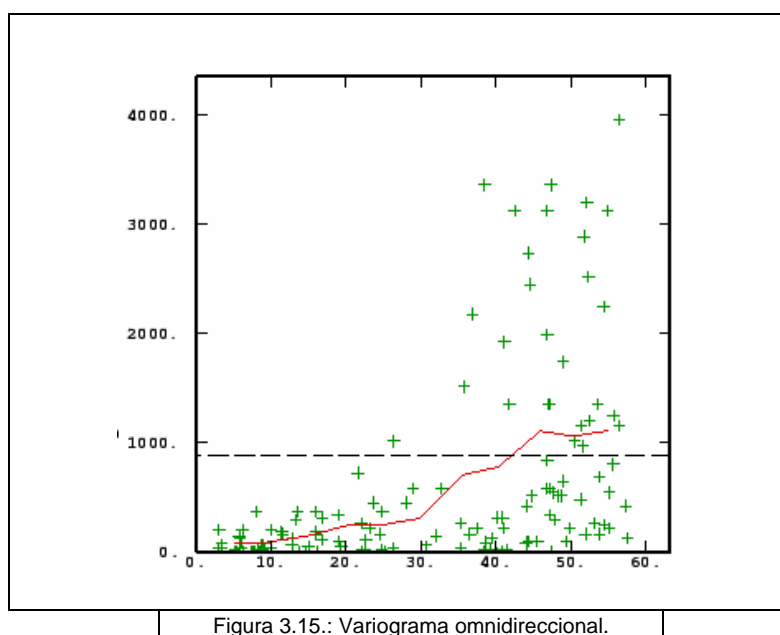


Figura 3.15.: Variograma omnidireccional.

3.5.5.3. Mètode d'obtenció del variograma experimental

El mètode d'obtenció que aquí es representa és general i tant pot ser utilitzat per a calcular el variograma direccional com el variograma omnidireccional.

1. S'escull una direcció θ .
2. S'escull una distància o *lag* h .
3. Es calcula γ^* pels valors de $h, 2h, 3h, \dots, nh$.
4. Es fa la gràfica de γ^* respecte als valors d' $h, 2h, 3h, \dots, nh$.

Per obtenir el valor de $l'h$ i del nombre de *lags* es poden utilitzar les regles següents:

- L' h serà la distància mitja entre punts en la direcció θ .
- El nombre de *lags* d'un domini D s'escollirà en funció $n < |D|/2h$.
- El semivariograma mostrat s'ha de considerar únicament per a petites distàncies i és aconsellable estimar-lo per a valors d' $|h|$ menors que la meitat de distància màxima $|h| < d_{max}/2$.
- El nombre de parelles per interval per a considerar el semivariograma estable és d'entre 30 i 50.
- Cal un mínim de 10 intervals per a poder estimar amb precisió el rang i el *sill*.

Si els valors es troben repartits irregularment, pot donar-se que no existeixin valors per una direcció i un *lag* determinat; és per això que per aquesta situació es relaxarà el *lag* h a un interval $[h - \Delta h, h + \Delta h]$. Si es pretén evitar que existeixin espais entre intervals o que els intervals s'encavalquin, el Δh ha de ser igual a $2h$.

Un procediment semblant a l'aplicat per aconseguir els *lags* s'ha de realitzar per aconseguir dividir l'espai de treball angularment. Si s'escull θ , com l'angle d'aplicació i es relaxarà l'angle θ a un interval $[\theta - \Delta\theta, \theta + \Delta\theta]$.

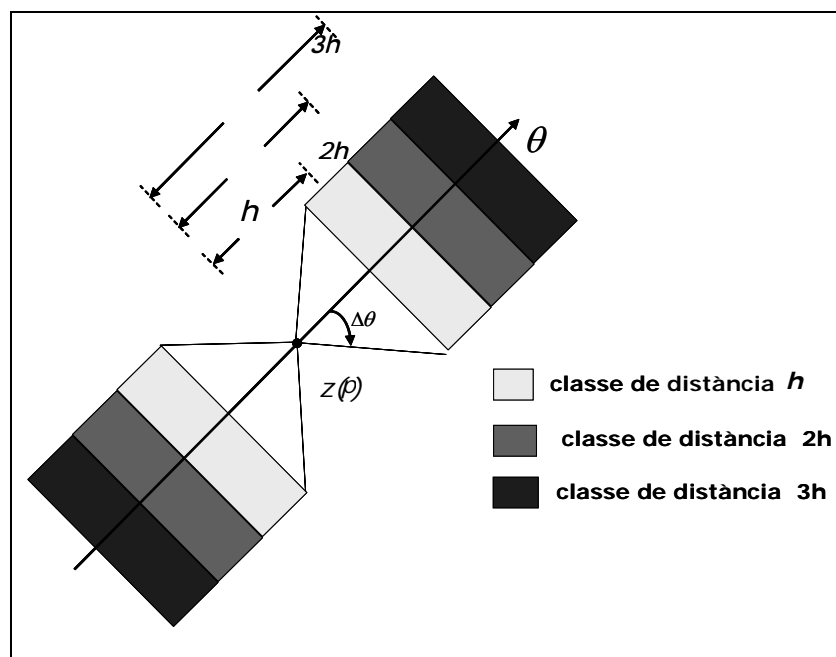


Figura 3.16.: Representació de la forma de càlcul del variograma.

3.5.5.4. Selecció del model teòric

Una vegada efectuats els variogrames experimentals o empírics, cal analitzar-ne les característiques i les possibles situacions d'anisotropia. Per poder dur a terme aquest procés cal trobar un model de variograma teòric que modelitzi correctament els variogrames empírics d'interès.

La necessitat d'utilitzar un model teòric de variograma en lloc del model empíric és degut a què no és possible utilitzar directament els estimadors del variograma experimental per a la estimació de la funció que defineix les relacions espacials entre les mostres. L'impossibilitat d'utilitzar el variograma empíric és degut a què es desconeixen els valors en distàncies o direccions intermitges, i per tant no es possible trobar el valor per a les posicions a predir. A més la interpolació entre els punts del variograma experimental no és suficient per a garantir l'existència i la unicitat de la solució del sistema de *kriging* (Matheron, 1970), donat que els mètodes d'interpolació no són condicionalment semidefinits negatius i això pot causar estimacions negatives de l'error en la mesura quadràtica de predicció.

El primer pas a fer és seleccionar d'entre els variogrames teòrics existents aquell amb les propietats més adients per al procés en qüestió.

Els variogrames teòrics més habituals són:

Model 'efecte llavor pur'

- Aquest model representa un fenomen completament aleatori, en el qual no hi ha correlació espacial.
- No importa quant a prop es trobin els valors de les variables, sempre seran no correlacionats.
- El seu variograma és:

$$\gamma(h) = \begin{cases} 0 & \text{si } |h|=0 \\ s & \text{si } |h|\neq 0 \end{cases}$$

i gràficament pren la forma següent:

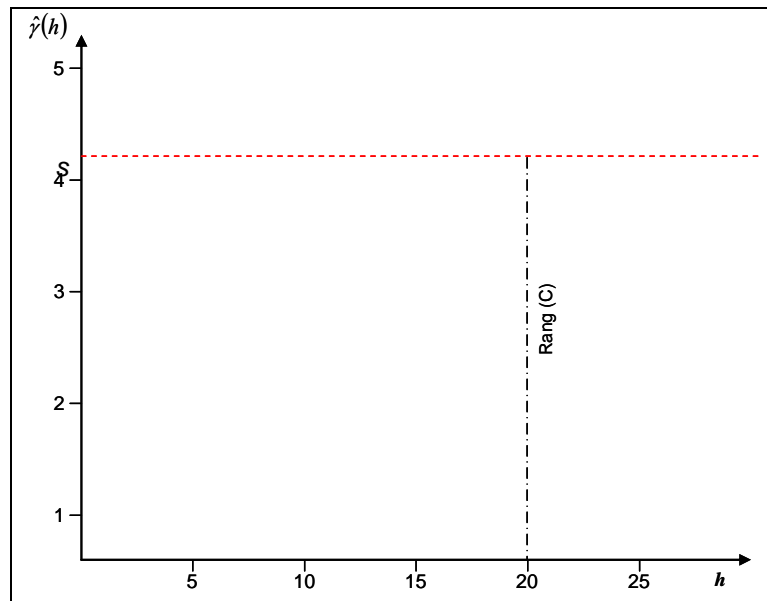


Figura 3.17.: Variograma del model efecte llavor pur.

Model exponencial

- El sill s s'assoleix de forma asimptòtica.
- Es pren com a rang el valor a .
- El rang experimental és igual a $3a$.
- Té un comportament lineal en l'origen.
- El pendent és igual a $3s/a$.
- Representa fenòmens continus però no diferenciables.
- El seu variograma és:

$$\gamma(h) = s \left(1 - \exp\left(-\frac{|h|}{a}\right) \right)$$

i gràficament es representa de la forma següent:

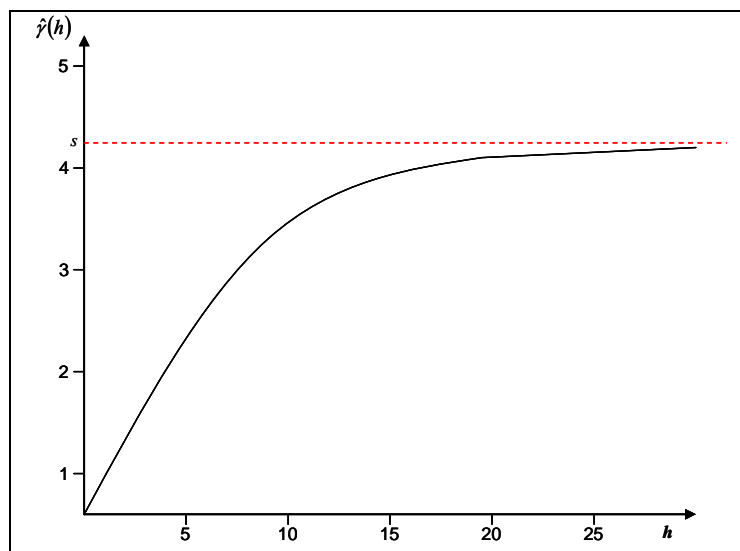


Figura 3.18.: Variograma del model exponencial.

Model gaussià

- El sill s s'assoleix de forma asintòtica.
- Es pren com a rang el valor a .
- El rang experimental és igual a $\sqrt{3a}$.
- Té un comportament quadràtic en l'origen.
- Representa fenòmens continus infinitament diferenciables (summament continus).
- El seu variograma és:

$$\gamma(h) = s \left(1 - \exp\left(-\frac{|h|^2}{a^2}\right) \right)$$

Gràficament és representat a la figura 3.19.

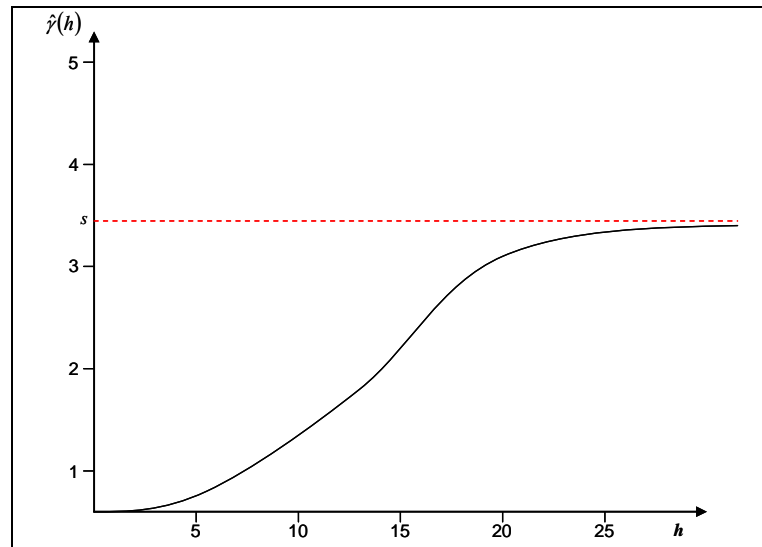


Figura 3.19.: Variograma del model gaussià.

Model potencial

- s se denomina factor d'escala.
- El valor de p es troba en l'interval $[0,2)$.
- El comportament en l'origen depèn del valor de p .
- Representa fenòmens no estacionaris.
- El seu variograma és:

$$\gamma(h) = s|h|^p$$

Model esfèric

- El valor del rang és a i el valor del sill s .
- Té un comportament lineal en l'origen.
- El pendent és igual a $1.5s/a$.
- Representa fenòmens continus però no diferenciables.

- El seu variograma és:

$$\gamma(h) = \begin{cases} s \left(\frac{3}{2} \frac{|h|}{a} - \frac{1}{2} \frac{|h|^3}{a^3} \right) & \text{si } |h| \leq a \\ s & \text{si } |h| > a \end{cases}$$

- La representació gràfica la podem veure a la plana següent.

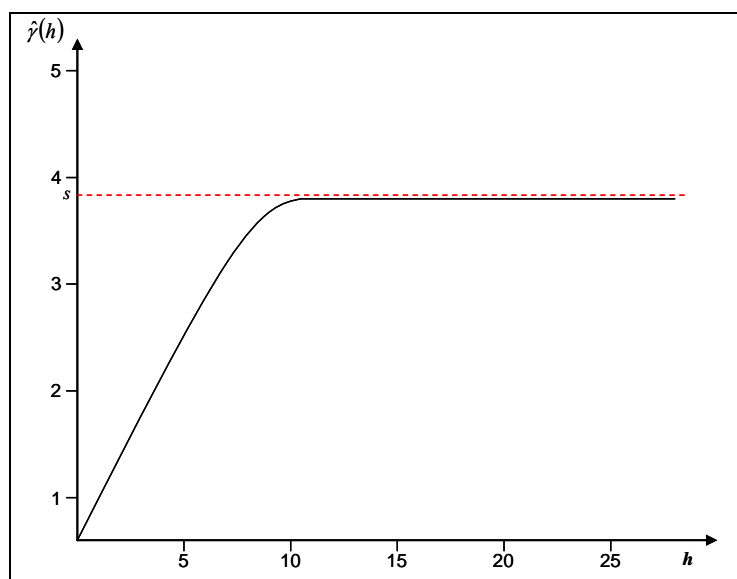


Figura 3.20.: Variograma del model esfèric.

Les propietats de tot variograma teòric són:

- $\gamma(0) = 0$.
- $\gamma(-h) = \gamma(h)$.

- Tot variograma és una funció definida positiva condicional.

Per a qualsevol n , sigui p_1, p_2, \dots, p_n punts en l'espai i $\lambda_1, \lambda_2, \dots, \lambda_n$

qualssevol valors tals que $\sum_{i=1}^n \lambda_i = 1$, aleshores es dona que:

$$-\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i - x_j) \geq 0$$

- Aquesta propietat permet calcular de manera consistent la variància de combinacions lineals de funcions aleatòries.

$$\text{var}(\lambda^\alpha V_\alpha) = -\lambda^\alpha \lambda^\beta \gamma_{\alpha\beta}$$

- Per funcions aleatòries estacionàries es té que

$$\gamma(h) = C(0) - C(h)$$

C és la funció de covariància.

- Si γ és el variograma de una funció aleatòria estacionària o intrínseca, aleshores

$$\lim_{h \rightarrow \infty} \frac{\gamma(h)}{|h|^2} = 0$$

En particular per a h suficientment gran existeix una constant c tal que

$$\gamma(h) \leq c|h|^2.$$

- Si $\gamma_1(h), \gamma_2(h), \gamma_3(h), \dots, \gamma_N(h)$ són models de variograma i $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N$ són valors positius, aleshores

$$\gamma(h) = \sum_{i=1}^n \alpha_i \gamma_i(h)$$

permet modelar les estructures aniuades i també l'anisotropia zonal.

- La funció del variograma teòric és:

$$\gamma(h) = \frac{1}{2} E [V(x) - V(x+h)]^2$$

- Permet analitzar el comportament espacial d'una propietat o variable sobre una zona determinada.
- Si $V(P)$ és estacionària o intrínseca, aleshores

$$Var[V(\mathbf{P}) - V(\mathbf{P}+\mathbf{h})] = E[V(\mathbf{P}) - V(\mathbf{P}+\mathbf{h})]^2$$

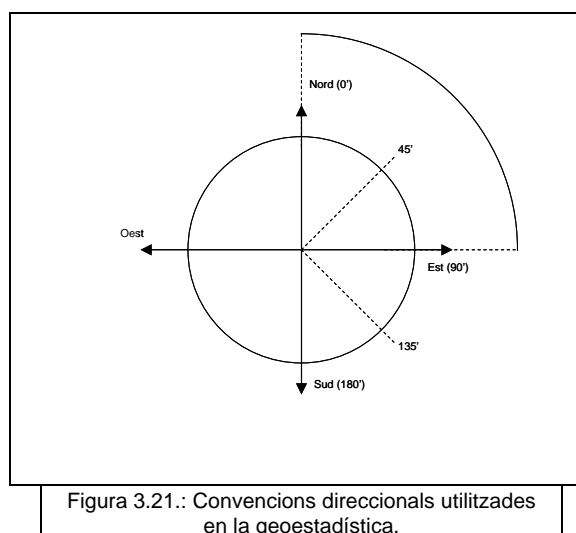
que representa el valor mitjà al quadrat dels valors mesurats en dos punts separats per una distància $|h|$.

- γ és independent de la localització.
- γ depèn del mòdul i de la direcció del vector h .

3.5.5.5. Isotropia i anisotropia

La isotropia és la propietat que indica que el comportament de la variable espacial té el mateix comportament en totes les direccions. Per tant, l'anisotropia és la propietat de tenir comportaments espacials diferents en direccions diferents.

La anisotropia pot ser fàcilment constatada mitjançant l'observació dels semivariogrames direccionals. Les convencions direccionals utilitzades en la geoestadística es mostren en la figura 3.21.



Per tant, s'estudiaran, per convenció, els semivariogrames obtinguts per a les direccions de 0° , 45° , 90° i 135° , il·lustrats a la figura 4.4.4. En aquest cas es dona una similitud força elevada entre ells. Aquesta és la representació de un cas simple i poc freqüent en què la distribució espacial del fenomen és denominada isotròpica. En aquest cas, un únic model és suficient per a descriure la variabilitat espacial del fenomen en estudi.

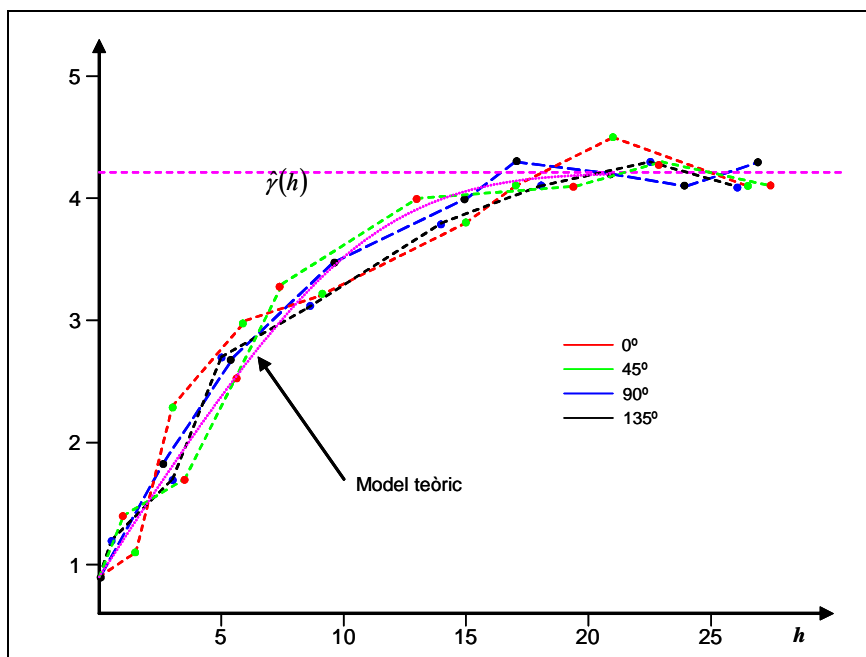


Figura 3.22.: Representació gràfica de semivariogrames isotròpics.

D'altra banda, si els semivariogrames no són iguals en totes les direccions, la distribució és denominada anisotròpica. Si l'anisotropia és observada i es reflexa pel mateix rang (C) amb diferent *sills* (a) del mateix model, aleshores es denomina anisotropia geomètrica.

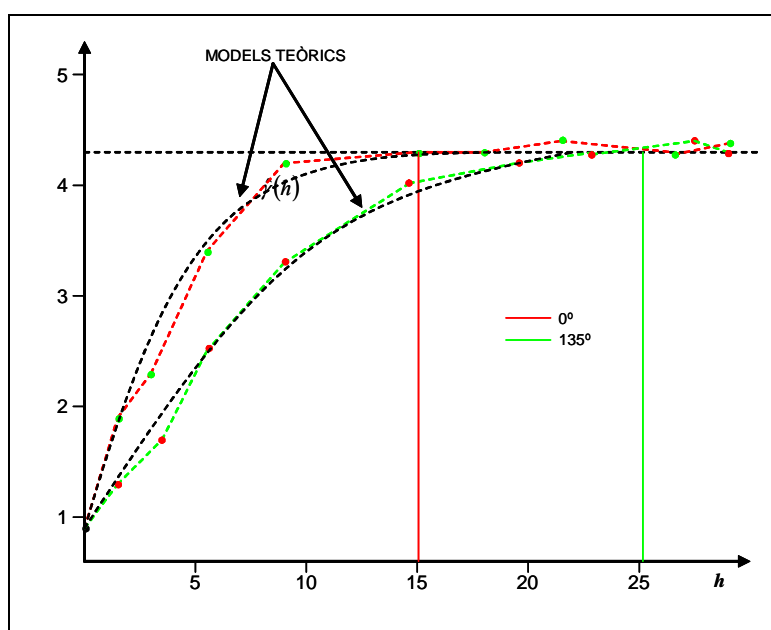


Figura 3.23.: Representació gràfica de anisotropia geomètrica.

En analitzar el semivariograma il·lustrat en la figura 3.23 s'observa que els punts units per línies de colors són els semivariogrames experimentals en dues direccions ortogonals. El semivariograma que primer assoleix la situació d'independència espacial és el que estudia la direcció 0° i el semivariograma amb major *sill* (línia verda) és el que explica el comportament en la direcció de 135° . Es defineix com la direcció de màxima dependència aquella en què el seu variograma té major *sill*. En el cas de l'exemple presentat el variograma platejat a partir de l'angle de 135° és la direcció de màxima dependència. Mentre el variograma generat a partir de les parelles relacionades amb l'angle 0° deixen de ser dependents en $k=15$, les dades a partir de les quals s'ha generat el variograma per a l'angle 135° deixen de ser dependents quan la k pren valor 25.

L'eix de màxima independència i el seu eix ortogonal defineixen una forma el·líptica (figura 3.24) que pot ser transformada a una forma circular a partir d'un procés de transformació dels eixos.

Las línies negres d'ambdues direccions són els models teòrics d'ajust dels semivariogrames experimentals.

Una manera molt directa de visualitzar i calcular els paràmetres (factor i angle) de l'anisotropia geomètrica és a través de l'esbós gràfic d'una el·lipse, calculada a través dels *sills* obtinguts en diferents direccions (figura 3.23). Les convencions que segueix aquest gràfic són les adoptades per Deutsch i Journel (1992). Per a l'eix major de l'el·lipse, denominat 'direcció de màxima continuïtat', s'aplica el major *sill* (a_1). L'angle de la direcció de màxima continuïtat queda definit a partir de la direcció Nord i en el sentit horari. El seu valor correspon a la direcció de major *sill*. L'eix menor defineix el *sill* (a_2) en la direcció de menor continuïtat, essent aquest ortogonal a la direcció principal (Deutsch i Journel, 1992).

El gràfic direccional dels rangs forma una el·lipse en què l'eix menor és el rang de la direcció de més ràpida variació; i l'eix major és el de la variabilitat més lenta. La relació $\lambda=a_1/a_2$ és la mesura d'anisotropia.

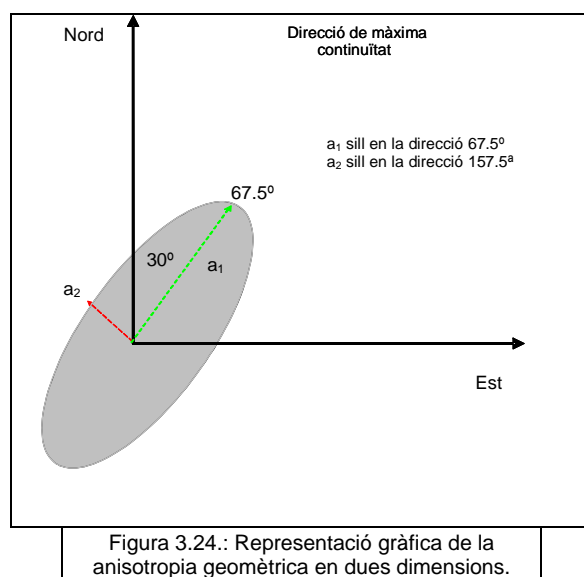


Figura 3.24.: Representació gràfica de la anisotropia geomètrica en dues dimensions.

El factor d'anisotropia geomètrica ve definit com la raó entre l'abast (*sill*) en la direcció de menor continuïtat (a_2) i l'abast en la direcció de major continuïtat (a_1). En aquest cas, el factor d'anisotropia geomètrica és sempre menor que la unitat, i l'angle d'anisotropia és igual a l'angle de la direcció de màxima continuïtat.

Existeix encara un altre tipus d'anisotropia. Aquesta es produeix quan els semivariogrames presenten els mateixos *sills* (a) i *rangs* (C) diferents. En aquest cas, l'anisotropia és denominada 'zonal'. L'anisotropia zonal és un cas poc freqüent, present no obstant en els fenòmens naturals. El més comú és trobar combinacions de l'anisotropia zonal i geomètrica, denominada anisotropia combinada.

Si s'observa el semivariograma presentat en la figura 3.25. els punts units amb línies discontinües de colors corresponen als semivariogrames experimentals en dues direccions ortogonals. El semivariograma amb major rang (vermell) es refereix a la direcció de 0° i el semivariograma amb menor rang (verd) es refereix a la direcció 135° . Els models d'ajust als semivariogrames estan representats per línies negres.

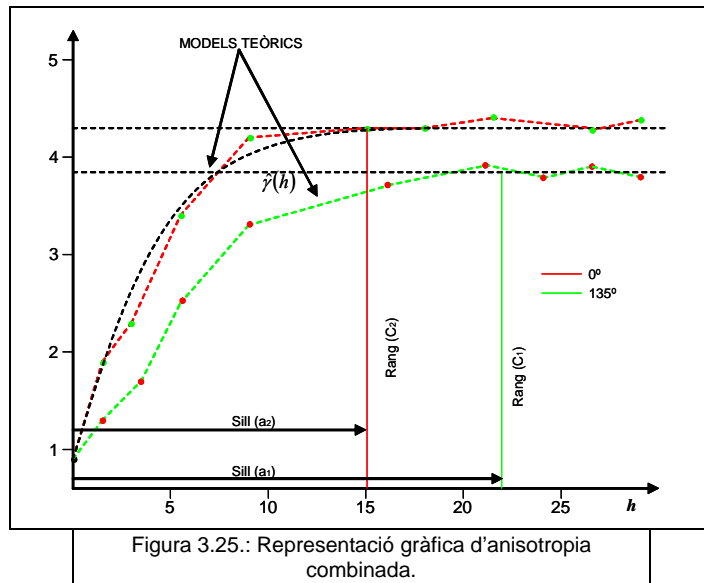


Figura 3.25.: Representació gràfica d'anisotropia combinada.

Segons Isaaks i Srivastava (1989), citats per Deutsch i Journel (1992, p. 25), l'anisotropia zonal es pot considerar com un cas particular de l'anisotropia geomètrica en suposar-se un factor d'anisotropia molt gran. Sota aquesta condició, el sill implícit en la direcció de menor continuïtat és molt gran. L'estructura del semivariograma és llavors afegida solament per a la direcció de major continuïtat.

3.5.5.6. Modelització del variograma experimental

Els mètodes més utilitzats habitualment per a la modelització del variograma experimental són:

- Mínims Quadrats Ordinaris (MCO).
- Mínims Quadrats Generalitzats (MCG).
- Mínims Quadrats Ponderats.

Tot i que, cal puntualitzar aquí que els més precisos pel que fa a la forma no lineal dels variogrames són els mètodes iteratius com Gauss-Newton o Newton-Rapson.

Criteri de selecció del model de variograma teòric

Normalment el criteri de selecció del model es basa en el criteri d'informació d'Akaike (variació de l'estimador de màxima versemblança (Johnston, J., 1987) que penalitza l'entrada de paràmetres en el model), ja que els models del variograma teòric poden tenir un nombre diferent de paràmetres a estimar. En el cas que aquí es presenta tots els models utilitzats tenen el mateix nombre de paràmetres.

El Criteri d'informació d'Akaike (*AIC*), que es defineix com:

$$AIC = -2 \ln(\text{máx. Versemblança}) + 2 (\text{nombre de paràmetres})$$

i es pot estimar a partir de

$$(AIC)^* = \left\{ n \ln\left(\frac{2\pi}{n}\right) + n + 2 \right\} + n \ln(R) + 2p$$

com que $n \ln\left(\frac{2\pi}{n}\right) + n + 2$ és constant per a tots els models provats, tan sols es pot emprar la part final del càlcul $n \ln(R) + 2p$.

n és el nombre de valors estimats $\{\gamma^*(h_i), i = 1, \dots, n\}$ del variograma mostral; R és la suma residual dels quadrats de les diferències entre els valors experimentals $\gamma^*(h_i)$ i els del model ajustat $\gamma(h_i)$, és a dir:

$$R = \sum_{i=1}^n (\gamma(h_i) - \gamma^*(h_i))^2$$

i p és el nombre de paràmetres del model de variograma $\gamma(h_i)$.

El model que presenti un menor *AIC* serà el que ajusti millor. L'*AIC* va ser seleccionat com a mètode d'estimació de l'error, davant de la possibilitat

d'utilitzar models teòrics amb diferent nombre de paràmetres. Però pels casos presentats el mètode de la màxima versemblança s'acaba reduint a un suma dels errors quadràtics, ja que el tots els models escollits tenen el mateix nombre de paràmetres.

3.5.6. El *kriging*

El terme *kriging* es deriva del seu 'creador espiritual' Daniel G. Krige, qui va ser el pioner en introduir l'ús de mitjanes mòbils per evitar la superestimació sistemàtica en les reserves de mineria. Però, materialment, va ser Matheron als anys seixanta qui va formalitzar i generalitzar en l'àmbit matemàtic el conjunt de tècniques desenvolupades per Krige. Aquestes tècniques utilitzaven la correlació espacial com a base per a la realització de les prediccions.

Les tècniques basades en *kriging* més àmpliament utilitzades són: el *kriging* simple, el *kriging* ordinari, el *kriging* universal i el *co-kriging*.

La diferència entre els mètodes *kriging* i la resta de mètodes d'interpolació radica en la manera com s'atribueixen els pesos a les diferents mostres. En el cas d'interpolació lineal simple, per exemple, els pesos són tots iguals a $1/N$ (N = nombre de mostres); en la interpolació basat en l'invers del quadrat de les distàncies, els pesos són definits com l'invers del quadrat de la distància que separa el valor interpolat dels valors observats. En els mètode basats en *kriging* els pesos són determinats a partir d'una anàlisi espacial, que es basa en una eina de suport coneguda com semivariograma experimental. Aquesta eina permet representar quantitativament la variació d'un fenomen regionalitzat en l'espai (Huijbregts, 1975).

Per tant, el procés de predicció espacial realitzat a partir de la metodologia *kriging* es durà a terme, bàsicament, en dos passos.

1. Realització dels models de covariància o del semivariograma del procés espacial, que permetrà escollir tant els paràmetres matemàtics com els valors associats als paràmetres.

2. L'ús dels resultats obtinguts a partir del variograma per a resoldre el sistema *kriging* per a un conjunt específic de punts. El resultat d'aquest procés suposa l'obtenció dels valors predits i els errors estàndards associats.

L'objectiu principal del procés *kriging* és arribar a estimar el valor d'un fenomen en una posició $V_0 = Z(p_0)$ a partir del conjunt d'observacions disponibles $Z(p_1), \dots, Z(p_n)$.

Aquesta predicció prendrà la forma següent:

$$Z^*(p_0) = \sum_{i=1}^N \lambda(p_i) Z(p_i)$$

és a dir, l'estimació de $Z(p_0)$ s'aconseguirà a partir de la combinació lineal de les observacions disponibles.

El criteri de selecció dels pesos $\lambda_i(p)$ es basa en la minimització de:

$$\text{var}[Z(p) - Z^*(p)]$$

Donat que les dades observades són una realització parcial del procés $Z(p)$ caldrà efectuar certes suposicions sobre el procés de manera que sigui possible la inferència sobre els paràmetres del model.

Una de les característiques del *kriging* és que l'error que es pugui produir no dependrà directament de les dades, sinó de la seva continuïtat espacial (Anselin, 1988).

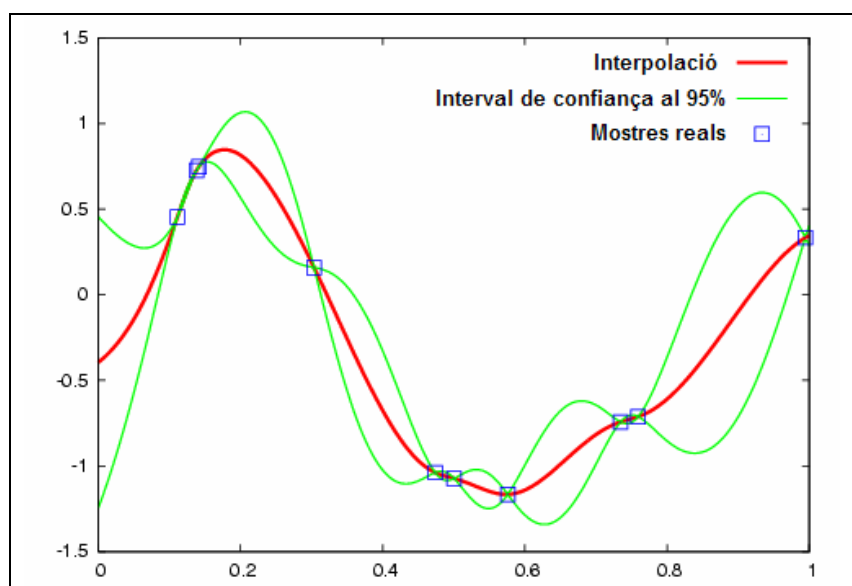


Figura 3.26.: Representació estimació *kriging* en 2D.

3.5.6.1. *Kriging* simple

El cas més senzill és l'anomenat *kriging* simple.

Les hipòtesis d'aquest model són:

- El model és estacionari.
- La mitjana de la funció aleatòria és coneguda

$$E(Z(u)) = m$$

En aquest context, es poden donar dues situacions:

1. La mitjana és coneguda i igual a zero.

Sota aquestes condicions l'estimador *kriging* és no esbiaixat

$$E(Z^*(p)) = \sum_{i=1}^N \lambda_i(p) E(Z(p_i)) = 0 = E(Z(p))$$

Els pesos que satisfaran la condició de variància mínima (Cassel et al., 2000): són el valor de la correlació existent entre el punt a estimar i els punts utilitzats per a realitzar l'estimació.

El procés de minimització fa possible solucionar el sistema d'equacions:

$$\begin{bmatrix} C(0) & C(p_1 - p_2) & \dots & C(p_1 - p_N) \\ C(p_2 - p_1) & C(0) & \dots & C(p_2 - p_N) \\ \dots & \dots & \dots & \dots \\ C(p_N - p_1) & C(p_N - p_2) & \dots & C(0) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} C(p_1 - p_0) \\ C(p_2 - p_0) \\ \dots \\ C(p_N - p_0) \end{bmatrix}$$

La primera matriu defineix la relació entre les observacions disponibles, mentre que la matriu de la part dreta de la igualtat defineix la relació entre les observacions disponibles i el punt a estimar.

La variància de l'error o variància del *kriging* és :

$$\text{var}[Z(p) - Z^*(p)] = \sigma^2 - \sum_{i=1}^N \lambda_i C(p_0 - p_i)$$

2. La mitjana és coneguda i diferent de zero.

En aquesta situació cal considerar una nova funció aleatòria de mitjana zero:

$$Y(p) = Z(p) - m$$

i aleshores el procediment a seguir és el descrit al punt 1.

El sistema d'equacions del *kriging* simple s'haurà de resoldre utilitzant la funció de covariància i no el variograma, mentre que el *kriging* ordinari i el *kriging* tendencial es podran resoldre tant a partir de la funció de covariància com a partir del variograma.

En aquest punt, els interrogants que sorgeixen són:

Què passaria si les observacions disponibles fossin incorrelacionades entre si?

$$\begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} C(p_1 - p_0) \\ C(p_2 - p_0) \\ \dots \\ C(p_N - p_0) \end{bmatrix}$$

I la solució fos

$$\lambda_i = \frac{C(p_0 - p_i)}{\sigma^2} = \rho(p_0 - p_i)$$

Què passaria si les observacions disponibles fossin incorrelacionades amb el punt a estimar?

$$\begin{bmatrix} C(0) & C(p_1 - p_2) & \dots & C(p_1 - p_N) \\ C(p_2 - p_1) & C(0) & \dots & C(p_2 - p_N) \\ \dots & \dots & \dots & \dots \\ C(p_N - p_1) & C(p_N - p_2) & \dots & C(0) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

Donat que la matriu referent a les condicions establertes per a obtenir solució és invertible, aleshores tots els λ_i han de ser iguals a zero i l'estimador del *kriging* simple és aleshores:

$$Z^*(p) = m$$

3.5.6.2. Kriging ordinari

Generalment, el valor de la mitjana m és desconegut i per tant no es pot utilitzar el *kriging* simple.

El *kriging* ordinari estableix el no esbiaixament de l'estimador que acaba suposant en termes matemàtics la condició:

$$\sum_{\alpha=1}^N \lambda_{\alpha}(u) = 1$$

que aplicada sobre les equacions d'estimació del *kriging*:

$$E(Z^*(p)) = \sum_{i=1}^N \lambda_i(p) E(Z(p_i)) = m \sum_{i=1}^N \lambda_i(p) = m$$

Per tant, la forma de trobar els λ_i es basarà en un problema de minimització amb restriccions que es resoldrà per la tècnica dels multiplicadors de Lagrange, ja que en el sistema d'equacions s'haurà d'incloure la restricció establerta sobre els valors λ_i .

$$\Phi(\lambda_1, \dots, \lambda_N, \mu) = \text{var}(Z(p) - Z^*(p)) + 2\mu \left(1 - \sum_{i=1}^N \lambda_i \right)$$

La minimització d'aquesta funció la constituïran per tant els valors dels λ_i que minimitzin la variància i que sumin 1.

El sistema d'equacions a resoldre serà el següent:

$$\begin{bmatrix} C(0) & C(p_1 - p_2) & \dots & C(p_1 - p_N) & 1 \\ C(p_2 - p_1) & C(0) & \dots & C(p_2 - p_N) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ C(p_N - p_2) & C(p_N - p_2) & \dots & C(0) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_N \\ -\mu \end{bmatrix} = \begin{bmatrix} C(p_0 - p_1) \\ C(p_0 - p_2) \\ \dots \\ C(p_0 - p_N) \\ 1 \end{bmatrix}$$

i la variància de l'error serà:

$$\text{var}[Z(p) - Z^*(p)] = \sigma^2 - \sum_{i=1}^N \lambda_i C(p_0 - p_i) - \mu$$

El sistema d'equacions representat en funció dels valors del variograma és:

$$\begin{bmatrix} 0 & \gamma(p_1 - p_2) & \dots & \gamma(p_1 - p_N) & 1 \\ \gamma(p_2 - p_1) & 0 & \dots & \gamma(p_2 - p_N) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \gamma(p_N - p_2) & \gamma(p_N - p_2) & \dots & 0 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_N \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma(p_0 - p_1) \\ \gamma(p_0 - p_2) \\ \dots \\ \gamma(p_0 - p_N) \\ 1 \end{bmatrix}$$

i els resultats que se n'obtidran seran exactament els mateixos.

Què passaria si les observacions disponibles fossin incorrelacionades entre si?

$$\begin{bmatrix} C(0) & 0 & \dots & 0 & 1 \\ 0 & C(0) & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & C(0) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_N \\ -\mu \end{bmatrix} = \begin{bmatrix} C(p_0 - p_1) \\ C(p_0 - p_2) \\ \dots \\ C(p_0 - p_N) \\ 1 \end{bmatrix}$$

i la solució fos:

$$\lambda_j = \rho(u - u_j) + \frac{\mu}{C(0)} \quad j = 1, 2, \dots, N$$

i sota la condició de no biaix:

$$\mu = \frac{C(0)}{N} \left[1 - \sum_{j=1}^N \rho(u - u_j) \right]$$

Què passaria si, a més, les observacions disponibles fossin incorrelacionades amb el punt a estimar?

$$\rho(p_0 - p_j) = 0 \quad \forall j$$

I per tant:

$$\mu = \frac{C(0)}{N} \quad \text{i} \quad \lambda_j = \frac{1}{N} \quad j = 1, 2, \dots, N$$

i l'estimador de kriging ordinari fos:

$$Z^*(p) = \frac{1}{N} \sum_{i=1}^N Z(p_i)$$

3.5.6.3. *Kriging* universal o *kriging* tendencial

El *kriging* universal assumeix que la funció aleatòria Z es pot descompondre de la manera següent:

$$Z(p) = R(p) + m(p)$$

on R és una funció aleatòria estacionària d'ordre 2 amb $E(R(p))=0$ i m és una funció no aleatòria que depèn de la localització p .

Sota aquestes hipòtesis caldrà:

$$E(Z(p)) = E(R(p)) + m(p) = m(p)$$

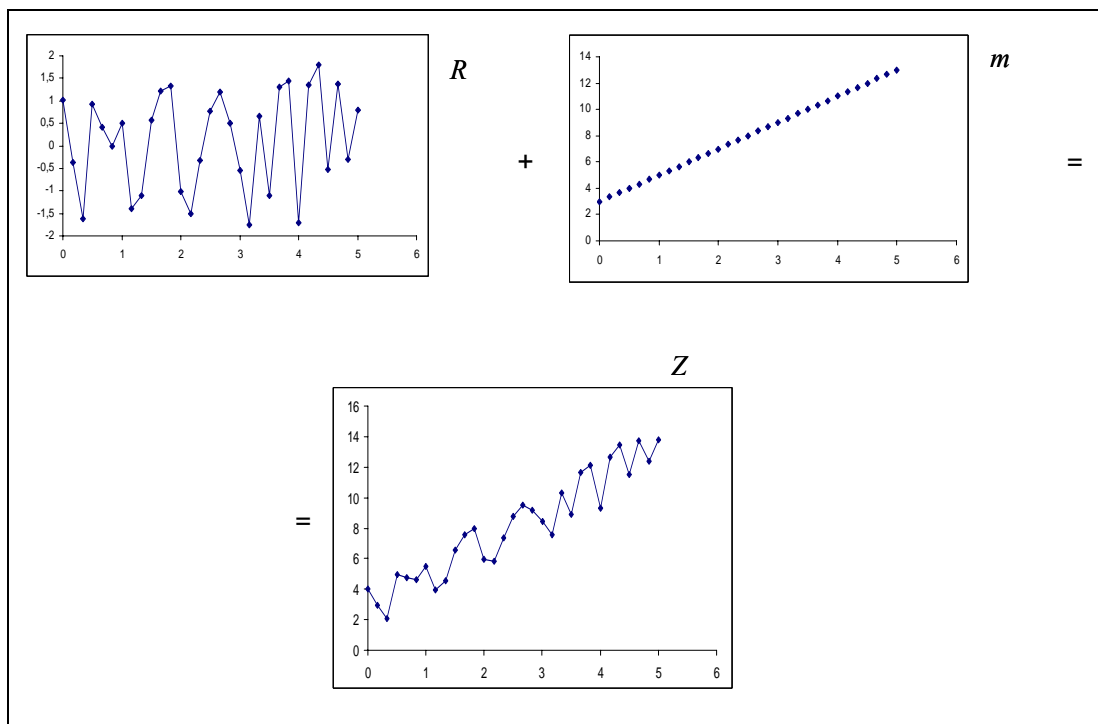


Figura 3.27.: Descomposició de la funció aleatòria en base al *kriging* universal.

Si s'assumeix la descomposició, aleshores per poder realitzar l'estimació de $Z(p)$ caldrà eliminar la tendència i treballar amb el procés estacionari d'ordre 2 amb mitjana zero que es genera:

$$R(p) = Z(p) - m(p)$$

Els valors $R(p)$ es designaran com valors residuals.

Es pot donar el cas en què la funció de tendència sigui coneguda o que no ho sigui; en cas que no ho sigui aquesta s'estimarà a partir de les observacions disponibles.

Una vegada eliminada la tendència es podrà utilitzar la tècnica plantejada en el *kriging* ordinari.

1. Cas de la funció de tendència coneguda

Si la funció de tendència és coneguda, aleshores es pot calcular

$$R(p_i) = Z(p_i) - m(p_i)$$

i realitzar el procés d'estimació d' $R(p_i)$ a partir del *kriging* simple.

Una vegada finalitzat el procés d'estimació cal obtenir el valor de $Z^*(p)$ a partir de:

$$Z^*(p) = R_{ks}^*(p) + m(p)$$

2. Cas de la funció de tendència desconeguda

Quan la funció de tendència és desconeguda és habitual assumir que aquesta podrà ser expressada de la manera següent:

$$m(p) = \sum_{j=0}^L a_j f_j(p) \quad f_0 \equiv 1$$

Les funcions f són conegudes i els paràmetres a són desconeguts, i per tant la funció de tendència és desconeguda.

Les funcions f_j , anomenades 'de base', hauran de ser escollides en funció de la naturalesa del problema i no pas arbitràriament.

La idea és procedir com en el *kriging* ordinari. És a dir, imposar condicions per a filtrar el valor desconegut de la mitjana.

L'estimador que es proposa és:

$$Z^*(p) = \sum_{i=1}^N \lambda_i Z(p_i)$$

i per tant:

$$E(Z^*(p)) = \sum_{i=1}^N \lambda_i m(p_i) = \sum_{j=0}^L a_j \sum_{i=1}^N \lambda_i f_j(p_i)$$

Tot imposant la condició:

$$\sum_{i=1}^N \lambda_i f_j(p_i) = f_j(p) \quad \forall j$$

s'obté un estimador no esbiaixat.

La forma de variància de l'error serà:

$$\text{var}[Z(p) - Z^*(p)] = C(R(p), R(p)) + \sum_{i,j=1}^N \lambda_i \lambda_j C(R(p_i), R(p_j)) - 2 \sum_{i=1}^N \lambda_i C(R(p_i), R(p))$$

i a causa de les restriccions establertes, la funció a minimitzar serà:

$$\Phi(\lambda_1, \dots, \lambda_N, \mu_0, \dots, \mu_L) = \text{var}[Z(p) - Z^*(p)] + 2 \sum_{j=0}^L \mu_j \left(\sum_{i=1}^N \lambda_i f_j(p_i) - f_j(p_0) \right)$$

i derivant respecte els λ_i , $i=1, \dots, N$ i els μ_j , $j=0, \dots, L$

$$\begin{bmatrix} C(0) & C_{12} & \dots & C_{1N} & f_0(p_1) & f_1(p_1) & \dots & f_L(p_1) \\ C_{21} & C(0) & \dots & C_{2N} & f_0(p_2) & f_1(p_2) & \dots & f_L(p_2) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ C_{N1} & C_{N2} & \dots & C(0) & f_0(p_N) & f_1(p_N) & \dots & f_L(p_N) \\ f_0(p_1) & f_0(p_2) & \dots & f_0(p_N) & 0 & 0 & \dots & 0 \\ f_1(p_1) & f_1(p_2) & \dots & f_1(p_N) & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ f_L(p_1) & f_L(p_2) & \dots & f_L(p_N) & 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_N \\ \mu_0 \\ \mu_1 \\ \dots \\ \mu_L \end{bmatrix} = \begin{bmatrix} C_{10} \\ C_{20} \\ \dots \\ C_{N0} \\ f_0(p_0) \\ f_1(p_0) \\ \dots \\ f_L(p_0) \end{bmatrix}$$

que expressat en forma matricial correspon a:

$$\begin{bmatrix} \Sigma & F \\ F' & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} C_0 \\ F_0 \end{bmatrix}$$

La unicitat de la solució dependrà de la matriu F , la qual està subjecte a la configuració de les observacions disponibles.

3.5.7. Altres mètodes d'estimació espacial

Els mètodes que es presentaran a continuació no han estat implementats.

Tot i això, s'ha considerat escaient per tal de donar una visió completa del treball realitzat per aquesta tesi incloure'n una breu presentació.

3.5.7.1. Mètode dels veïnatsges

Existeixen diverses variacions del mètode dels veïnatsges; totes elles àmpliament utilitzades. Les més destacables són:

- El mètode dels k veïns més propers.
- El mètode dels veïns dins del radi.

El primer, el dels k veïns més propers, acota el nombre de participants en el càlcul, mentre que el segon el que acota és el radi màxim a utilitzar.

L'algoritme de càlcul és el següent:

1. Es càlcul la distància al punt a estimar per a totes les mostres.
2. S'ordenen les mostres en funció de la distància.
- 3a. Si el mètode utilitzat és el dels k veïns més propers, s'escullen els primers k elements de la llista.
- 3b. Si el mètode utilitzat és el del radi, s'escullen totes aquelles mostres amb distància inferior al radi triat.
4. Es calcula el valor en el punt a estimar a partir de l'estadístic decidit.

Encara que l'estimador utilitzat no quedi definit pel mètode, l'utilitzat més habitualment sol ser la mitjana, encara que es possible utilitzar la distància inversa o qualsevol altre estadístic.

Com es pot observar a la figura 3.28, el radi utilitzat en el cas dels k veïns més propers depèn de la densitat de mostres al voltant del punt.

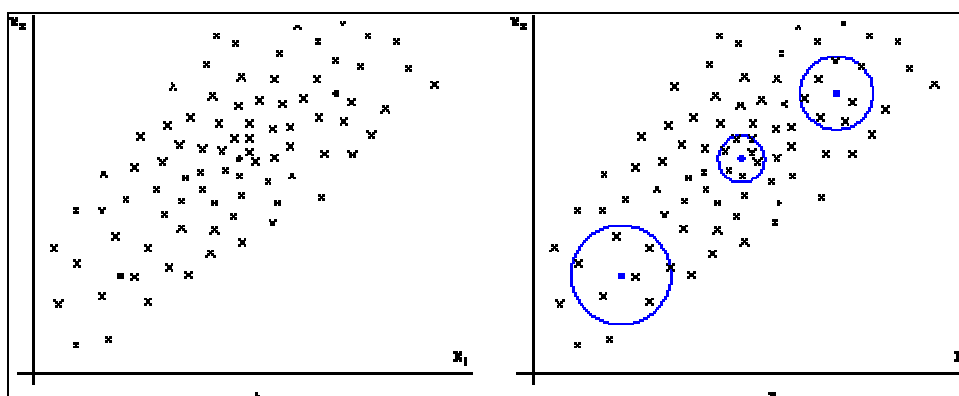


Figura 3.28.: Mostres en l'espai considerant els 7 veïns més propers.

3.5.7.2. Mètode de triangulació

El mètode de triangulació es basa en la creació de triangles que a partir dels punts mostrals que compleixen la condició de triangulació Delaunay. Aquesta condició estableix que la circumferència circumscrita de cada triangle de la xarxa no ha de contenir cap vèrtex de cap altre triangle. El resultat es un *patchwork* de cares de triangles sobre el conjunt de punts.

Aquest mètode és un interpolador exacte. Cada triangle generat defineix un pla sobre el conjunt de punts que delimiten el triangle, el pendent del triangle, així com l'elevació del triangle respecte al pla quedarà determinat per les tres mostres que han servit per a definir el triangle. Tots els punts que es trobin dins de l'espai mostral quedaran definits per la superfície triangular generada.

La majoria de programaris comercials utilitzen aquest mètode per a la realització de les imatges d'interpolacions espacials.

El requeriment legal de treballar amb un mínim de sis mostres descarta automàticament aquest mètode com a possible mètode d'estimació.

3.6. Mètodes d'estimació no lineals

3.6.1 Introducció

Habitualment els paràmetres del variograma s'obtenen totalment o parcial de forma visual. El procediment presentat en aquesta tesi no admet intervenció humana i és per això que cal incorporar-hi mètodes d'estimació que aconseguixin modelitzar el variograma, tant pel que fa a la forma (família de corbes a què pertanyen) com pel que fa als paràmetres.

Les funcions que defineixen els variogrames teòrics no són lineals, és per això que una estimació lineal podia donar solucions de poca qualitat. En base a aquest plantejament es va decidir utilitzar procediments iteratius d'estimació no lineals.

3.6.2. Mètode d'obtenció del variograma teòric

En base a les propietats dels diferents variogrames teòrics i a la bibliografia específica d'aplicació d'aquests models a l'àmbit immobiliari (Chica, 1994) es va acabar limitant els possibles models de variogrames teòrics utilitzables a tres models: l'esfèric, el gaussià i l'exponencial.

El procés d'estimació s'havia de realitzar en dos passos.

- Realitzar el procés d'estimació dels paràmetres per a cada una de les formes de variograma teòric.
- Escollir a partir d'algun estadístic de comparació la millor de les tres estimacions.

Per a realitzar el procés d'estimació dels paràmetres dels diferents variogrames teòrics es va utilitzar el mètode dels mínims quadrats ordinaris combinat amb l'algoritme d'optimització no lineal de Gauss-Newton. Amb aquest procediment s'obtenia la millor corba de cada una de les tres famílies de variogrames.

El procés de selecció entre les diferents corbes es va realitzar amb els resultats del criteri d'informació d'Akaike.

El variograma teòric seleccionat era, doncs, la millor estimació d'entre les tres famílies de variogrames.

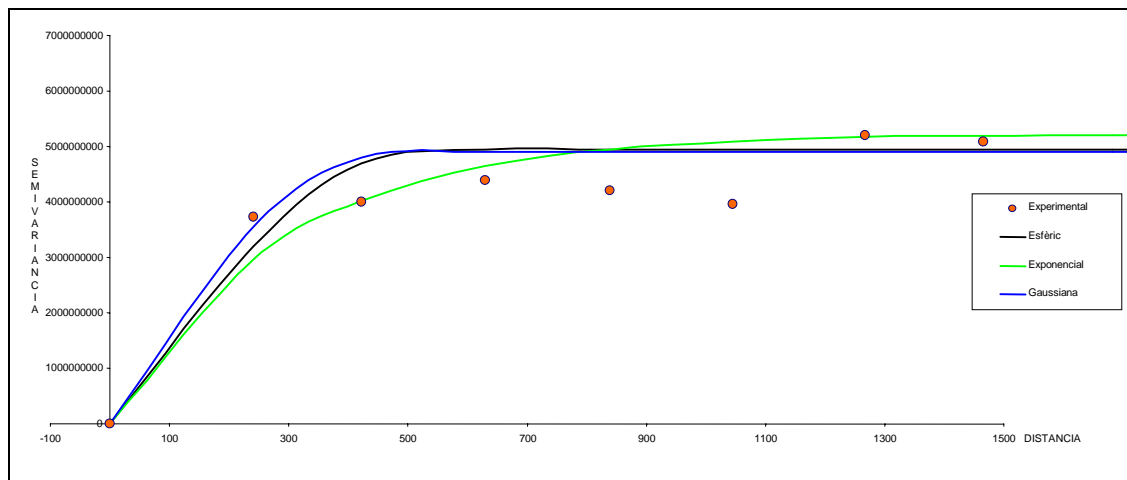


Figura 3.29. Representació gràfica de les tres millors estimacions.

3.6.3. Mètode dels mínims quadrats ordinaris i l'algorisme de Gauss-Newton

S'entén per funció no lineal aquella en què al realitzar la primera derivada respecte als seus paràmetres els resultats que s'obtenen són funcions no lineals.

A continuació, el model general:

$$y_i = f(X_i, \beta) + e_i \quad i=1,2,\dots,n \quad (a)$$

L'estimador de mínims quadrats no lineals serà aquell que minimitzi la suma de quadrats residuals:

$$S(\beta) = \sum e_i^2 = \sum_{i=1}^N [y_i - f(x_i, \beta)]^2 \quad (b)$$

Les condicions de primer i segon ordre, necessàries i suficients per a la minimització (òptim local), respectivament, venen donades per:

$$\frac{\partial S(\beta)}{\partial \beta} = -2 \sum_{i=1}^N \frac{\partial f(x_i, \beta)}{\partial \beta} [y_i - f(x_i, \beta)] = 0 \quad (c)$$

$$\frac{\partial^2 S(\beta)}{\partial \beta \partial \beta'} = 2 \left(\sum_{i=1}^N \frac{\partial f(x_i, \beta)}{\partial \beta} \frac{\partial f(x_i, \beta)}{\partial \beta'} - \sum_{i=1}^N (y_i - f(x_i, \beta)) \frac{\partial^2 f(x_i, \beta)}{\partial \beta \partial \beta'} \right) \quad (d)$$

en què la matriu de l'equació (d) ha de ser definida positiva.

La majoria de les vegades aquestes equacions no tenen una solució analítica i aleshores cal utilitzar un mètode iteratiu. El mètode de Gauss-Newton arriba a una solució a partir d'un algoritme numèric.

Si es considera a continuació el mateix model plantejat en (a) i la funció de minimització de (b), es planteja una solució per mínims quadrats.

$$y_i = f(X_i, \beta) + e_i \quad i=1, 2, \dots, n$$

$$S(\beta) = \sum e_i^2 = \sum_{i=1}^N [y_i - f(x_i, \beta)]^2$$

Suposant que e segueix una distribució normal de mitjana 0 i variància σ^2 .

La condició de primer ordre per trobar un mínim de la funció és:

$$\frac{\partial S(\beta)}{\partial \beta} = -2 \sum_{i=1}^N \frac{\partial f(x_i, \beta)}{\partial \beta} [y_i - f(x_i, \beta)] = 0 \quad (e)$$

El problema de l'estimació és trobar els valors de β que compleixin la condició de primer ordre anterior i que corresponguin a un mínim global de la funció de suma de quadrats. Perquè això es produeixi es canvia la $f(x_i, \beta)$ per una aproximació desenvolupada a partir de la sèrie de Taylor de primer ordre al voltant del paràmetre β_0 :

$$f(\mathbf{x}, \beta) \approx f(\mathbf{x}, \beta_0) + \sum_{k=1}^K \frac{\partial f(\mathbf{x}, \beta)}{\partial \beta_k} (\beta_k - \beta_k^0) \quad (f)$$

en què β_k^0 és una solució inicial;

i donat que $y = f(x_i; \beta) + \varepsilon_i$, es pot escriure l'expressió 3.5.6. com:

$$Y_i = f(\mathbf{x}, \beta_k^0) + \sum_{k=1}^K \frac{\partial f(\mathbf{x}, \beta)}{\partial \beta_k} (\beta_k - \beta_k^0) + \varepsilon_i \quad (\text{g})$$

i si es calcula $Y_i^0 = Y_i - f(\mathbf{x}, \beta_k^0)$ l'equació 3.5.7. quedarà com:

$$Y_i^0 = \sum_{k=1}^K \frac{\partial f(\mathbf{x}, \beta)}{\partial \beta_k} (\beta_k - \beta_k^0) + \varepsilon \quad (\text{h})$$

i fent

$$\mathbf{g}^0 = \left(\left(\sum_{k=1}^K \frac{\partial f(\mathbf{x}, \beta)}{\partial \beta_k} \right)^T \left(\sum_{k=1}^K \frac{\partial f(\mathbf{x}, \beta)}{\partial \beta_k} \right) \right)^{-1} \left(\sum_{k=1}^K \frac{\partial f(\mathbf{x}, \beta)}{\partial \beta_k} \right)^T Y^0 \quad (\text{i})$$

Es pot obtenir el vector d'estimadors utilitzables per obtenir coeficients revisats de la manera següent:

$$\beta_k^1 = \beta_k^0 + \mathbf{g}_k^0$$

i β_k^1 serà el vector d'estimadors de β_k un cop finalitzada la primera iteració.

Per comprovar si els ajusts s'estan realitzant apropiadament s'emprarà el criteri dels mínims quadrats.

A partir del vector d'estimadors inicials β^0 s'obindrà la suma quadràtica residual (SSE). Aquest mateix procediment es repetirà a partir dels β^1 .

Si el $SSE_1 < SSE_0$, els coeficients revisats seran millors que els inicials, per tant, es pot afirmar que l'algoritme està treballant correctament.

El procediment torna al punt 3.5.6. prenent ara com a valors inicials els estimadors β_k^1 , i així successivament fins que les diferències entre els SSE de les dues estimacions successives siguin inferiors al criteri de convergència establert.

3.7. L'estimació de la tendència

El models *kriging* que s'utilitzen en aquesta tesi parteixen de la condició de la no existència tendència global, per tant abans de dur a terme qualsevol dels procediments de càlcul presentats a partir de les metodologies *kriging* caldrà realitzar una anàlisi de la tendència i en cas d'existir-ne caldrà extreure-la primer i tornar-la a incorporar en finalitzar el procés.

Per tal d'avaluar si existeix comportament tendencial en alguna de les direccions s'utilitzaran el mètodes regressius.

Donat que el nombre de models possibles és baix s'emprarà un mètode exhaustiu d'avaluació basat en l'idea del subconjunts regressius (Neter, Wasserman i Kutner, (1985), per tant s'avaluarà:

$$VH = \beta_1 \cdot x + \beta_2 \cdot y + \beta_3 \cdot x \cdot y$$

$$VH = \beta_1 \cdot x$$

$$VH = \beta_1 \cdot x + \beta_3 \cdot x \cdot y$$

$$VH = \beta_2 \cdot y$$

$$VH = \beta_2 \cdot y + \beta_3 \cdot x \cdot y$$

Si existeix més d'un model que rebutgi la hipòtesi nul·la i tingui paràmetres estadísticament diferents de zero, s'escollirà aquell model que minimitzi el C_p de Mallows.

$$C_p = \frac{SSE_P}{S^2} - N + 2P$$

És possible utilitzar altres estadístic com el Criteri d'informació d'Akaike, ja utilitzat anteriorment per a seleccionar el model de variograma.

4. El mètode de comparació de mercat basat en un sistema expert

4.1. Introducció

Una vegada establert el marc legal i presentades totes les eines i els procediments teòrics amb detall suficient, s'escau plantejar-ne els usos en la construcció d'una metodologia de càlcul per passos que tingui com a solució el valor de taxació l'objecte immoble.

Molts dels processos utilitzats no seran desenvolupats amb detall, donat que no són l'objectiu d'aquesta tesi, no obstant això, la informació presentada és del tot suficient i entenedora com per a poder seguir sense dificultat la metodologia plantejada. La seva presentació s'ha cregut necessària per a poder establir un punt de partida del procés d'homogeneïtzació espacial.

4.2. Estudi de necessitats i obligacions

El primer pas per crear una base de dades i un programari de qualitat és estudiar les necessitats d'informació que es poden presentar. És en base a aquesta idea que es va dur a terme el repàs i l'extracció d'informació de l'ECO/805/2003 amb què posteriorment es van definir els camps necessaris sota perspectiva legal i els procediments de càlcul a utilitzar.

A l'article 21 de l'ECO/805/2003 s'han definit els requisits necessaris per a dur a terme el mètode de comparació.

Bàsicament i succint, s'hi assenyala que per tal d'aplicar el mètode de comparació cal que s'acompleixin dues condicions:

- que existeixi un nombre suficient d'immobles comparables, amb preus coneguts;
- que les mostres escollides per a l'obtenció del valor siguin representatives del mercat dels immobles similars.

A l'article 22 es designen els procediments a realitzar per trobar el valor per comparança.

1. Estudi del mercat de comparables.
2. Elecció de les mostres més significatives.
3. Homogeneïtzació de les mostres.
4. Determinació del valor de mercat.

Donat que tots aquest passos havien de ser observables i explicables pels agents participants i l'objectiu empresarial de CATSA era crear un sistema expert que simulés el procés humà, però que alhora complís amb la llei esmentada, es va fer palesa la necessitat que el sistema expert donés per a cada un dels passos les explicacions i les sortides necessàries que els agents participants requerien.

A continuació, es presenta la recopilació de les necessitats que es van palesar a partir d'aquest estudi:

- Necessitat de definició de totes les característiques participants.
- Establiment dels criteris de selecció de mostres comparables en base a filtres excloents.
- Creació d'un sistema d'homogeneïtzació de les característiques constructives de l'habitatge.
- Definir un sistema de selecció de mostres homogeneïtzades basat en els diferencials provocats pel procés d'homogeneïtzació.
- Creació d'una metodologia d'homogeneïtzació espacial.
- Recerca de mètodes de la funció de valor.

4.3. Selecció dels camps participants en el mètode de comparació

La selecció de les característiques participants ens els diferents processos de càlcul es va basar en els criteris experimentats del conjunt d'experts de CATSA, en l'ECO/805/2003, en els informes de les agències de la propietat immobiliària i en les variables utilitzades en Ceular i Caridad (2001), Gallego Mora-Esperanza (2004), P. Garcia (2004), Guimet i Urrutia (2000), i Roca Cladera (1986).

Les característiques més presents en les referències anteriorment citades per a la realització del càlcul del valor de taxació van ser la localització, la superfície, l'antiguitat, l'estat de conservació, les reformes realitzades, el nombre de banys i habitacions. La majoria de referències fan esment de característiques com les instal·lacions (calefacció, aire condicionat, energia solar, ascensor, etc.) o les qualitats constructives (acabats, banys, etc.). Altres característiques que també van ser considerades en algunes situacions específiques van ser:

- Externalitats (per exemple: la línia de mar)
- Orientació
- Nombre de façanes
- Longitud de les façanes
- Forma
- Nombre d'habitacions interiors
- Altura
- Situacions especials extrínseques

Altres característiques com la jerarquia social, els indicadors de qualitat de les infraestructures i les comunicacions, que molts autors (Garcia, 2005) utilitzen en els processos de càlcul del valor, perden una part important de la seva significació en passar a treballar amb mètodes basats en la localització per coordenades de l'immoble, ja que aquesta informació és inherent a la localització.

El conjunt de característiques requerides pel procés és molt més extens que el conjunt de característiques que seran utilitzades durant el procés d'obtenció del valor pel mètode de comparació, ja que existeixen una gran quantitat de camps relacionats amb els requeriments legals establerts a l'ECO805/2003.

Totes les característiques utilitzades en el procés de càlcul van haver de ser tecnificades i protocol·litzades perquè poguessin ser utilitzades pel sistema expert.

A continuació es presenten les característiques que són d'obligada utilització en el procés de càlcul del valor per comparació, acompanyades d'una explicació del seu ús.

4.3.1. Tipologia i subtipologia

En base a la bibliografia estudiada (BOC, M2, Construct) es va decidir utilitzar les definicions tipològiques següents:

Residencial plurifamiliar	% cost de construcció dedicat a acabats
Bloc aïllat (tipus VPO) estr. mur fàbrica / alt<4 pl.	36%
Bloc aïllat (tipus VPO) pilars i reticular h.a. / alt>4 pl.	34%
Bloc aïllat (tipus apartaments de 40 m ² útils)	42%
Edifici entre mitjanes (renta normal)	43%
Bloc aïllat (tipus renda de luxe)	47%

Taula 4.1.: Tipus de residencials plurifamiliars i percentatges de cost de construcció dels acabats.

Una vegada decidides les subtipologies es va haver d'estudiar el percentatge que cada una de les subtipologies dedicava en cost als acabats (taules 4.1 i 4.2).

Les taules a partir de les quals es van extreure aquests resultats es ressenyen a l'annex I.

L'ús dels percentatges sobre acabats va permetre calcular el cost destinat als acabats i els diferencials monetaris que es van calcular de la diferència

d'estat i qualitat. Per tant, aquests percentuals s'utilitzen per a realitzar l'homogeneïtzació de la qualitat i l'estat.

Tanmateix, la tipologia estableix un primer filtre, donat que els habitatges plurifamiliars només seran comparables entre ells.

4.3.2. Definició de l'habitatge

L'ECO805/2003, en els articles 70 i 71, requereix una descripció de tots aquells elements que poden afectar el valor final de l'immoble. A més, especifica la necessitat d'indicar les dades relacionades amb les reformes, les distribucions i les superfícies de l'immoble.

Per tant, la definició de les superfícies, el tipus de superfície adoptada, el nombre d'elements, els acabats, les qualitats, l'estat de conservació i les reformes realitzades, així com l'any de construcció són part dels requeriments legals.

A la taula 4.2. es mostren, per una banda, la definició dels camps utilitzats, i per l'altra, les categories establertes per a cada camp amb l'objectiu de definir les característiques de l'habitatge.

Parts de l'habitatge	Camps requerits
Banys	Nombre
	Superfícies
	Paviments
	Sanitaris
	Parets
	Sostres
Cuina	Nombre
	Superfícies
	Paviments
	Mobiliari de cuina
	Parets
	Sostres
Dormitoris, sales, trasters i distribuïdors	Nombre
	Superfícies
	Paviments
	Parets
	Sostres

Taula 4.2.: Definició d'elements que conformen l'habitatge (1/2).

Parts de l'habitatge	Camps requerits
Galeries, terrasses i balcons	Nombre
	Superfícies
	Paviments
	Parets

Taula 4.2.: Definició d'elements que conformen l'habitatge (2/2).

A més, es van definir elements generals com: fusteries exteriors, fusteries interiors, vidres i instal·lacions elèctriques.

L'estudi dels materials i la corresponent agrupació van ser desenvolupats pel cos de tècnics de CATSA i se'n mostren alguns exemples a les taules 4.3. i 4.4.

Per a cada una d'aquestes categories es va haver d'establir un cost de col·locació genèric. Aquest procés es va dur a terme a partir d'extreure la informació localitzada a la base '*Precios de Edificación y Obra Civil En España*' (PREOC) per a diferents anys, la base de l'ITEC, la base CENTRO i les informacions disponibles al BOC. El departament tècnic va ser el designat d'establir-ne les categories i de definir la categoria a què pertanyia cada un dels materials.

Donat que coexistien bases de diferents anys es van haver de realitzar les actualitzacions necessàries per a aconseguir una igualació en any. Aquestes actualitzacions es van fer en base a l'índex de preus estatals dels materials de la construcció i de la mà d'obra proporcionats per l'INE.

Una vegada homogeneïtzades totes les bases en any, les informacions es van agrupar en una única base i se'n va realitzar l'explotació estadística necessària per a obtenir un cost per categoria.

Categoria	Funció	Origen	Tècniques	Tipus genèric	Tipus detallat	Preu/m ²
PAV. 01	Revestiment	Natural	Additiva (hidràulica)	Pedra	Primera	91
PAV. 02	Revestiment	Natural	Additiva (hidràulica)	Pedra	Segona	60
PAV. 03	Revestiment	Natural	Additiva (hidràulica)	Pedra	Tercera	50

Taula 4.3.: Exemple d'agrupacions de materials. Estudi de paviments (1/2).

Categoria	Funció	Origen	Tècniques	Tipus genèric	Tipus detallat	Preu/m²
PAV. 04	Revestiment	Natural	Additiva (encolat), Fixació (clavada, flotant)	Fusta	Primera	76
PAV. 05	Revestiment	Natural	Additiva (encolat), Fixació (clavada, flotant)	Fusta	Segona	40
PAV. 06	Revestiment	Natural	Additiva (encolat), Fixació (clavada, flotant)	Fusta	Tercera	36
PAV. 07	Revestiment	Natural	Additiva (encolat), Fixació (clavada, flotant)	Fusta	Suro	34
PAV. 08	Revestiment	Artificial	Additiva (hidràulica)	Petris	Gres	30
PAV. 09	Revestiment	Artificial	Additiva (hidràulica)	Petris	Gres porcellànic	40
PAV. 10	Revestiment	Artificial	Additiva (hidràulica)	Petris	Terratzo	30
PAV. 11	Revestiment	Artificial	Additiva (hidràulica)	Petris	Hidràulic	20
PAV. 12	Revestiment	Artificial	Additiva (hidràulica)	Petris	Formigó	17
PAV. 13	Revestiment	Artificial	Additiva (hidràulica)	Petris	Fang	41
PAV. 14	Revestiment	Artificial	Additiva (encolat)	Orgànic	Fenòlic - Melamínic	45
PAV. 15	Revestiment	Artificial	Fixació (Flotant)	Orgànic	Resina	17
PAV. 16	Revestiment	Artificial	Estès – Projectat	Orgànic	Vinílic	26
PAV. 17	Revestiment	Artificial	Additiva (hidràulica)	Teixit	Moqueta	33

Taula 4.3.: Exemple d'agrupacions de materials. Estudi de paviments (2/2).

Les reformes plantejades són les següents:

Lloc	Reforma	Parts afectades	Recuperació
Interior	Pintura	Parets i sostres pintats	100%
Interior	Reforma paviments	Tots els paviments	100%
Interior	Reforma cuina	Tots els elements de cuina	100%
		Aigua, gas, electricitat	30%
Interior	Reforma banys	Tot els elements del bany	100%
		Aigua, gas, electricitat	30%
Interior	Reforma cuina i banys	Tots els elements de cuina i banys	100%
		Aigua, gas, electricitat	80%
Interior	Reforma fusteria interior	Fusteria interior	100%
Interior	Reforma fusteria exterior	Fusteria exterior	100%
Interior	Reforma instal·lacions	Aigua, gas, electricitat	100%

Taula 4.4.: Definició de reformes (1/2).

Lloc	Reforma	Parts afectades	Recuperació
Interior	Reforma integral sense estructura	Totes les parts de l'habitatge menys estructura	100%
Interior	Reforma integral amb estructura	Tot	100%
Edifici	Façana	Façana	100%
Edifici	Coberta	Coberta	100%
Edifici	Instal·lacions	Aigua, gas, electricitat, comunicacions, telefonia, sanejament	100%
Edifici	Integral sense estructura	Façana, coberta, instal·lacions, interior i elements comuns	100%
Edifici	Integral amb estructura	Tot l'edifici i els elements comuns	100%

Taula 4.4.: Definició de reformes (2/2).

Amb la informació tractada en aquest punt es van generar diferents índexs que van permetre estudiar diversos diferencials de cost i utilitat:

- Diferencials de cost deguts a la superfície.
- Diferencials de cost degut a la qualitat i l'estat.
- Diferencials de cost degut al nombre d'habitacions en funció de la superfície.
- Diferencials de cost degut al nombre de banys.

Una de les característiques de l'habitatge de fonamental importància és la planta. Els procediments realitzats sobre aquesta característica es presentaran en el punt següent conjuntament amb la característica de l'edifici – existència d'ascensor.

Les qualitats, els estats, el nombre de banys o estances no seran filtres excloents i per tant no seran utilitzat en el procés de selecció de comparables.

4.3.3. Caracterització de les parts específiques de l'edifici

Inicialment, les parts de la construcció considerades comunes no havien de ser tractades pel procés; de manera que la imputació de valor s'havia d'efectuar de manera directa en funció de la tipologia i la subtipologia establerta.

Tanmateix, existeixen certes parts comunes que apareixen de forma habitual a les indicacions de reformes dels processos valoratius. Degut a això i al fet que aquestes reformes poden afectar al valor final de l'immoble es va decidir caracteritzar-les. Per tant, en cas d'existir aquestes reformes el seu 'estat' s'haurà de definir en funció de l'any de reforma.

Així doncs, es va haver de realitzar la caracterització dels elements de façana, coberta i instal·lacions de l'edifici.

Elements definits per l'edifici	Reformes		
Moviment de terres			
Fonaments			
Estructura			Integral amb estructura
Sanejament (horitz. i vertic.)		Integral sense estructura	Integral amb estructura
Feines de paleta gruixudes		Integral sense estructura	Integral amb estructura
Lampisteria	Instal·lacions	Integral sense estructura	Integral amb estructura
Energia solar (ACS)	Instal·lacions	Integral sense estructura	Integral amb estructura
Electricitat	Instal·lacions	Integral sense estructura	Integral amb estructura
Calefacció	Instal·lacions	Integral sense estructura	Integral amb estructura
Instal·lacions especials	Instal·lacions	Integral sense estructura	Integral amb estructura
Ascensors	Instal·lacions	Integral sense estructura	Integral amb estructura
Feines de paleta terrats i impermeabilitat	Reformes de coberta	Integral sense estructura	Integral amb estructura
Feines de paleta acabats de façana	Reformes de façana	Integral sense estructura	Integral amb estructura
Feines de paleta ajudes a industrials		Integral sense estructura	Integral amb estructura
Feines de guix		Integral sense estructura	Integral amb estructura

Taula 4.5.: Elements d'edifici i reformes considerades.

Gràcies a aquesta informació es van poder estudiar els diferencials de cost imputables a l'edifici.

Les característiques 'ascensor' i 'planta' tenen dos usos dins del procediment de càlcul del valor de taxació:

- Participen en càlculs de diferencial de valor, és a dir, són part del procés d'homogeneïtzació;
- Tenen un comportament de filtre pel que fa a la selecció de comparables.

4.3.4. La localització per geoposició

La possibilitat d'utilització de la localització exacta -indicada en coordenades- tant de l'objecte a valorar com de les mostres utilitzades ha estat un gran avanç pel que fa a l'estudi de les relacions i dependències entre els preus en funció de les localitzacions exactes dels diferents elements participants en el càlcul.

Es va definir la distància com un dels filtres principals per a seleccionar les mostres que participen en els càlculs inicials.

La geoposició juntament amb els valors homogeneïtzats seran les variables que participaran en el procés d'homogeneïtzació espacial i en l'obtenció de la funció del valor.

4.4. Establiment dels criteris de selecció de mostres comparables en base a filtres excloents

Totes les característiques de l'immoble esmentades anteriorment participen en el mètode de comparació com a variables de càlcul i cada una d'elles va ser analitzada per veure quins límits de comparabilitat establia en base als principis de substitució i de mercat. Tanmateix, es va fer evident que no era suficient amb l'estudi de característiques una a una, sinó que calia realitzar un estudi de comportaments conjunts, ja que la combinació de dues o més variables podia produir situacions de no comparabilitat, no imputables a partir dels estudis individuals o a la inversa.

El primer que es va haver de fer va ser analitzar quan una mostra deixava de ser comparable a causa d'una determinada característica. Per tant, calia establir quina era la distància màxima d'acceptació de la comparabilitat.

Les preguntes a què calia donar resposta en aquest context van ser les següents:

- Quines característiques són utilitzables en la recerca de comparables com a filtres excloents i quines participen únicament en el càlcul dels diferencials de valor?
- Com establir la distància màxima de comparabilitat per a cada una de les característiques amb filtres excloents?

Una primera aproximació a aquestes qüestions han estat les definicions d'objectes comparables:

- En base al principi de substitució, s'ha designat com a mostra comparable aquella considerada com a equivalent o substituïble (Roca, 1986).
- En base al mercat, es considerarà que dos béns són comparables quan tinguin mercats iguals o molt semblants (Roca, 1986).

Per estudiar els límits de la comparabilitat i les característiques que participen en aquest procés, el departament tècnic de CATSA va realitzar una enquesta exhaustiva entre els tècnics.

Les característiques que a partir de l'enquesta es van definir com a filtres excloents van ser:

- Tipologia
- Antiguitat
- Distància
- Superfície
- Ascensor i planta

A mode de puntualització, el filtre per tipologia es va definir com un filtre d'exclusió directa en el cas dels habitatges plurifamiliars.

El filtre per antiguitat diferencia únicament l'obra nova de la de segona mà.

El filtre per distància també es va definir com d'exclusió directa. Aquest filtre pot superar les limitacions municipals i es va pensar com una manera de reduir el nombre de mostres amb la intenció de fer el procés de càlcul més lleuger. En aquest sentit, es va establir la doble condició:

- distància al objecte < 2000 metres
- nombre de mostres ordenades per distància < 200 mostres

A partir dels resultats de l'enquesta es va definir un sistema de corbes per a les diferents tipologies. La corba de comparabilitat per habitatges plurifamiliars es mostra a continuació a la figura 4.1:

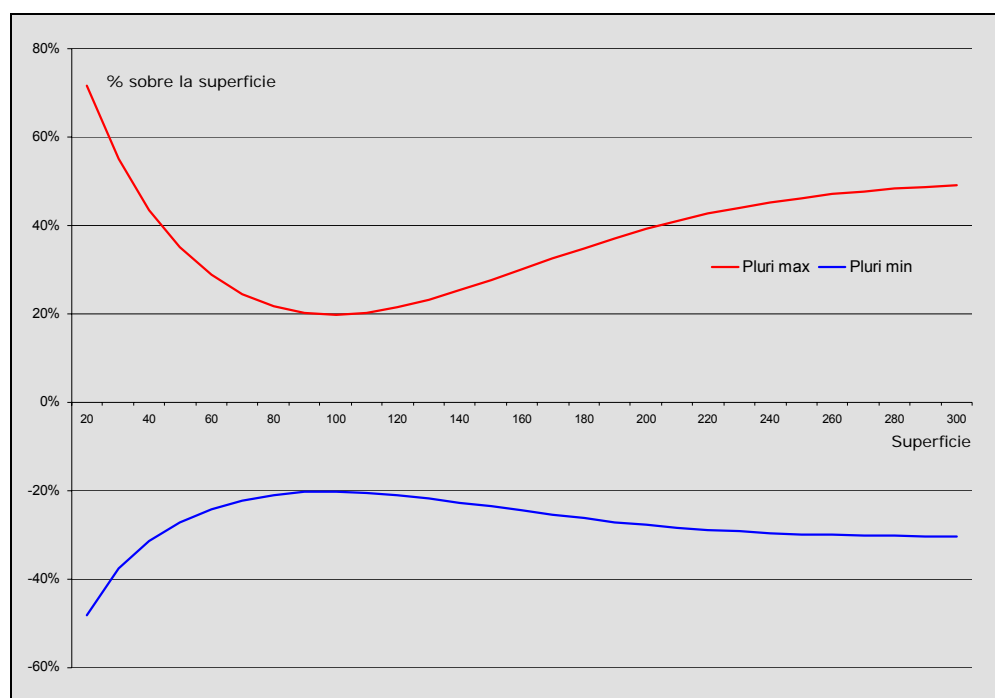


Figura 4.1: Comparabilitat en funció de la superfície.

El filtre per ascensor i planta també es va establir com un filtre d'exclusió amb les condicions següents:

- Si la mostra i l'objecte estan situats en una planta inferior o igual a dos, aleshores són comparables i l'existència o no d'ascensor no hi influeix.
- Si la mostra i l'objecte tenen ascensor, aleshores són comparables sense rellevància de la planta.
- Si la mostra o l'objecte estan situats en plantes superiors a la dos, caldrà disposar d'ascensor a tots dos elements perquè puguin ser comparables.
- Si la mostra es troba en una planta inferior a dos i no té ascensor i l'objecte es troba en una planta superior a dos i té ascensor també són comparables. Això mateix serà aplicable en cas invers.
- La mostra o l'objecte es troben en una planta superiors a dos i l'altre element té ascensor, aleshores no són comparables.
- La mostra i l'objecte es troben en una planta superiors a dos i a un màxim de dues plantes de distància entre elles.
- La resta de casos no són comparables.

4.5. Creació d'un sistema d'homogeneïtzació de les característiques constructives de l'habitatge

L'homogeneïtzació s'obté a partir de l'estudi dels diferencials de valor provocats per les diferències en les característiques que imputen valor a l'immoble. En base a aquesta idea l'objectiu va ser establir pel conjunt de mostres comparables els mètodes d'homogeneïtzació per a cada característica o conjunt de característiques per a tots els diferencials possibles.

La situació ideal hagués estat poder estudiar a partir de residuals el comportament de cada una de les característiques involucrades en el procés d'homogeneïtzació a partir d'informacions de mercat, però mentre que hi ha variables com la superfície o la planta en que es possible aconseguir amb certa facilitat, situacions pel càlcul del residual. Pel que fa a la majoria de característiques a analitzar difícilment es donaran situacions perfectes en les quals l'única variable diferenciada fos la d'interès.

La solució va venir donada novament per la realització d'un estudi estadístic i la comprovació de les dades de la base de dades de CATSA i l'experiència del experts per acabar acceptant les solució plantejades.

Les característiques que van participar en el procés homogeneïtzació van ser: la superfície, l'estat i la qualitat, la planta i existència d'ascensor, el nombre de banys, el nombre de dormitoris i la seva relació amb la superfície.

La definició de les diferents característiques es va obtenir a partir d'una enquesta realitzada pels serveis tècnics de CATSA a la totalitat dels tècnics validadors de l'empresa, i que va permetre definir els diferents paràmetres que participen en el procés d'homogeneïtzació.

Aquesta enquesta es va plantejar a mode de vint exercicis d'homogeneïtzació a resoldre per part dels tècnics, en què les diferents situacions plantejades permetien aïllar cada un dels paràmetres en diferents situacions.

En cada un dels vint exercicis es demanava als tècnics que homogeneïtzessin sis habitatges amb valor desconegut en funció d'un habitatge amb preu conegut.

4.5.1. Homogeneïtzació de la superfície

Les diferències per superfície entre l'objecte a taxar i les mostres es va obtenir a partir d'una funció del cost de construcció, la qualitat i el percentatge del cost dels acabats sense equipament.

4.5.2. Homogeneïtzació de l'estat i la qualitat

El procés de càlcul del valor de mercat s'haurà de dur a terme dues vegades durant el procés de valoració; el primer l'ha de realitzar el tècnic visitador a partir de mostres procedents de diverses fonts com ara de llocs web, diaris, anuncis o informació dels promotors. El segon l'ha de realitzar el tècnic valorador a partir de les taxacions ja existents realitzades amb anterioritat a CATSA.

Per tant, va caler el tractament de la informació a diferents nivells en funció de la quantitat d'informació disponible.

El sistema que es presenta a continuació es base en els estudis realitzats per A. Garcia i H. DiPasqua (2007), plantejat per a treballar en situacions d'informació mínima (com és el cas de les mostres) per poder estudiar els diferencials derivats de la qualitat, l'estat i l'antiguitat.

El Coeficient d'antiguitat, reformes i qualitats (ARQ) es va definir a partir de l'antiguitat, les reformes realitzades i les qualitats dels materials utilitzats; definides exactament de la mateixa manera que el Coeficient d'estat i qualitat material (CEQM) ho fa pel procés que es realitza a partir de les taxacions ja solucionades. Es va decidir definir l'estat de l'immoble sota les possibilitats que oferien els formats habituals d'ofertes que es podrien trobar tant a la xarxa com als diaris:

Estat	
Nou	NU
Seminou	SN
D'origen sense reformar	DO
Algunes reformes necessàries	RN
Totalment reformat	TR
A reformar obligatòriament	AR

Taula 4.6.: Categorització de l'estat de l'habitatge.

La qualitat va quedar definida com a: alta, mitjana – alta, mitjana, mitjana – baixa i baixa.

Per desenvolupar d'una manera entenedora el procés d'obtenció de l'ARQ, el punt de partida havia de ser l'estudi de les corbes que definien l'ARQ de l'immoble totalment reformat i el de la corba que definia l'ARQ de l'immoble a reformar obligatòriament, ja que aquestes corbes havien d'establir l'interval en què es podien moure la resta de corbes.

A les figures 4.2. i 4.3. l'antiguitat és representada a l'eix d'abscisses i l'ARQ a l'eix d'ordenades.

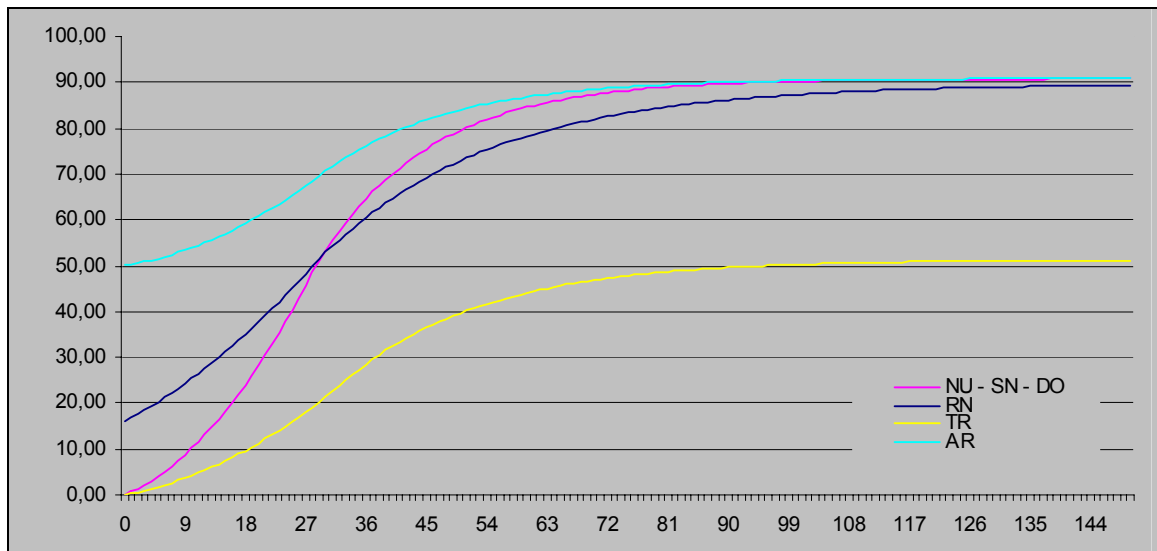


Figura 4.2.: Corbes de l'ARC, segons estat.

Si es prenia la qualitat com a únic objecte d'estudi, les corbes a definir es poden veure a la figura 4.3.

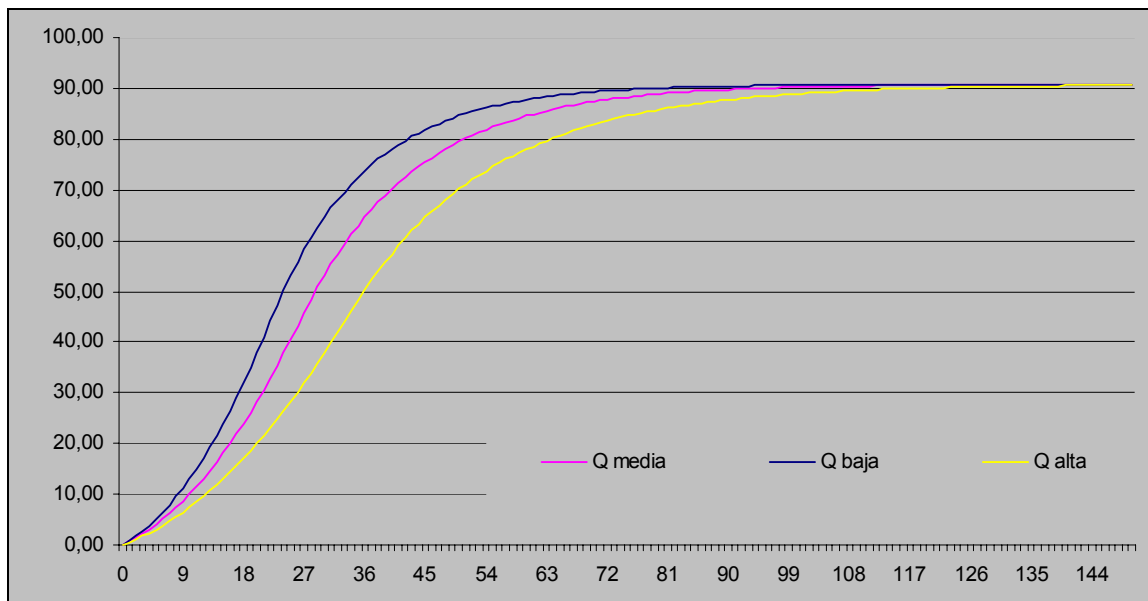


Figura 4.3.: Corbes de l'ARC, segons qualitat.

Quan finalment es va realitzar la modelització conjunta dels conceptes 'qualitat' i 'estat', el resultat va ser la composició de les dues corbes definides anteriorment per a cada una de les combinacions. El resultat mostra que per cada un dels estats definits s'obté una família de corbes en funció de la qualitat.

En la figura següent es representen les qualitats: alta, mitjana i baixa per a totes les definicions d'estat.

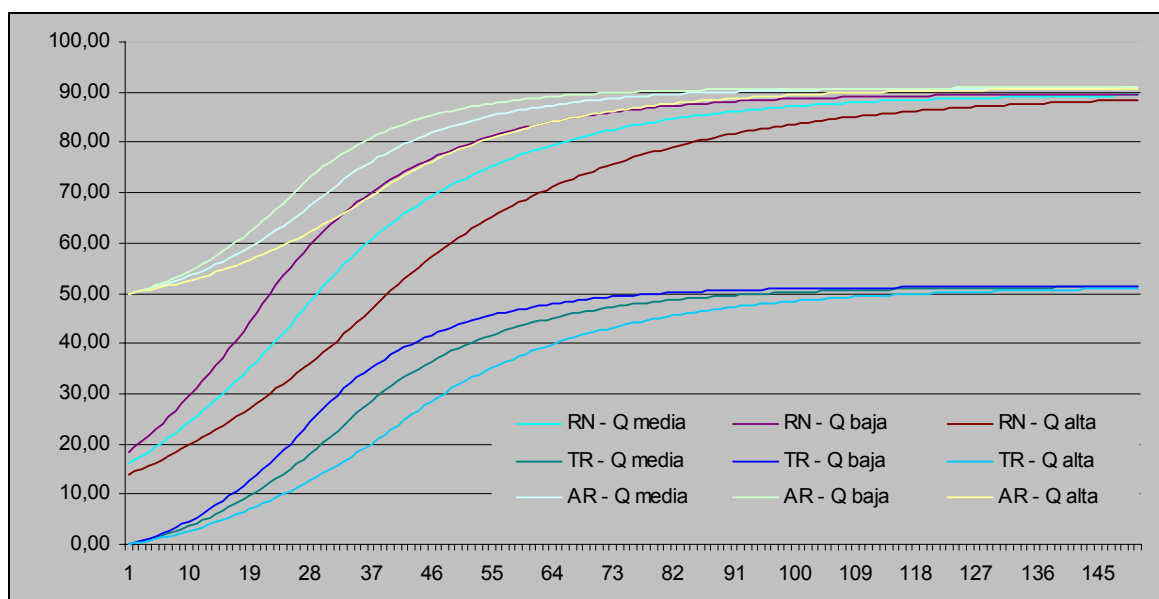


Figura 4.4.: Corbes de l'ARC, segons estat i qualitat.

El diferencial entre l'ARC de l'objecte i l'ARC de cada una de les mostres, multiplicat per una funció de cost servirà per a homogeneïtzar les diferències de qualitat i estat.

El CEQM (Barris, 2007) per la seva part realitza un desenvolupament molt més detallat a partir d'un mètode constructivista, basat en la pèrdua de valor dels acabats en base a les seves vides útils i a les funcions de pèrdua definides.

4.5.3. Homogeneïtzació del nombre de banys

El diferencial de valor provocat per la diferència entre el nombre de banys de l'objecte i el nombre de banys de les mostres es va construir a partir d'una funció d'utilitat, cost i el percentatge de cost dedicat als banys.

La taula d'utilitats dels banys obtinguda a partir de les enquestes realitzades és la següent:

Banys	Utilitat
1	1
2	2
3	2,2
més	2,2

Taula 4.7.: Relació de nombre de banys i utilitat

4.5.4. Homogeneïtzació del nombre de dormitoris i la seva relació amb la superfície.

Per homogeneïtzar la diferència de dormitoris es va utilitzar, en funció del coeficient d'utilitat, el percentatge del cost dedicat a les habitacions i el cost total. La taula d'utilitats va ser generada pel cos tècnic de CATSA.

4.5.5. Homogeneïtzació de la planta i existència d'ascensor

El procés d'homogeneïtzació de la planta i l'ascensor encara es troba en procés de desenvolupament a hores d'ara i no ha pogut ser utilitat en el càlculs inicials d'aquesta tesi.

4.5.6. Càlcul del valor homogeneïtzat

Un cop finalitzat el procés d'homogeneïtzació de cada una de les característiques el resultat és un conjunt de diferències de valor provocades a partir de l'homogeneïtzació. Aquestes diferències són el resultat de l'estudi realitzat pels serveis tècnics de CATSA.

Diferencial de valor produït per:		
Diferències de superfície	D_{sup}	$Abs(D_{sup})$
Diferència en el nombre de banys	D_{ban}	$Abs(D_{ban})$
Diferència en el nombre d'habitacions	D_{hab}	$Abs(D_{hab})$
Diferències d'ARC	D_{arc}	$Abs(D_{arc})$
Diferències de qualitat	D_{qua}	$Abs(D_{qua})$
Diferències de planta i ascensor	D_{asc}	$Abs(D_{asc})$

Taula 4.8. Diferencials per característiques.

Aleshores, el valor homogeneïtzat de les mostres s'obté de la manera següent:

$$VH_i = PM_i + D_{i,sup} + D_{i,ban} + D_{i,hab} + D_{i,arc} + D_{i,qua} + D_{i,asc}$$

VH_i és el valor homogeneïtzat de la mostra i PM_i és el preu inicial de la mostra.

Aquest valor correspon ara a habitatges 'idèntics' a l'objecte a homogeneïtzar.

4.6. Estudi de comparabilitat i homogeneïtzació global

Els procediments d'homogeneïtzació de les característiques constructives es van aconseguir gràcies al departament tècnic de CATSA que va dur a terme una sèrie d'estudis a partir d'informacions extretes de publicacions com el *BOC*, el *Construct* o el *m²*, a més de realitzar una anàlisi exhaustiva de les característiques en base a productes d'obra nova.

El conjunt de procediments presentats en els apartats 4.5.1. – 4.5.6. van permetre calcular per a cada una de les característiques les diferències monetàries provocades per la homogeneïtzació. La suma de totes aquestes diferències monetàries és la homogeneïtzació a realitzar per aconseguir habitatges 'iguals' en característiques físiques, d'estat i qualitat.

$$D_{i,sup} + D_{i,ban} + D_{i,hab} + D_{i,arc} + D_{i,qua} + D_{i,asc} = VH_i - PM_i$$

Per estudiar si el procediment de selecció de comparables i el procediment d'homogeneïtzació plantejat per parts es corresponien amb els procediments realitzats pels tècnics es va realitzar un exercici d'anàlisi de la comparació i de l'homogeneïtzació amb vint-i-un tècnics de CATSA.

L'exercici el van constituir 120 processos de comparació, separats en vint fitxes amb sis comparacions per fitxa.

Els resultats més destacables de l'anàlisi van ser els següents:

- Quan l'acord en la comparabilitat objecte-mostra entre els tècnics superava el 67%, aleshores el percentatge d'acord entre els tècnics i el sistema expert (S.E.) assolía el 86'6%.
- La concordança en signes respecte al valor de l'homogeneïtzació entre la mitjana dels tècnics i el valor decidit pel sistema expert va ser del 89'93%. Valors molt semblants es van obtenir en utilitzar la mediana (88'1%).

Objeto 80m² 3 hb 1 bñ 3º Asc 1950 Algunas reformas necesarias 		Ficha Nº: 1 de 20 Valor medio del Objeto = 250.000 euros siguiente Comentarios, reglas aplicadas (opcional)			
Muestra 1 85m² 3 hb 1 bñ 3º Asc 1950 Algunas reformas necesarias 	Muestra 2 120m² 3 hb 1 bñ 3º Asc 1950 Algunas reformas necesarias 	Muestra 3 100m² 3 hb 1 bñ 3º Asc 1950 Algunas reformas necesarias 	Muestra 4 50m² 3 hb 1 bñ 3º Asc 1950 Algunas reformas necesarias 	Muestra 5 60m² 3 hb 1 bñ 3º Asc 1950 Algunas reformas necesarias 	Muestra 6 75m² 3 hb 1 bñ 3º Asc 1950 Algunas reformas necesarias
Comparable? <input type="checkbox"/> No comparable <input checked="" type="checkbox"/> No comparable <input type="checkbox"/> No comparable <input checked="" type="checkbox"/> No comparable <input type="checkbox"/> No comparable <input type="checkbox"/> No comparable					
Homogeneización La muestra es MEJOR La muestra es MEJOR La muestra es MEJOR La muestra es PEOR La muestra es PEOR La muestra es PEOR					
Ajuste en euros Máximo <input type="text" value="12000"/> Mínimo <input type="text" value="6000"/>		Ajuste en euros Máximo <input type="text"/> Mínimo <input type="text"/>		Ajuste en euros Máximo <input type="text" value="30000"/> Mínimo <input type="text" value="18000"/>	
Ajuste en euros Máximo <input type="text"/> Mínimo <input type="text"/>		Ajuste en euros Máximo <input type="text"/> Mínimo <input type="text"/>		Ajuste en euros Máximo <input type="text" value="18000"/> Mínimo <input type="text" value="30000"/>	
Ajuste en euros Máximo <input type="text"/> Mínimo <input type="text"/>		Ajuste en euros Máximo <input type="text"/> Mínimo <input type="text"/>		Ajuste en euros Máximo <input type="text" value="6000"/> Mínimo <input type="text" value="12000"/>	

Figura 4.5.: Pàgina de recollida de dades.

- A partir de les mostres considerades comparables pel sistema expert es va realitzar un contrast d'hipòtesi de comparació de dades aparellades de la variable 'variació del preu per metre quadrat':

$$H_0 : \mu_{tecnicos} = \mu_{S.E.}$$

$$H_1 : \mu_{tecnicos} \neq \mu_{S.E.}$$

No es va poder rebutjar la hipòtesi nul·la, i per tant, es va poder considerar que no existien diferències estadísticament significatives entre la manera de procedir dels conjunt dels tècnics i la manera de procedir del S.E.

El coeficient de correlació entre el valor mitjà dels tècnics i el valor del S.E. per la variable 'variació del preu metre quadrat' va ser del 0,89.

- Es va realitzar el mateix contrast tècnic a tècnic. Dels 21 contrastos d'hipòtesis realitzats, el 80'9% (17) no es va poder rebutjar la hipòtesis nul·la d'igualtat de mitjanes.

En funció d'aquests resultats es va establir els valors pels diferents paràmetres del procés d'homogeneïtzació.

Es va establir la condició que definia la comparabilitat entre mostra i objecte a partir dels diferencials de valor produïts a l'homogeneïtzació. Les condicions que deriven del conjunt de tests realitzats es mostren a la taula 4.9.

$$\text{Si } (Abs(VH_i - P_i)/P_i > 0.17 \text{ O } (Abs(D_{i,sup}) + Abs(D_{i,ban}) + Abs(D_{i,hab}) + Abs(D_{i,arc}) + Abs(D_{i,qua}) + Abs(D_{i,asc}))) > 0.32 \cdot P_i \text{ ALESHORES la mostra no és comparable.}$$

Taula 4.9: Regla d'homogeneïtzació acceptable.

4.7. Creació d'una metodologia d'homogeneïtzació espacial

Les característiques de les dades que participen en el procés d'homogeneïtzació són les següents:

- Totes les mostres són de la mateixa tipologia que l'objecte a taxar.
- Les mostres es troben a una distància màxima de 2000 metres de l'objecte, que pot venir donat per un model de densitats o establint un nombre màxim de dades a utilitzar.
- Totes les mostres han superat el filtre per superfície.
- Totes les mostres han superat el filtre per antiguitat.
- Totes les mostres han superat el filtre basat en l'existència d'ascensor i planta.
- La variació entre el valor homogeneïtzat i el valor real no supera el 17% en valor absolut i la variació no supera el 32% del valor absolut de cada una de les parts (taula 4.9).

S'ha treballat amb diverses maneres d'homogeneïtzació de la localització en funció dels diferents mètodes d'estimació i interpolació espacial i dels processos de validació creuada presentats en el capítol 3, amb l'objectiu de trobar una metodologia d'obtenció del valor de taxació.

Es van definir tres algorismes de treball per tal de comparar la precisió entre ells.

El primer procés executa tots els mètodes espacials una sola vegada i en base a un procés de validació creuada escull com a solució l'estimació resultant del mètode amb menor coeficient de validació creuada.

El segon procés busca la solució a partir d'un procés iteratiu. La definició d'aquests processos va ser la següent:

El criteri de selecció inicial és la proximitat en distància (distància mínima).

Els criteris de parada es limita a un màxim de 5 iteracions o al compliment dels criteris, d'aquesta manera en el pitjor dels casos una de les mostres

escollides inicialment en base a la distància quedarà en el conjunt mostral final.

El procés iteratiu s'aturarà abans de les 5 iteracions si les validacions creuades de les 6 mostres seleccionades compleixen el criteri establert per al EVC_{4i} i el conjunt de mostres compleix amb el criteri establert per l' EVC_4 .

Criteri de substitució: en cas de no complir-se els criteris d'acceptació de les mostres, en cada iteració es substituirà una, i només una, de les sis mostres utilitzades durant el procés de càlcul.

Les regles de substitució de mostres podien ser de dos tipus:

- Eliminar aquella mostra amb major error de validació creuada.
- Eliminar la mostra més propera que no supera el criteri de validació creuada establert.

La primera regla pot provocar que no es produeixin millores en les diferents iteracions si les mostres amb major error són poc influents en el càlcul de l'estimació. L'avantatge d'aquesta regla és que no cal definir un criteri de validació creuada.

L'avantatge de la segona regla d'eliminació és que pot eliminar mostres influents amb errors superiors als desitjats. El desavantatge és que pot provocar la no eliminació de mostres amb errors grans i influència elevada.

Els criteris d'inclusió de noves mostres van ser els següents:

- Proximitat a l'objecte. Inclou en la selecció de càlcul la següent mostra més propera que prèviament hagi superat tots els criteris de comparabilitat i homogeneïtzació.
- Criteri de similitud. S'escull aquella mostra que dins del conjunt de mostres comparables i homogeneïtzables tingui valor de $abs(P T_i - V H_i) / P T_i$ inferior al 0,12.

Amb aquest procediment es va aconseguir:

- Millorar la selecció d'ofertes.
- Eliminar ofertes que es troben fora de mercat.

Donat que inicialment es va realitzar una selecció per proximitat sota la restricció $abs(PT_i - VH_i)/PT_i \leq 0,17$ és possible que una oferta que es trobi fora de mercat entri en el conjunt de mostres a utilitzar per a la realització de l'estimació del valor de l'objecte. La identificació d'aquesta mostra com a fora de mercat es realitzarà a partir del procediment de validació creuada que donarà un resultat molt elevat en el rati:

$$abs(VH_i - VCH_i)/VH_i$$

PT_i : Preu trobat a l'oferta per a la mostra i.

VH_i : Valor homogeneïtzat per a la mostra i.

VCH_i : Valor homogeneïtzat obtingut a partir de la validació creuada.

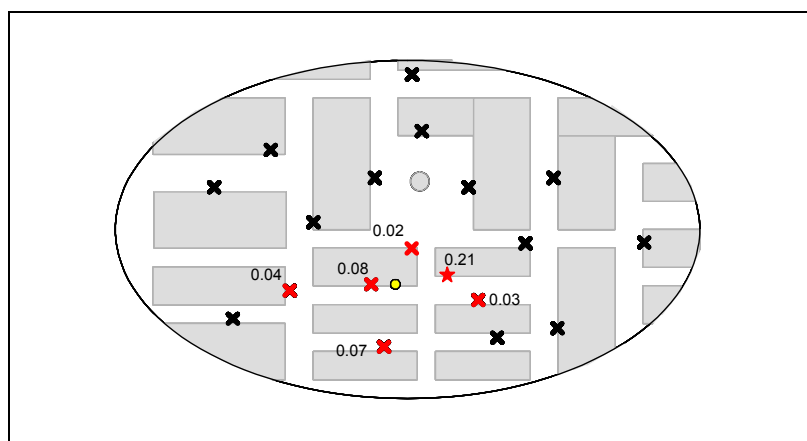


Figura 4.6: Oferta fora de mercat que participa en el procés d'estimació (1/2).

L'oferta fora de mercat és eliminada i entra la oferta més propera que compleix amb les restriccions.

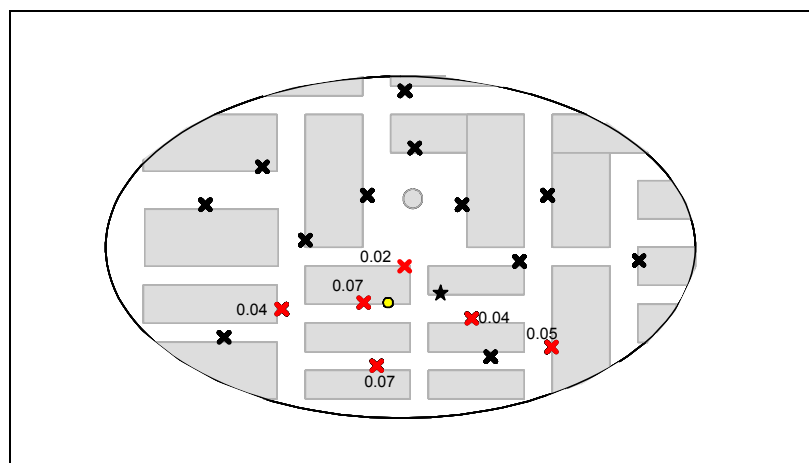


Figura 4.6: Oferta fora de mercat que participa en el procés d'estimació (2/2).

- Detecció de comportaments de mercats diferents.

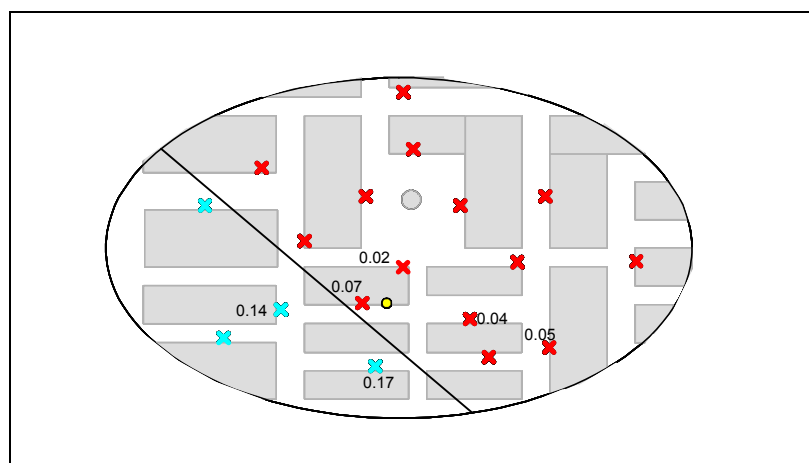


Figura 4.7: Existència de dos mercats propers.

La solució consistirà en l'eliminació del mercat que provoca major error.

Per tant, es considerarà que una mostra es troba fora de mercat quan aquesta no pertany al mateix pla de valors que la resta de les mostres escollides.

4.7.1. Primer procés – comparació de mètodes

El primer procés consisteix en la realització de totes els mètodes d'interpolació i predicció plantejats en el capítol 3.

Per a conjunts mostrals d'entre 6 i 14 dades es compararan els mètodes de la distància inversa, el mètode de les funcions radials i el mètode de les mitjanes mòbils.

Per a conjunts mostrals d'entre 15 i 32 s'incorporarà als mètodes anteriors el mètode basat en el *kriging* sota hipòtesi de isotròpic.

Per a conjunts de més de 32 dades s'incorporarà als mètodes anteriors el mètode basat en el *kriging* sota hipòtesi de anisotròpic.

En cas d'obtenir més de 6 mostres comparables, aleshores es durà a terme el procés següent:

es seleccionen les 6 mostres més properes i es calculava:

- les mitjanes mòbils
- la distància inversa
- les funcions radials

Si el nombre de mostres obtingudes és major de 15 mostres:

- es calcula el variograma omnidireccional
- amb les 6 mostres més properes es realitza un càlcul mitjançant *kriging*.

Si el nombre de mostres obtingudes es major de 32 mostres:

- Es calculen els variogrames direccionals en les direccions de 22,5°; 67,5°; 112,5°; 157,5°.
- Es troba la direcció de màxima dependència i s'aplicava la transformació de coordenades corresponent.
- Es troba el nou variograma omnidireccional en funció de les dades transformades.

- Es seleccionen les 6 dades més properes en funció de la distància transformada.
- Es realitza el càlcul mitjançant *kriging*.

La qualitat de les diferents solucions s'ha d'avaluar a partir d'un procés de validació creuada de les mostres, dels quals es va seleccionar el mètode amb validació creuada menor.

4.7.2. Segon procés – procés iteratiu de recerca de solucions

Les mostres utilitzades durant el procés de valoració són ofertes i com a tals el seu preu d'oferta pot distar del seu valor real. És per aquesta característica de les ofertes que es va considerar que no totes les ofertes són vàlides per a realitzar una valoració.

En aquest segon procés s'intentarà a partir d'un procediment iteratiu la identificació de les ofertes que es troben 'fora de mercat'. Es consideraran ofertes 'fora de mercat' aquelles mostres que incompleixen el criteri de validació creuada. Si una mostra incompleix el criteri de validació, aleshores es pot considerar que aquesta mostra no es troba en el mateix pla de valors que la resta de las mostres seleccionades, i per tant, que es troba fora de mercat.

El procediment utilitza els resultats de la validació creuada per estudiar la seva adequació al comportament espacial esperat.

El mètode de validació creuada escollit és el basat en el EVC_4 , ja que és el que ha donat resultats més estables en el conjunt de proves realitzades a CATSA, es pot veure un exemple del procediment de selecció a l'apartat 5.3.4.

El procediment presentat es realitzarà dues vegades en cada iteració, una per a les mostres que utilitzen els mètodes que no necessiten cap transformació de les dades: *kriging* isotròpic, mètode de la distància inversa, mètodes de les mitjanes mòbils i el mètode basat en les funcions radials. El procediment es repetirà per a les mostres utilitzades pel mètode *kriging*

anisotròpic, que degut a les transformacions espacials realitzades no són les mateixes que per als altres mètodes.

El procés consta dels següent passos:

1. La primera iteració es realitza a partir del procediment plantejat a l'apartat 4.7.1.
2. Una vegada obtinguda la solució caldrà cercar la mostra de menys qualitat i eliminar-la. Aquest procés es realitzarà a partir de la validació creuada de les mostres i en funció del mètode que minimitzi l'error en aquesta iteració pels mètodes isotròpic, de mitjanes mòbils, de funcions radials i distància inversa.
 - a. Si el EVC_4 de la mostra i -èsima és superior a 0,06, aleshores aquesta mostra serà candidata a sortir.
 - b. Entre totes les mostres candidates a sortir s'eliminarà aquella que es trobi a una distància menor.
 - c. En cas de cap mostra tingui un EVC_4 per sota del 0,06 aleshores s'eliminarà aquella mostra amb EVC_4 major.
3. La mostra escollida per participar a la següent iteració del procés es seleccionarà a partir del procediment següent:
 - a. D'entre totes les mostres amb una diferència entre el seu preu inicial i el seu preu homogeneïtzat inferior al 6% s'escollirà aquella que es trobi més propera a l'objecte a valorar.
 - b. Si no existeixen mostres amb diferencials d'homogeneïtzacions inferiors al 6%, aleshores s'anirà incrementant decimal a decimal fins a trobar la primera mostra propera amb homogeneïtzació mínima.
4. Un cop introduïda la nova mostra es repetirà el procediment de l'apartat 4.7.1.

Aquest procediment finalitzarà després de 5 iteracions.

Un cop finalitzades les cinc iteracions, s'hauran obtingut 25 solucions amb els seus corresponents EVC_4 . S'escollirà aquella solució amb menor EVC_4 .

Amb aquest procediment s'espera no només seleccionar correctament ofertes de mercat sinó també eliminar aquelles ofertes que es trobin fora de mercat. Les mostres que suposem fora de mercat són aquelles que no es troben en el mateix pla de valors que la resta de mostres.

En situacions d'anisotropia la transformació provoca una selecció diferent, i serà el càlcul de l'EVC₄ el que decideixi quin dels mètodes dona solucions més adients.

En situacions d'isotropia, els mètodes anisotròpics -els valors resultants del càlcul dels variogrames- donaran solucions semblants, fet que provocarà que les transformacions de coordenades siguin mínimes, i consegüentment, la selecció de les mostres i de les distàncies de treball seran quasi idèntiques, donant solucions molt semblants. Els errors que provoquin aquest càlculs també seran semblants.

4.8. Recerca de la funció de valor

Una forma diferent de calcular el valor podria ser a partir de la generació de la funció de densitat empírica del valor. El problema d'execució d'aquest procés ve donat per la despesa en temps de computació.

El procediment utilitzat s'exposa a continuació:

- Es seleccionen les 30 mostres més properes dins del conjunt de mostres que han superat tots els filtres indicats en el punt 9.6.1.
- Es generen tots els possibles remostreigs *bootstrap* sense repetició, és a dir s'aplica el mètode de les permutacions per a conjunts de 6 elements. Aquest càlcul generarà més de 500.000 permutacions diferents (Hesterberg, 2003).
- Es calcula pel mètode seleccionat l'estimació del valor per a totes i cada una de les mostres donant per a cada una de les mostres una solució.

Amb les més de cinc-centes mil solucions s'acaba produint una funció de probabilitat empírica del valor, a partir de la qual es poden extreure tota mena d'estimadors.

El procediment bootstrap es realitza en base al mètode escollit en l'execució de les algorismes presentats als apartats 4.7.1. i 4.7.2. D'aquesta manera el resultat obtingut a partir del procediment de càlcul del valor es podrà sobreposar directament sobre l'histograma de la funció de valor.

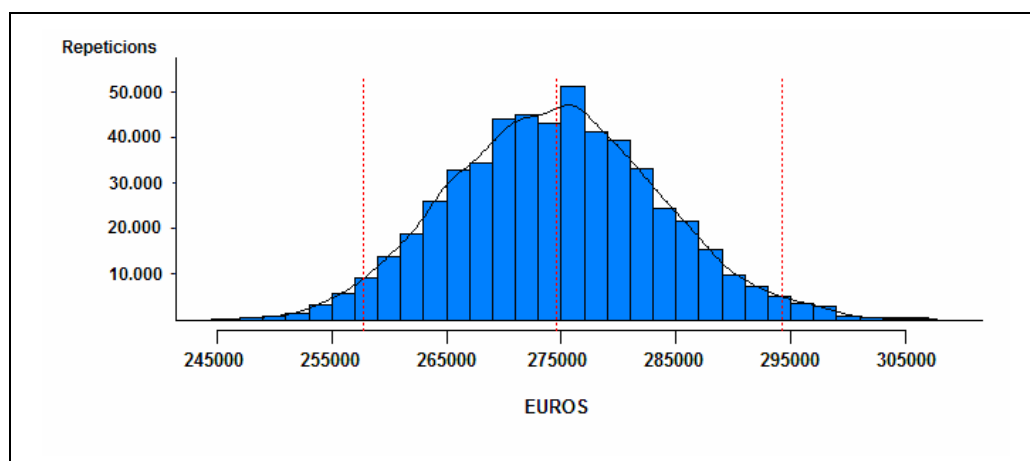


Figura 4.8.: Representació de distribució a partir de l'estimació per *bootstrap*.

L'histograma representarà la funció de densitat empírica obtinguda a partir de les mostres properes i representa totes les possibles permutacions de mostres que es podrien haver donat en base a la informació existent.

El procediment *bootstrap* aportarà tots els estadístics necessaris per a la realització d'interval de confiança i contrastos d'hipòtesi.

5. Funcionament i resultats

5.1. Introducció

Els procediments teòrics plantejats en el capítol tercer i que han estat definits com a processos de treball en el capítol quart necessiten d'un conjunt de proves que permetin validar-ne el funcionament, estudiar-ne els avantatges i definir noves necessitats amb vistes a treballs futurs.

En aquest capítol cinquè es presenten algunes de les proves que van permetre concloure amb èxit aquesta tesi.

Les bases de dades emprades per a realitzar les proves i els exercicis d'aquest capítol són les antigues bases de CATSA modificades adientment per tal d'aconseguir una definició correcta d'estat i qualitat.

Quan l'aplicatiu nou de CATSA es trobi en ple funcionament i la base estigui completament renovada caldrà efectuar un nou conjunt de proves per a determinar el guany en qualitat que suposa treballar amb la base optimitzada.

En un futur les bases de valoracions incorporaran reformes i estats de qualitats i materials dels diferents elements de l'habitatge.

El primer estudi que es va realitzar va tenir com a objectiu l'obtenció d'un sistema de selecció de mètodes i d'avaluació de l'error (apartat 5.3). Per realitzar aquest estudi va caler treballar amb resultats de valoracions realitzades a CATSA en lloc de mostres perquè fos possible avaluar l'error d'estimació. Sota aquesta condició es va suposar que els valors de la base eren vàlids, i per tant, que no calia la utilització de mètodes iteratius per a assolir els objectius d'aquest primer estudi.

Els resultats obtinguts van permetre definir el procediment de selecció del mètode d'estimació entre el conjunt de mètodes utilitzats.

Seguidament es va elaborar un procediment d'estimació del valor en base a les valoracions ja solucionades per CATSA (apartat 5.4.). La intenció d'aquest exercici era mostrar els avantatges del procés de selecció utilitzat i del algoritme iteratiu plantejat.

En aquest procés tots els elements formaven part tant del conjunt de mostres com del conjunt d'objectes a valorar. D'aquesta manera es va poder estudiar l'error de taxació produït en els diferents moments del procés iteratiu.

A l'apartat 5.4. es presenta un estudi de comprovació que permetrà avaluar la qualitat dels resultats obtinguts i les millores assolides a partir d'aplicar l'algoritme iteratiu.

El procediment d'estimació desenvolupat a partir de taxacions validades a l'apartat 5.4. és idèntic a l'utilitzat habitualment pels tècnics validadors de CATSA com a sistema d'ajuda a la valoració.

A l'apartat 5.5. s'exposa un procés de valoració basat en ofertes de mercat realitzat sobre la zona del mercat de Sant Antoni. Com s'ha esmentat en l'apartat 4.7.2., la utilització d'ofertes de mercat exigeix l'ús d'un sistema iteratiu de recerca de mostres, ja que cal identificar i eliminar les ofertes que es troben fora de mercat, així com aquelles que pertanyen a mercats diferents al de l'objecte a valorar. S'hi presenten dos resultats amb la intenció d'establir un estudi de comparació de la qualitat de les solucions

obtingudes. El primer resultat s'ha obtingut mitjançant la utilització del procés no iteratiu de comparació de mètodes (apartat 5.5.2.2), mentre que el segon és el resultat de l'execució del procés iteratiu amb cinc iteracions (apartat 5.5.2.3). En el procés com a tal es definirà únicament un objecte a valorar que a més es situarà en localitzacions diferents amb la intenció d'elaborar un mapa de preus per a la definició d'habitatge especificat.

En l'exemple de l'apartat següent (5.6.) es mostra un procés de valoració sobre un únic bé immoble. Aquest procés serà el més habitual a CATSA i equival al procés de càlcul que ha de realitzar el tècnic validador. Donada la importància d'aquest procés se n'exposaran amb ple detall els resultats juntament amb els resultats de la primera iteració i de la funció empírica d'estimació del valor generada a partir del remostreig *bootstrap*.

La informació de sortida ens donarà les mostres utilitzades, el valor estimat a partir d'aquestes i la funció de probabilitat empírica del valor obtinguda a partir de les 15 mostres més properes.

5.2. Definició de les bases

Durant el procés de càlcul i de proves es va treballar amb dues bases diferents. La primera base enclouïa valoracions realitzades a CATSA i va servir per a determinar l'indicador de la qualitat del mètode d'estimació en funció de la validació creuada i per estudiar la qualitat dels resultats obtinguts a partir de l'algoritme iteratiu. La segona base, formada per mostres de mercat, va servir per a determinar la capacitat del sistema per a obtenir solucions.

Camp	Tipus de camp	Possibles valors
Coordenada x	real	
Coordenada y	real	
Superfície	real	
Nombre de banys	enter	
Nombre de dormitoris	enter	
Any de construcció	enter	
Estat de conservació (calculat a partir dels camps de l'aplicatiu actual)	text	Nou, seminou, d'origen sense reformar, algunes reformes necessàries, totalment reformat, a reformar obligatòriament
Qualitat (calculat a partir dels camps de l'aplicatiu actual)	text	Alta, mitjana-alta, mitjana, mitjana-baixa, baixa
Preu d'oferta de la mostra / valor de taxació	real	
Ascensor	lògic	
Planta (calculat)	enter	

Taula 5.1: Definició dels camps utilitzats en el procés d'homogeneïtzació.

L'única diferència entre la informació proporcionada per la base de valoracions respecte a la informació proveïda per la base de mostres és que a la base de validacions el camp 'valor de l'immoble' assenyala el valor de taxació, mentre que a la base de mostres aquest mateix camp indica el preu de venda presentat per l'ofertent.

Una vegada estigui en funcionament l'aplicatiu de CATSA, la informació continguda a la base de valoracions serà molt més extensa, ja que aportarà informació sobre les reformes realitzades i sobre l'estat de l'edifici.

5.3 Estudi per a l'obtenció d'un mètode de selecció de la solució a partir d'una funció basada en la validació creuada

5.3.1. Objectiu de l'estudi

L'objectiu d'aquest estudi va ser la recerca d'una funció basada en els mètodes de validació creuada (Fox, 1997) que permetés escollir entre els diferents mètodes d'estimació plantejats.

5.3.2. Definició del marc de treball

Per realitzar aquest exercici es van utilitzar nombrosos conjunts de taxacions de diferents localitzacions i es van elaborar diverses definicions de l'índex de validació creuada.

Les taxacions que es van utilitzar per a la realització d'aquest estudi van ser extreptes de l'antiga base de CATSA, tot i conèixer els problemes existents en aquestes bases (P. Garcia, 2004).

Per minimitzar els problemes derivats de la manca d'informació a les bases es van fixar un seguit de filtres (Taula 5.2.).

Les expressions de les diferents funcions estudiades van ser les següents:

- $$EVC_1 = \sum_{i=1}^{Nc} (Vh_i - Vhe_{-i})^2 \cdot p_i$$
- $$EVC_2 = \sum_{i=1}^{Nc} \left(\frac{Vh_i - Vhe_{-i}}{Vh_i} \right)^2 \cdot p_i$$
- $$EVC_3 = \sum_{i=1}^{Nc} \text{abs}(Vh_i - Vhe_{-i}) \cdot p_i$$
- $$EVC_4 = \sum_{i=1}^{Nc} \text{abs} \left(\frac{Vh_i - Vhe_{-i}}{Vh_i} \right) \cdot p_i$$

$$\begin{aligned}
 \blacksquare \quad EVC_5 &= \sum_{i=1}^{Nc} (Vh_i - Vhe_{-i})^2 \cdot \left(\left(\sum_{i=1}^{Nc} dist_i \right) - dist_i \right) \Big/ \left(5 \cdot \sum_{i=1}^{Nc} dist_i \right) \\
 \blacksquare \quad EVC_6 &= \sum_{i=1}^{Nc} \left(\frac{Vh_i - Vhe_{-i}}{Vh_i} \right)^2 \cdot \left(\left(\sum_{i=1}^{Nc} dist_i \right) - dist_i \right) \Big/ \left(5 \cdot \sum_{i=1}^{Nc} dist_i \right) \\
 \blacksquare \quad EVC_7 &= \sum_{i=1}^{Nc} abs \left(\frac{Vh_i - Vhe_{-i}}{Vh_i} \right) \cdot \left(\left(\sum_{i=1}^{Nc} dist_i \right) - dist_i \right) \Big/ \left(5 \cdot \sum_{i=1}^{Nc} dist_i \right) \\
 \blacksquare \quad EVC_8 &= \sum_{i=1}^{Nc} abs(Vh_i - Vhe_{-i}) \cdot \left(\left(\sum_{i=1}^{Nc} dist_i \right) - dist_i \right) \Big/ \left(5 \cdot \sum_{i=1}^{Nc} dist_i \right)
 \end{aligned}$$

També es va intentar determinar el punt límit a partir del qual no es pogués considerar vàlid el resultat d'un mètode d'estimació. Aquest objectiu no es va poder assolir a partir de les funcions basades en la validació creuada. Els estudis realitzats amb l'objectiu de trobar una relació entre la validació creuada i l'error d'estimació van ser no significatives.

5.3.3. Algorisme de treball

L'algorisme de treball va ser el següent:

1. Es va realitzar un filtre en funció de la distància (2.000 metres).
2. Es van establir els límits de comparabilitat i càlcul de l'ARC de l'objecte.
3. Es va dur a terme un procés d'homogeneïtzació de les mostres.
4. Es va efectuar la recerca de comparables.
5. En cas de disposar de sis o més mostres comparables, es van realitzar els passos següents:
 - 5.1. Selecció inicial de les sis mostres més properes.
 - 5.2. Càlcul de l'estimació del valor a partir del mètode de la distància inversa.
 - 5.3. Càlcul de l'estimació de la validació creuada del mètode de la inversa.

5.4. Càlcul dels estadístics en base a l'estimació de la validació creuada.

Repetició dels punts 5.1, 5.2. i 5.3. per aconseguir els estimadors pels estimadors basats en mitjanes mòbils i funcions radials.

6. En cas de superar les 15 mostres comparables, es van definir els següents passos a realitzar:

6.1. Estudi i eliminació de tendències. En cas de donar-se alguna tendència estadísticament significativa es va treballar amb els residus.

6.2. Realització del variograma.

6.3. Càlcul de l'estimació del valor a partir del mètode del *kriging* isotròpic. En cas d'haver treballat amb els residus, va caler efectuar la conversió a l'escala inicial a partir de l'aplicació dels paràmetres del procés regressió del punt 6.1. sobre la localització de l'objecte.

6.4. Càlcul de l'estimació de la validació creuada per al mètode *kriging* isotròpic.

6.5. Càlcul dels estadístic en base a l'estimació de la validació creuada.

7. En cas de superar les 32 mostres comparables, es van establir els següents passos a realitzar:

7.1. L'estimació per *kriging* anisotròpic s'ha de realitzar a partir de les dades resultants del punt 6.1.

7.2. Realització del variograma direccional.

7.3. Realització de les transformacions de les coordenades de les dades en funció dels resultats dels variogrames direccional.

7.4. Càlcul del variograma a partir de les dades transformades.

7.5. Càlcul de l'estimació del valor a partir del mètode del *kriging* anisotròpic. En cas d'haver treballat amb els residus, caldrà realitzar la conversió a l'escala inicial a partir de l'aplicació dels

paràmetres del procés regressió del punt 7.1. sobre la localització de l'objecte.

7.6. Càlcul de l'estimació validació creuada del mètode del *kriging* anisotròpic.

7.7. Càlcul dels estadístics en base a l'estimació validació creuada.

5.3.4. Proves sobre el mètode de selecció

Del conjunt de proves efectuades en aquest apartat es presenta de forma detallada la prova realitzada a l'Hospitalet del Llobregat.

El marc mostral va ser el següent:

<p>K establerta pels tècnics $> 1'25$.</p> <p>Coefficient de depreciació establert pels tècnics $< 25\%$.</p> <p>Sense especificació de reformes.</p> <p>Habitatges en edificis plurifamiliars sense terrasses, trasters ni garatges.</p> <p>Any d'edificació posterior a 1965.</p> <p>Cost de construcció entre 800 € i 1000 €</p> <p>Procediment realitzat a partir de les valoracions de la base de CATSA.</p>
--

Taula 5.2.: Marc mostral.

El nombre total d'immobles emprats en aquest estudi va ser de 262.

Es va realitzar un estudi de les característiques bàsiques dels immobles amb l'objectiu de conèixer el mercat representatiu de la zona.

	N	Mínim	Mitjana	Mediana	Moda	Màxim	Desviació
Superfície	262	45,6	71,86	70,81	61,00	115,68	11,87
Any d'edificació	262	1960	1968	1966	1960	2007	8,14
Valor m ²	262	2473	3980	3977		5144	154
Valor	262	179584	285918	282507		517023	48593

Taula 5.3.: Característiques de les mostres. Exemple 1 (1/9).

Dormitoris	N	%
1 dormitori	8	3,05
2 dormitoris	55	20,99
3 dormitoris	184	70,23
4 dormitoris	15	5,73

Taula 5.3.: Característiques de les mostres. Exemple 1 (2/9).

Banys	N	%
1 bany	258	96,95
2 banys	8	3,05

Taula 5.3.: Característiques de les mostres. Exemple 1 (3/9).

Sales	n	%
1 sala	257	98,09
2 sales	5	1,91

Taula 5.3.: Característiques de les mostres. Exemple 1 (4/9).

Ascensor	n	%
Amb ascensor	115	43,89
Sense ascensor	147	56,11

Taula 5.3.: Característiques de les mostres. Exemple 1 (5/9).

Qualitat fusteria	n	%
Baixa	29	11,07
Mitjana – baixa	2	0,76
Mitjana	182	69,47
Mitjana - alta	6	2,29
Alta	43	16,41

Taula 5.3.: Característiques de les mostres. Exemple 1 (6/9).

Qualitat cuina	n	%
Baixa	11	4,20
Mitjana – baixa	2	0,76
Mitjana	167	63,74
Mitjana - alta	8	3,05
Alta	74	28,24

Taula 5.3.: Característiques de les mostres. Exemple 1 (7/9).

Qualitat material elèctric	n	%
Baixa	9	3,44
Mitjana – baixa	2	0,76
Mitjana	228	87,02
Mitjana - alta	5	1,91
Alta	18	6,87

Taula 5.3.: Característiques de les mostres. Exemple 1 (8/9).

Qualitat sanitaris	n	%
Baixa	9	3,44
Mitjana – baixa	1	0,38
Mitjana	206	78,63
Mitjana - alta	7	2,67
Alta	39	14,89

Taula 5.3.: Característiques de les mostres. Exemple 1 (9/9).

Un cop realitzades les homogeneïtzacions no es va detectar cap mena de tendència general entre les 'mostres equivalents' (Montero i Larraz, 2006).

Una vegada finalitzat el procés d'estudi de la tendència global aquest procés es va executar únicament la primera iteració de l'algoritme de càlcul.

De les 262 valoracions inicials se'n van poder realitzar un total de 197, per tant, en 65 casos (24,8%) no es van trobar els 6 comparables mínims, dins l'espai proper que permetessin realitzar el procés d'estimació. En aquests casos el sistema hauria de demanar al tècnic que introduís nous comparables amb l'objectiu de poder realitzar la valoració correctament.

Es va realitzar un estudi de correlació entre les solucions seleccionades a partir de l'estimador òptim (resultat del model que minimitza l'error) amb les solucions seleccionades a partir dels diferents estimadors de validació creuada. Els resultats van ser els següents:

Zona	N	ECV ₁	ECV ₂	ECV ₃	ECV ₄	ECV ₅	ECV ₆	ECV ₇	ECV ₈
Hospitalet	197	0,95	0,95	0,95	0,95	0,94	0,94	0,94	0,94

Taula 5.4.: Estudi dels diferents ECV per a la mostra de l'Hospitalet del Llobregat

El mateix procés es va dur a terme per a diferents localitzacions obtenint els següents resultats:

Zona	N	ECV ₁	ECV ₂	ECV ₃	ECV ₄	ECV ₅	ECV ₆	ECV ₇	ECV ₈
Hospitalet	196	0,95	0,95	0,95	0,95	0,94	0,94	0,94	0,94
Barcelona	23	0,95	0,95	0,96	0,95	0,98	0,98	0,96	0,97
Badalona	65	0,90	0,90	0,90	0,90	0,89	0,89	0,89	0,89
Santa Coloma	63	0,76	0,76	0,81	0,81	0,72	0,72	0,73	0,73
Cornellà	14	0,97	0,96	0,96	0,97	0,93	0,93	0,94	0,93
El Prat	16	0,77	0,77	0,77	0,77	0,85	0,85	0,85	0,85
Esplugues	26	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99
Mataró	27	0,98	0,98	0,98	0,98	0,98	0,98	0,98	0,98
Ripollet	20	0,92	0,92	0,91	0,91	0,92	0,92	0,92	0,92
Rubí	19	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96
Sabadell	29	0,97	0,97	0,97	0,97	0,97	0,97	0,97	0,97

Taula 5.5.: Estudi dels diferents ECV per a diferents mostres.

En funció dels resultat es va decidir utilitzar l'ECV₄, tot i que s'hagués pogut triar qualsevol estimador entre l'ECV₄ i l'ECV₈. L'elecció es va fer en base al fet que tant la mitjana com la mediana dels coeficients de correlació obtinguts eren de valor superior a l'EVC₄ que a la resta de ECV's, i que la desviació típica de la mitjana dels coeficients de correlació era la de menor magnitud. A més, es va prendre en consideració el fet que el mínim del coeficient de correlació entre l'estimador òptim i l'ECV₄ era el més elevat.

Reprement l'exemple de l'Hospitalet del Llobregat, es van comparar els resultats obtinguts a partir del selector òptim (resultat del model que minimitza l'error) amb els resultats obtinguts a partir del selector escollit, ECV₄ (taula 5.4.), i els resultats van ser els següents:

	Òptim		ECV ₄	
	n	%	N	%
Menys del 2%	107	54,3	66	33,5
Menys del 6%	168	85,3	152	77,2
Menys del 10%	188	95,4	180	91,4
Més del 10%	9	4,6	17	8,7

Taula 5.6.: Relació percentual entre la selecció òptima i l'ECV₄.

	Òptim		ECV ₄	
	n	%	N	%
Menys de 3000	82	41,6	37	18,8
Menys de 6000	107	54,3	73	37,1
Menys de 12000	154	78,2	131	66,5
Menys de 24000	181	91,9	174	88,3
més de 24000 €	15	7,6	23	11,7

Taula 5.7.: Relació monetària entre la selecció òptima i l'ECV₄.

Un exemple gràfic de la relació entre les solucions aportades pel EVC₄ i els valors taxats es mostra en la figura 5.1.

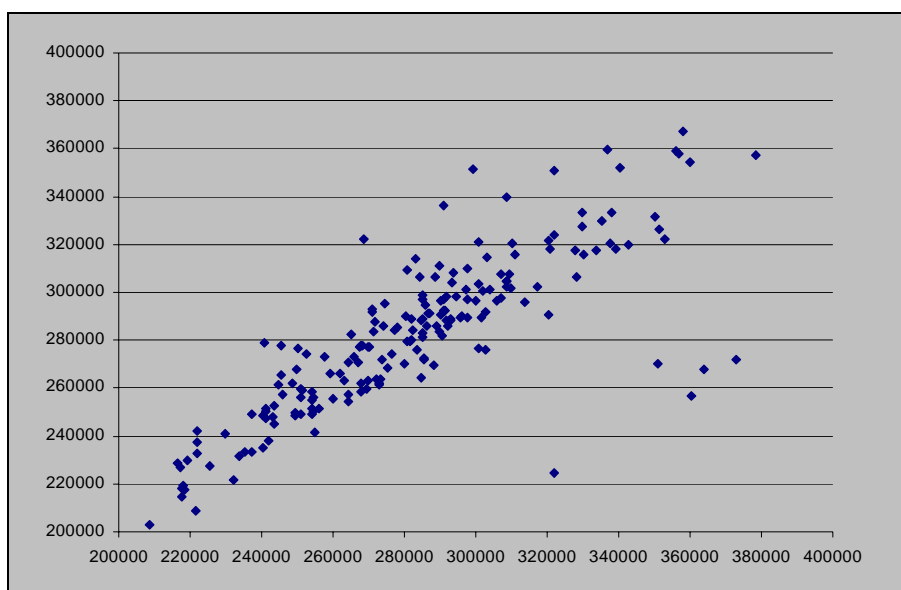


Figura 5.1.: Valor taxat (x) – Valor estimat (y).

5.4. Estudi de la millora de resultats assolida a partir de la utilització de l'algoritme iteratiu de treball plantejat

5.4.1. objectius

L'objectiu d'aquest estudi va ser estudiar les millores que podia suposar l'ús d'un procés iteratiu.

5.4.2. Definició del marc de treball

Es va treballar amb un conjunt de 410 habitatges que constitueixen una selecció aleatòria representativa d'habitatges plurifamiliars ubicats a l'àrea de Badalona, Barcelona, L'Hospitalet del Llobregat, Santa Coloma de Gramenet i Sant Adrià del Besòs, i que, a més, compleixen l'única restricció requerida de no disposar d'elements a valorar addicionals com ara garatges, trasters, terrasses, etc.

Es va dur a terme un procés basat únicament en taxacions validades. Els habitatges del conjunt mostral seleccionat es van utilitzar com a mostres i alhora com a objectes del procés de valoració.

5.4.3. Descripció de la mostra

La descripció de la mostra de treball és la següent:

	N	Mínim	Mitjana	Mediana	Moda	Màxim	Desviació
Superfície	410	33,00	79,42	74,90	73,00	183,85	21,27
Any d'edificació	410	1840	1959	1965	1960	2007	25
Valor m ²	410	2874	4099	3991	-	6424	518
Valor	410	149894	328944	297577	-	1002014	119115

Taula 5.8.: Característiques de les mostres. Exemple 2 (1/11).

Dormitoris	n	%
1 dormitori	21	5,2
2 dormitoris	87	21,2
3 dormitoris	239	58,9
4 dormitoris	62	15,2
5 dormitoris	1	0,4

Taula 5.8.: Característiques de les mostres. Exemple 2 (2/11).

Banys	n	%
1 bany	344	83,90
2 banys	64	15,61
3 banys	2	0,49

Taula 5.8.: Característiques de les mostres. Exemple 2 (3/11).

Sales	n	%
1 sala	394	96,10
2 sales	15	3,66
3 sales	1	0,24

Taula 5.8.: Característiques de les mostres. Exemple 2 (4/11).

Ascensor	n	%
Amb ascensor	217	52,93
Sense ascensor	193	47,07

Taula 5.8.: Característiques de les mostres. Exemple 2 (5/11).

Qualitat fusteria	n	%
Baixa	32	7,80
Mitjana – baixa	14	3,41
Mitjana	321	78,29
Mitjana - alta	7	1,71
Alta	36	8,78

Taula 5.8.: Característiques de les mostres. Exemple 2 (6/11).

Qualitat cuina	n	%
Baixa	12	2,93
Mitjana – baixa	15	3,66
Mitjana	311	75,89
Mitjana - alta	13	3,17
Alta	59	14,39

Taula 5.8.: Característiques de les mostres. Exemple 2 (7/11).

Qualitat material elèctric	n	%
Baixa	10	2,44
Mitjana – baixa	11	2,68
Mitjana	358	87,32
Mitjana - alta	8	1,95
Alta	23	5,61

Taula 5.8.: Característiques de les mostres. Exemple 2 (8/11).

Qualitat sanitaris	n	%
Baixa	13	3,17
Mitjana – baixa	12	2,93
Mitjana	328	80,00
Mitjana - alta	11	2,68
Alta	46	11,22

Taula 5.8.: Característiques de les mostres. Exemple 2 (9/11).

Reformes	n	%
S'han realitzat reformes	40	9,76
No s'han realitzat reformes	370	90,24

Taula 5.8.: Característiques de les mostres. Exemple 2 (10/11).

Poblacions	n	%
Badalona	29	7,07
Barcelona	247	60,24
L'Hospitalet de Llobregat	102	24'88
Sant Adrià del Besòs	8	1'95
Santa Coloma de Gramenet	24	5'85

Taula 5.8.: Característiques de les mostres. Exemple 2 (11/11).

5.4.4. Resultats de l'estudi

En aquest apartat es mostren a mode de resum els resultats obtinguts a la primera iteració i les solucions obtingudes després de realitzar cinc iteracions. Amb aquesta informació serà possible avaluar les millores produïdes pel procés iteratiu.

Es va acceptar com a solució vàlida en cada una de les iteracions del procés tots aquells resultats del procés de valoració que complissin les restriccions següents:

- Cap de les sis mostres utilitzades produeix diferències superiors al 17% entre el seu valor i el valor homogeneïtzat.
- Cap de les mostres supera la distància màxima permesa, entenenent-se aquesta com la distància d'independència (rang) obtinguda a partir del variograma isotròpic –en cas d'existir-ne– o, en cas de no existir-ne, una distància màxima de 850 metres. Aquesta restricció es va aplicar únicament als mètodes directes, ja que els mètodes basats en *kriging* estan capacitats per gestionar aquest aspecte del càlcul.

En base a aquestes conjunt de restriccions es van obtenir les solucions següents:

Iteracions	Solucions vàlides	%
Una iteració	275	67,07
Cinc iteracions	376	91,71

Taula 5.9.: Solucions vàlides obtingudes.

Mètode escollit	Una iteració		Després de cinc iteracions	
	Solucions	%	Solucions	%
Anisotròpic	75	27,27	94	25,00
Dist. inversa	89	32,36	143	38,03
Isotròpic	32	11,64	70	18,62
Mitjanes mòbils	33	12,00	20	5,32
Radials	46	16,73	49	13,03

Taula 5.10.: Mètodes i solucions escollides.

Donat que la metodologia *kriging* requereix un mínim de 15 dades en el cas del mètode isotròpic i de 32 en el cas de l'anisotròpic, el nombre de processos d'estimació realitzats amb aquests procediments va ser inferior als realitzats a partir dels mètodes directes.

Percentils	% Error absolut	
	1 iteració (n=275)	Després de 5 iteracions (n=376)
100%	19,76	21,28
99%	18,24	19,25
95%	13,63	14,49
90%	10,51	11,51
75%	7,67	7,67
50%	4,11	4,22
25%	2,12	1,69
10%	0,98	0,77
5%	0,49	0,32
1%	0,06	0,10
0%	0,05	0,02

Taula 5.11.: Error absolut i percentils.
Tots els casos possibles.

Pel que fa a les valoracions que van obtenir solució a la primera iteració, els resultats van ser els següents:

Percentils	% Error absolut	
	1 iteració (n=275)	Després de 5 iteracions (n=275)
100%	19,76	19,34
99%	18,24	17,91
95%	13,63	13,54
90%	10,51	11,12
75%	7,67	7,13
50%	4,11	4,08
25%	2,12	1,69
10%	0,98	0,77
5%	0,49	0,42
1%	0,06	0,11
0%	0,05	0,05

Taula 5.12.: Error absolut i percentils.
Casos solucionats a la primera iteració.

La correlació existent entre el valor de taxació i els valors estimats va ser la següent:

	1 iteració (n=275)	Després de 5 iteracions (n=376)
Valor real	0,96866	0,97222

Resultats molt semblant s'obtenen si es comparen únicament aquells exercicis de valoració que van obtenir solució a la primera iteració.

	1 iteració (n=275)	Després de 5 iteracions (n=376)
Valor real	0,96866	0,97012

El resultat de la regressió entre el valor real i els valors estimats a partir de la primera iteració s'assenyalen a la taula següent:

Anàlisi de la variància				
Font	DF	Quadrats mitjans	F	p-valor
Model	1	$1,685 \cdot 10^{12}$	4152,14	<0,0001
Error	273	405859223		
Total	274			

R-quadrat	0,9383	R-quadrat ajustat	0,9381
------------------	--------	--------------------------	--------

Paràmetres estimats					
Variable	DF	Paràmetre	Error estàndard	T	p-valor
terme indep.	1	1747	4836,5	0,36	0,7182
Solució inicial	1	1,01513	0,01575	64,44	<0,0001

Taula 5.13.: Relació lineal valor taxat – valor estimat
Primera iteració.

Els resultats de la regressió entre el valor real i els valors estimats a partir de l'algoritme iteratiu es mostra a la taula següent:

Anàlisi de la variància				
Font	DF	Quadrats mitjans	F	p-valor
Model	1	$2,9174 \cdot 10^{12}$	6451,34	<0,0001
Error	374	452226569		
Total	375			

R-quadrat	0,9452	R-quadrat ajustat	0,9451
------------------	--------	--------------------------	--------

Paràmetres estimats					
Variable	DF	Paràmetre	Error estàndard	t	p-valor
Intercept	1	12614	3897,59	3,24	0,0013
Solució inicial	1	0,98197	0,01223	80,32	<0,0001

Taula 5.14.: Relació lineal valor taxat – valor estimat
Resultat després de cinc iteracions.

La taula 5.15. mostra els resultats categoritzats en relació amb l'error absolut existent entre l'estimació i el valor real segons els criteris de CATSA.

Diferència en valor absolut (DVA)		
$0 < DVA \leq 3\%$		Òptima
$3\% < DVA \leq 6\%$		Correcte
$6\% < DVA \leq 10\%$		Acceptable
$DVA > 10\%$		Insuficient

	1 iteració (n=275)	
	Freq.	%
Òptim	97	35,27
Correcte	82	29,82
Acceptable	61	22,18
Insuficient	35	12,73

	Després de 5 iteracions (n=376)	
	Freq.	%
Òptim	149	39,63
Correcte	90	23,94
Acceptable	82	21,81
Insuficient	55	14,63

Taula 5.15.: Solucions categoritzades.

Avaluant els resultats obtinguts, es pot concloure que ambdós procediments funcionen amb una precisió suficientment elevada per ser emprats com elements de suport a la valoració.

Com a aportació principal del mètode iteratiu destaca l'augment del nombre de valoracions realitzades; i en segon terme, s'aprecia una sensible millora en la precisió global del procés.

5.5. Valoracions realitzades a partir d'un objecte de característiques fixes.

5.5.1. Introducció

El procés que s'exposa a continuació es basa en mostres de mercat i té com a objectiu final simular el procediment que realitza el tècnic valorador quan efectua una valoració amb objectiu hipotecari pel mètode de comparació de mercat.

L'anàlisi es va centrar en la zona del mercat de Sant Antoni de Barcelona on conflueixen tres teixits urbans ben dispars: l'Eixample, Ciutat Vella i Sants-Montjuïc.

5.5.2. Exemple: zona mercat de Sant Antoni

A la zona del mercat de Sant Antoni (any 2006) es van obtenir un total de 247 mostres. Aquestes mostres no es van sotmetre a cap mena de restricció ni control.

Es va definir un objecte tipus (Taula 5.15) amb l'objectiu d'utilitzar-lo per a realitzar un seguit de valoracions individuals que permetessin generar un mapa de preus -entre d'altres gràfics- per contrastar de forma gràfica el funcionament del procés.

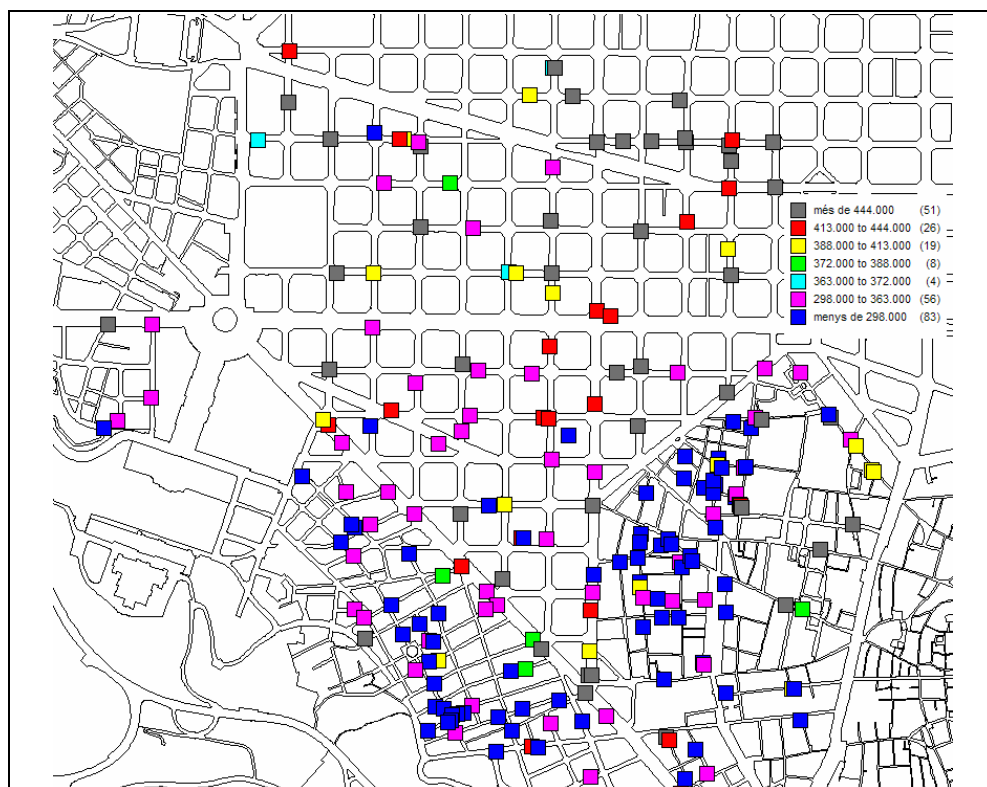


Figura 5.2.: Localització de les mostres inicials amb preus no homogeneïtzats.

Característica	Valor
Superfície	90
Nombre de banys	1
Nombre de dormitoris	3
Any de construcció	1955
Estat de conservació	Algunes reformes necessàries
Qualitat	Mitjana - baixa

Taula 5.16.: Objecte de treball.

Un cop definit l'objecte tipus només 80 mostres van ser acceptades com a comparables per les seves característiques. La identificació sobre el mapa d'aquestes mostres es mostra a la figura 5.3.:



Figura 5.3.: Localització de les mostres comparables al objecte creat.

Les restriccions que es van establir per poder acceptar la solució com a vàlida van ser les següents:

- Que cap de les sis mostres utilitzades no produís diferències superiors al 17% entre el seu valor i el valor homogeneïtzat.
- Que la distància entre la mostra més propera i l'objecte no superés els 350 metres.
- Cap de les mostres supera la distància màxima permesa, entenent-se aquesta com la distància d'independència (rang) obtinguda a partir del variograma isotròpic –en cas d'existir-ne– o, en cas de no existir-ne, una distància màxima de 850 metres. Aquesta restricció es va aplicar únicament als mètodes directes, ja que els mètodes basats en *kriging* estan capacitats per gestionar aquest aspecte del càlcul.

El criteri de distància mínima és un criteri imposat a l'exercici. La metodologia per definir els espais no homogeneïtzables en funció de la distància restarà pendent per desenvolupar en futurs treballs.

Tot aplicant aquestes restriccions es va aconseguir realitzar 300 valoracions d'un total de 328 punts creats.

5.5.2.1. Estadístics bàsics sobre la mostra seleccionada

Els estadístics que defineixen les principals característiques dels conjunt de mostres comparables van ser:

	N	Mínim	Mitjana	Mediana	Màxim	Desviació
Superfície	80	70,00	89,54	84,00	107,00	10,87
Any d'edificació	80	1877	1952	1962	1985	30,20

Taula 5.17.: Característiques de les mostres (1/7).

Dormitoris	n	%
1 dormitori	3	3,75
2 dormitoris	11	13,75
3 dormitoris	54	67,50
4 dormitoris	11	13,75
5 dormitoris	1	1,25

Taula 5.17.: Característiques de les mostres (2/7).

Ascensor	n	%
Sense ascensor	30	37,50
Amb algun ascensor	50	62,50

Taula 5.17.: Característiques de les mostres (3/7).

Bany	n	%
1 bany	57	71,25
2 banys	23	28,75

Taula 5.17.: Característiques de les mostres (4/7).

Qualitat	n	%
Baixa	13	16,25
Mitjana – baixa	18	22,50
Mitjana	49	61,25

Taula 5.17.: Característiques de les mostres (5/7).

Estat	n	%
A reformar obligatòriament	2	2,50
Algunes reformes necessàries	63	78,75
D'origen sense reformar	8	10,00
Totalment reformat	7	8,75

Taula 5.17.: Característiques de les mostres (6/7).

Les característiques de relació amb el mercat van ser les següents:

	N	Mínim	Mitjana	Mediana	Màxim	Desviació
Valor m ²	80	2453	4203	4210	6493	646
Valor total	80	238000	361747	360800	494868	60543
Valor homogeneïtzat	80	212036	369132	366256	559748	58616
VH/m ² habitatge tipus	80	2356	4101	4070	6219	651

Taula 5.17.: Característiques de les mostres (7/7).

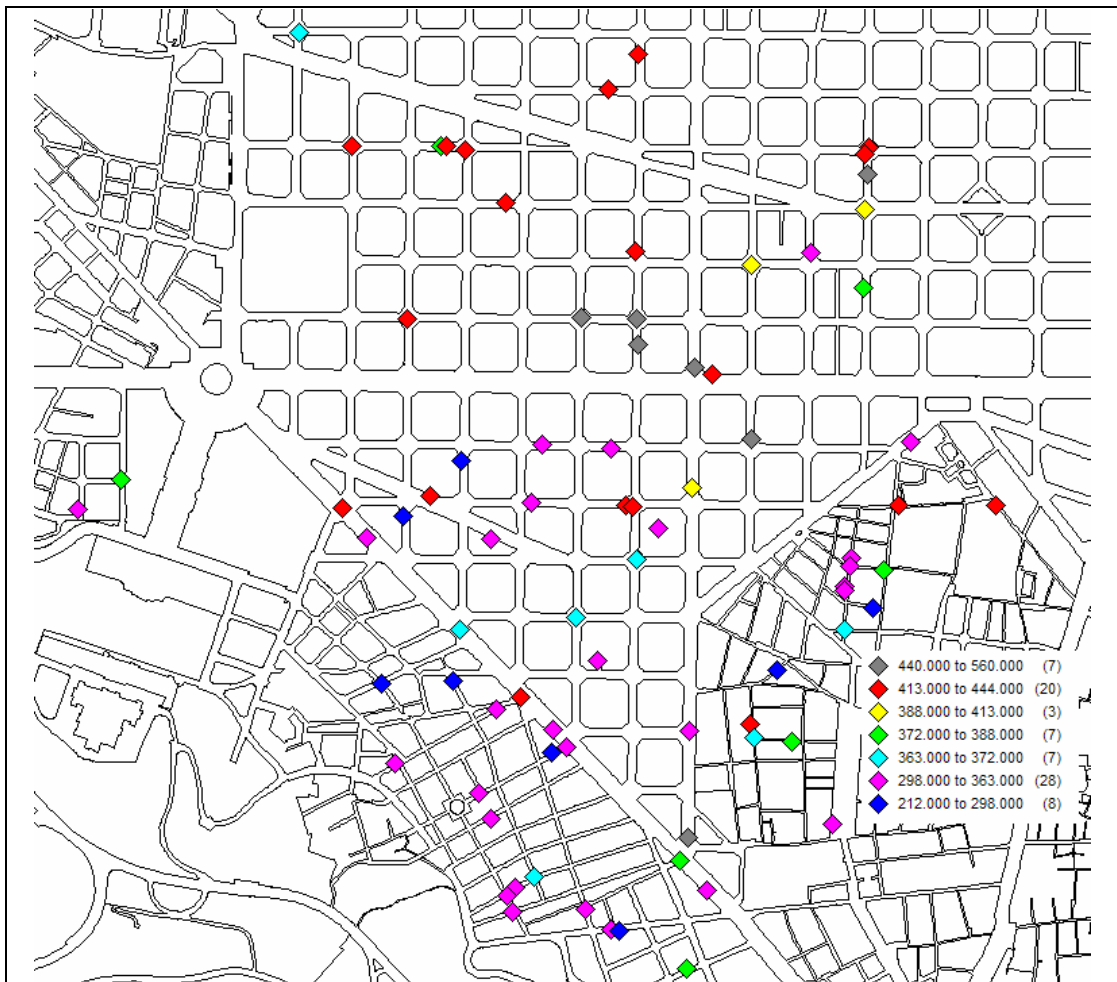


Figura 5.4.: Localització de les mostres comparables homogeneïtzades respecte l'objecte tipus.

Es va dur a terme un estudi de regressió sobre la relació lineal existent entre les variables homogeneïtzades i les coordenades amb la finalitat d'estudiar-ne les tendències globals. Els resultats van ser els següents:

Model	p-valor model	p-valor β_1	p-valor β_2	p-valor β_3
$vh = \beta_0 + \beta_1 x$	0,0045	0,0045		
$vh = \beta_0 + \beta_2 y$	0,0010		0,0010	
$vh = \beta_0 + \beta_1 x + \beta_2 y$	0,0007	0,0537	0,0115	0,0007

Taula 5.18.: Estudi de tendència sobre el conjunt de dades homogeneïtzades.

El fet de treballar amb subconjunt de dades va provocar que les tendències fossin eliminades en la major part dels casos.

5.5.2.2. Resultats per comparació de mètodes

El procés a una iteració té com a finalitat cercar la millor solució entre tots els models provats en base a l'índex de validació creuada EVC_4 .

El mètode amb una única iteració va resoldre 167 (50,91%) processos de valoració.

Mètode	Freq.	%	Validació creuada sobre les mostres escollides				
			5%	25%	50%	75%	95%
Anisotròpic	51	30,54	0,043	0,086	0,144	0,181	0,210
Isotròpic	22	13,17	0,029	0,043	0,832	0,106	0,169
Mitjanes mòbils	18	10,78	0,027	0,076	0,131	0,170	0,214
Distància inversa	55	32,93	0,026	0,043	0,078	0,090	0,150
Radials	21	12,57	0,064	0,134	0,155	0,169	0,198

Taula 5.19.: Qualitat dels mètodes d'estimació.

Tal com s'assenyala a la taula, el mètode de la distància inversa ha estat escollit com a òptim en el 32,93% dels casos. Els mètodes de càlcul basats en *kriging* han estat escollits en 73 (43,71%) ocasions, mentre que els basats en mètodes directes han estat escollit en 94 (56,29%) ocasions.

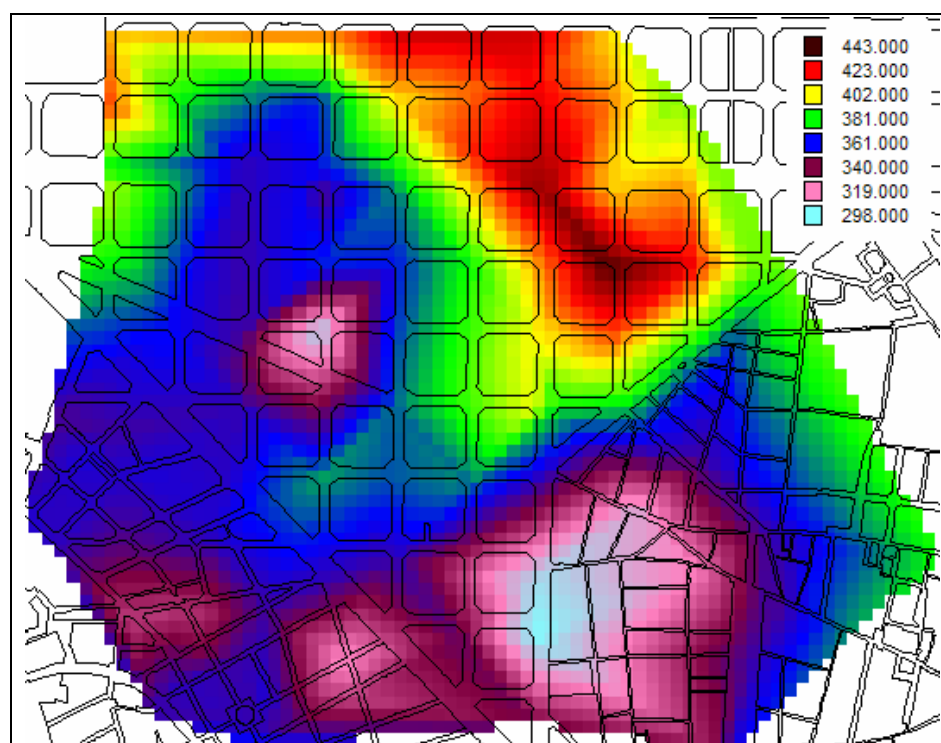


Figura 5.5.: Mapa de preus i espai interpolat mitjançant la tècnica dels triangles de Delaunay. Primera iteració.

A continuació, es presenta el gràfic d'influències de la mostra localitzada al carrer Compte Borrell.

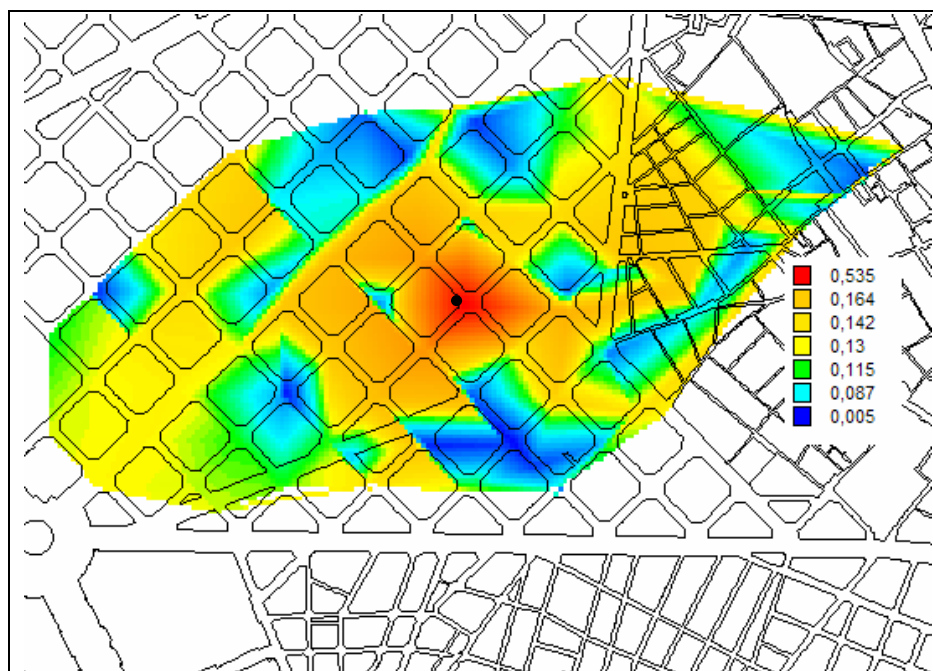


Figura 5.6.: Mapa d'influències de les mostres.
Mostra de Comte Borrell.

5.5.2.3. Resultats del procés iteratiu

Durant l'execució de l'algoritme de valoració iteratiu poden arribar a intervenir fins a 11 mostres. El procés no realitzarà les 5 iteracions en les situacions següents:

- No existeixen més mostres comparables homogeneïtzables.
- Dins del conjunt de mostres homogeneïtzables a utilitzar, cap no ha de presentar un percentatge d'homogeneïtzació inferior al 12%.

A més, s'aplicaran les restriccions inicials per determinar si la solució es pot considerar com a vàlida.

Mètode	%		Diferència entre validacions creuades				
			5%	25%	50%	75%	95%
Anisotròpic	81	27,00	0,030	0,045	0,104	0,133	0,179
Isotròpic	68	22,67	0,023	0,026	0,059	0,107	0,164
Mitjanes mòbils	22	7,33	0,019	0,029	0,075	0,117	0,166
Distància inversa	112	37,33	0,017	0,032	0,049	0,063	0,090
Radials	17	5,67	0,019	0,040	0,070	0,148	0,158

Taula 5.20.: Diferències entre els indicadors de qualitat a una volta i a cinc voltes sense aplicació de l'algoritme.

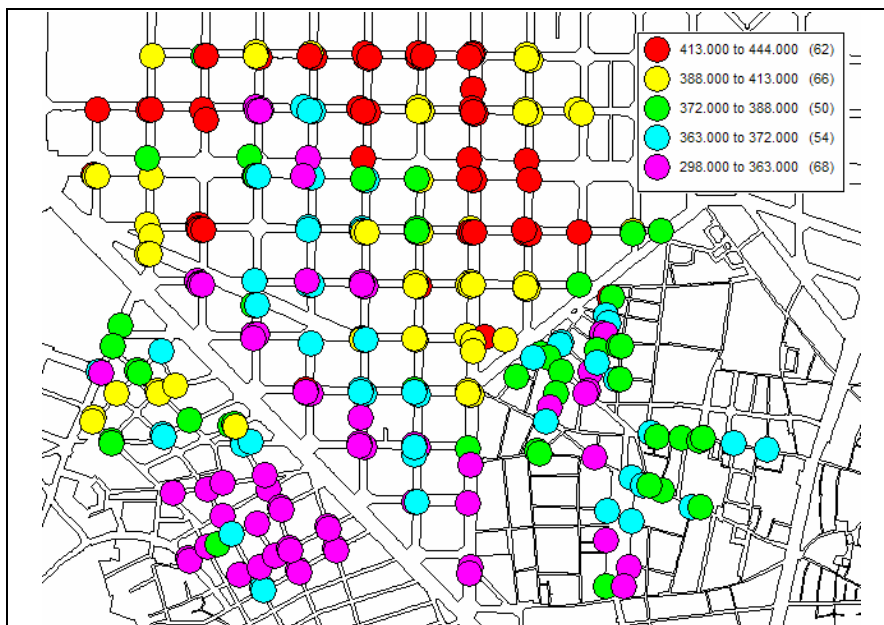


Figura 5.7.: Preus estimats després de cinc iteracions.

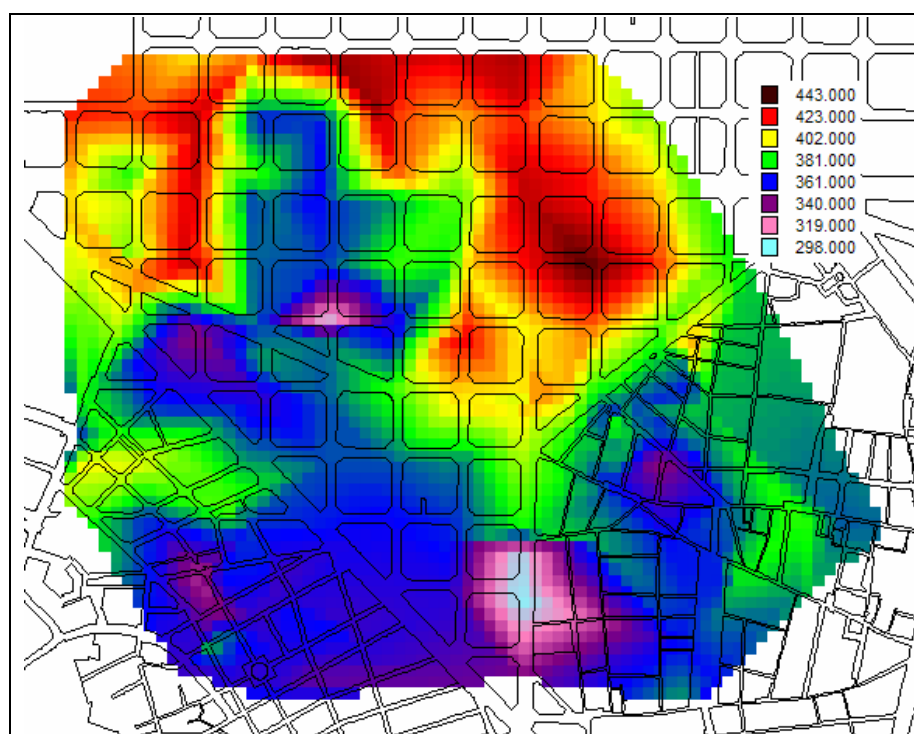


Figura 5.8.: Mapa de preus i espai interpolat mitjançant la tècnica dels triangles de Delaunay. Cinquena iteració.

Amb el procés iteratiu es va aconseguir augmentar el nombre de valoracions realitzades a partir de les mostres utilitzades de 167 (55,66%) a 300 (92,52%).

Els resultats per carrers van ser els següents:

Via	N	Mitjana	Desv. típica	Mínim	Màxim	Interval
Radas	2	345,323	8,273	339,473	351,173	17 - 53
Ronda de Sant Pau	7	352,794	43,834	298,435	401,355	12 - 80
Margarit	6	354,191	8,620	344,225	363,570	14 - 58
Concòrdia	4	354,682	14,955	332,255	362,630	18 - 62
Blasco de Garay	7	355,419	3,122	351,181	360,743	10 - 50
Creu Molers	8	358,749	9,317	351,162	377,594	14 - 85
Parlament de Catalunya	4	361,849	2,303	359,969	364,932	17 - 39
Riera Alta	5	363,351	22,813	337,780	384,851	28 - 69
Rambla del Raval	6	365,202	5,943	358,642	374,580	4 - 79
Jaume Fabra	3	365,567	382	365,292	366,003	12 - 16
Sant Gil	2	367,252	372	366,989	367,516	14 - 19
Bisbe Laguarda	3	367,459	6,406	361,727	374,374	1 - 10
Lleida	4	370,046	14,255	350,947	383,254	7 - 32

Taula 5.21.: Valors obtinguts (1/2).

Via	N	Mitjana	Desv. típica	Mínim	Màxim	Interval
Príncep de Viana	3	370,328	7,578	365,952	379,079	16 - 28
Hospital	7	371,018	8,244	361,869	380,369	56 - 126
Lluna	4	372,364	5,748	363,808	375,798	11 - 31
Botella	5	374,186	21	374,163	374,201	30 - 57
Carme	7	375,590	7,257	368,272	384,987	22 - 80
Calàbria	11	375,916	29,404	352,016	434,325	17 - 137
Floridablanca	13	376,299	34,016	305,599	412,416	11 - 145
Rocafort	10	376,816	25,524	352,447	431,898	19 - 125
Ricart	5	376,823	10,847	364,667	385,084	22 - 52
Mansó	8	377,779	18,767	351,337	413,117	16 - 78
Sant Vicenç	10	379,167	26,456	361,920	428,453	11 - 59
Viladomat	16	386,032	26,197	357,560	428,227	18 - 155
Teodor Bonaplata	3	387,873	5,560	384,663	394,294	5 - 13
Font Honrada	6	388,855	11,957	368,957	403,140	11 - 50
Tamarit	7	389,382	26,079	352,447	414,107	109 - 193
Gran Via de les Corts Catalanes	11	391,931	30,211	353,779	433,373	390 - 523
Compte Borrell	16	394,193	28,194	344,049	437,705	14 - 159
Sepúlveda	16	395,482	25,895	363,029	440,877	13 - 189
Diputació	14	397,811	25,705	361,940	433,417	45 - 178
Vallhonrat	3	398,447	4,449	393,370	401,666	20 - 30
Entença	8	400,301	40,561	338,233	431,581	17 - 88
Llançà	3	404,268	15,120	395,374	421,726	13 - 32
Vilamarí	7	410,822	7,435	398,088	419,697	8 - 78
Villaroel	10	417,380	14,856	404,427	443,308	16 - 94
Comte d'Urgell	12	417,939	13,209	395,851	432,212	10 - 93
Consell de Cent	24	421,781	10,461	406,548	434,325	102 - 209

Taula 5.21.: Valors obtinguts (2/2).

5.6. Exemples d'informes i processos. L'obtenció de la funció de probabilitat empírica del valor objecte

5.6.1. Introducció

L'exemple que es presenta a continuació és la representació del procés que habitualment realitzen els tècnics validadors de CATSA. Aquest procés ha de permetre al tècnic decidir si la selecció, l'homogeneïtzació i el valor proposat pel sistema expert són correctes o si, per contra, hi cal efectuar modificacions o incorporacions de noves mostres, tot elaborant una selecció diferent a la proposada o executant una adaptació manual del procés.

5.6.2. Exemple de valoració

L'objecte a valorar presenta les següents característiques:

Població	Hospitalet de Llobregat (L')	Codi Postal	08905
Tipologia	Habitatge Individual en edifici plurifamiliar		
Adreça	Carrer de Sant Francesc Xavier, XX segon pis		
Habitacions	3	Banys	1
Construïts	55,31	Útils	49
Ascensor	no	Reformes	no
Estat	Algunes reformes necessàries	Qualitat	Mitjana

Taula 5.22.: Taxació a valorar.

Les mostres proposades per la base i que van ser emprades per iniciar el procés de càlcul s'assenyalen a les pàgines següents:

	Hab.	Bany	m ²	Any	Adreça	Estat	Qualitat	Asc.	Planta	Preu	Vh	Dist.
Mostra 1	3	1	59	1974	Sant Francesc Xavier	Algunes reformes necessàries	Mitjana	No	3	228271	233046	24
Mostra 2	3	1	65	1960	Sant Francesc Xavier	Algunes reformes necessàries	Mitjana	No	1	254410	249379	41
Mostra 3	2	1	5..15	1965	Sant Francesc Xavier	Algunes reformes necessàries	Mitjana	No	3	210445	220222	48
Mostra 4	3	1	59	1974	Pins	Algunes reformes necessàries	Mitjana	No	2	231575	236514	94
Mostra 5	3	1	69,41	1960	Pins	Algunes reformes necessàries	Mitjana	No	3	249321	261293	97
Mostra 6	1	1	57,49	1960	La Florida	Algunes reformes necessàries	Mitjana	No	1	223155	236697	100

Taula 5.23.: Selecció inicial sense transformació de l'espai.

	Hab.	Bany	m ²	Any	Adreça	Estat	Qualitat	Asc.	Planta	Preu	Vh	Dist.
Mostra 1	2	1	56,38	1966	Illes Canàries	Algunes reformes necessàries	Mitjana	No	3	217289	232028	121
Mostra 2	3	1	59	1974	Sant Francesc Xavier	Algunes reformes necessàries	Mitjana	No	3	228271	233046	24
Mostra 3	3	1	69,41	1960	Pins	Algunes reformes necessàries	Mitjana	No	2	278126	261293	97
Mostra 4	3	1	77	1960	Abedul	Algunes reformes necessàries	Mitjana	No	3	293678	258912	211
Mostra 5	3	1	71,05	1970	Abedul	Algunes reformes necessàries	Mitjana - alta	Sí	2	274608	239768	217
Mostra 6	3	1	59	1974	Pins	Algunes reformes necessàries	Mitjana	No	2	231575	236515	94

Taula 5.24.: Selecció inicial amb transformació de l'espai en funció del varíograma anisotrópic.

Mètode	Valor de Taxació	EVC4	Pesos						Errors valoracions creuades					
			M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5	M6
Radials	237267	0,058	0,333	0,223	0,189	0,089	0,085	0,081	-0,023	0,079	-0,103	-0,061	0,097	0,003
Mitjanes mòbils	237318	0,062	0,351	0,209	0,176	0,091	0,087	0,085	-0,261	0,089	-0,112	-0,067	0,097	-0,001
Distància inversa	235726	0,063	0,558	0,198	0,139	0,037	0,035	0,033	-0,031	0,106	-0,124	-0,098	0,095	0,010
Isotrópic	239446	0,062	0,173	0,173	0,180	0,127	0,147	0,200	-0,042	0,077	-0,106	-0,041	0,098	-0,014
Anisotrópic	243221	0,046	0,186	0,153	0,160	0,160	0,167	0,175	-0,043	0,001	-0,109	-0,016	0,072	-0,034

Taula 5.25.: Resultats inicials pels diferents mètodes.

El model amb millor ajust va ser l'anisotrópic. Entre els models sense transformacions el de menor error va ser el mètode de les funcions radials. Després de les cinc iteracions, els resum de resultats va ser el següent:

Mètode	VT ₁	EVC4 ₁	VT ₂	EVC4 ₂	VT ₃	EVC4 ₃	VT ₄	EVC4 ₄	VT ₅	EVC4 ₅
Radials	237267	0,058	242401	0,052	244691	0,049	241361	0,040	244973	0,030
Mitjanes mòbils	237318	0,062	241412	0,058	241933	0,062	239646	0,053	245844	0,032
Distància inversa	235726	0,063	238363	0,067	238432	0,070	237450	0,065	247401	0,040
Isotrópic	239446	0,062	239446	0,049	239446	0,043	239446	0,030	239446	0,019
Anisotrópic	243221	0,046	244789	0,040	244788	0,036	244789	0,026	244789	0,016

Taula 5.26.: Resultats finals pels diferents mètodes.

	Hab.	Bany	m ²	Any	Adreça	Estat	Qualitat	Asc.	Planta	Preu	Vh _{cont}	%Vh _{cont}	Dist.	%Vh _{app}
Mostra 1	3	1	62	1960	Muses	AR	Mitjana	No	2	243040	245720	1,1%	633	-0,4%
Mostra 2	3	1	61,57	1960	Renclusa	AR	Mitjana	No	1	241108	244890	1,6%	631	0,0%
Mostra 3	2	1	62,14	1965	Font	AR	Mitjana	No	1	242035	242579	0,2%	601	0,9%
Mostra 4	3	1	65	1960	Sant Fr. Xavier	AR	Mitjana	No	2	254410	249379	-2,0%	41	-1,8%
Mostra 5	3	1	62,06	1965	Teide	AR	Mitjana	No	3	248488	249420	0,4%	183	-1,9%
Mostra 6	3	1	59	1974	Pins	AR	Mitjana	No	2	231575	236514	2,1%	94	3,5%

Taula 5.27.: Selecció inicial amb transformació de l'espai en funció del variograma anisotrópic.

AR = Algunes reformes necessàries

El resultat del procés de càlcul va donar com a solució final 244.789 €

El resultat del procés *bootstrap* a partir de les 15 dades més properes i després de la transformació en base al mètode *kriging* anisotròpic va donar els següents estadístics.

N	5005	Mitjana	245834
Desviació típica	11810	Mediana	244508
Rang	96192	Màxim	313226
Rang interquartílic	11321	Mínim	217033

Taula 5.28.: Resultat *bootstrap* sobre 15 mostres.

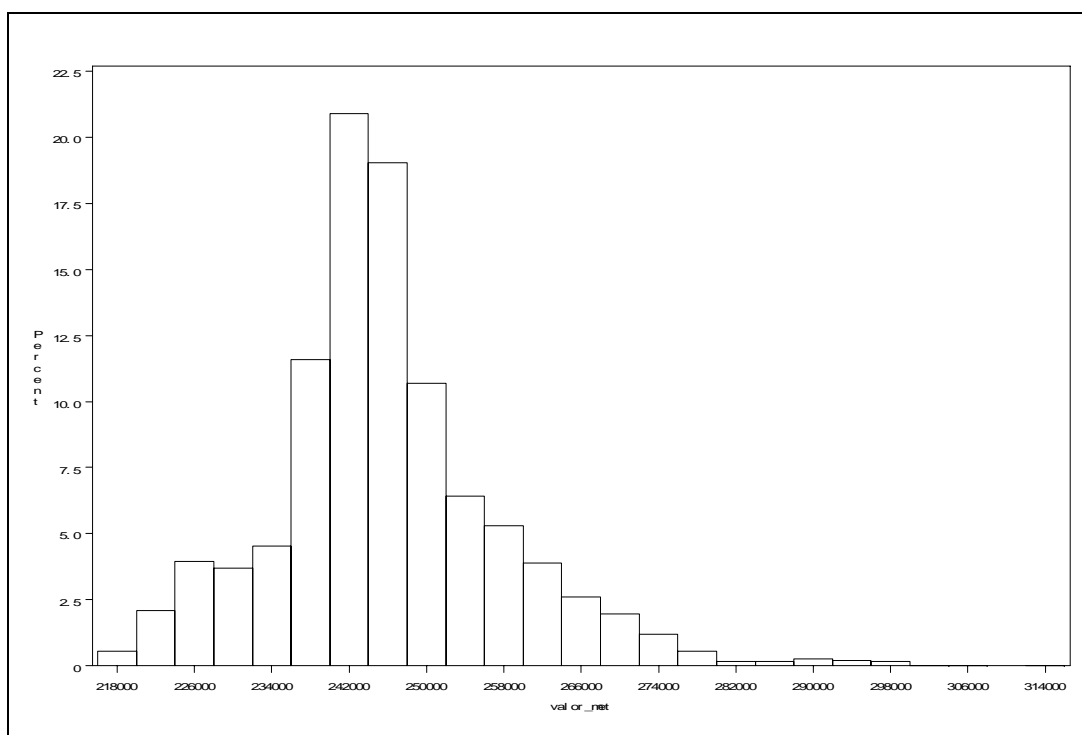


Figura 5.9.: Gràfic *bootstrap*: resultat de les valoracions realitzades a partir del remostreig de 15 mostres properes utilitzades de 6 en 6.

Es va comprovar que en aquest cas les estimacions obtingudes a partir del procés anisotròpic escollit queden compreses dins de l'interval de confiança proposat pel procés *bootstrap* anisotròpic realitzat a partir de

les quinze mostres més properes després de ser transformades espacialment.

L'informe tècnic obtingut a partir del mètode iteratiu plantejat va ser el següent:

Població	Hospitalet de Llobregat (L')		Codi Postal	08905							
Tipologia	Habitatge Individual en edifici plurifamiliar										
Adreça	Carrer de Sant Francesc Xavier, XX segon pis										
Habitacions	3		Banys	1							
Construïts	55,31		Útils	49,00							
Ascensor	no		Reformes	no							
Estat	Algunes reformes necessàries		Qualitat	Mitjana							
Valor determinat pel mètode de mercat		244789 €	Desviació típica		11810						
Valor mitjà determinat en base a la zona		245834 €	Interval de confiança 90%								
Valor medià determinat en base a la zona		244508 €	225478	271401							
Mostres utilitzades											
N. Hab.	N. Banys	m ²	Any	Adreça	Est at	Qualitat	A s c	P i s	Preu	%Vh	P e s
3	1	62	1960	Muses	AR	Mitjana	No	2	243040	0,7%	1/6
3	1	61,57	1960	Renclusa	AR	Mitjana	No	1	241108	1,5%	1/6
2	1	62,14	1965	Font	AR	Mitjana	No	1	242035	1,1%	1/6
3	1	65	1960	Sant Fr. Xavier	AR	Mitjana	No	2	254410	-3,8%	1/6
3	1	62,06	1965	Teide	AR	Mitjana	No	3	248488	-1,5%	1/6
3	1	59	1974	Pins	AR	Mitjana	No	2	231575	5,7%	1/6

Taula 5.29.: Resultats del procés de valoració.

Cal destacar que l'interval de confiança obtingut no era simètric, ja que va ser establert a partir de la funció de densitat empírica obtinguda a partir del procés *bootstrap*.

5.6.3. Exemple d'informe

El procés presentat es troba en aquests moments en fase de programació, depuració d'error i establiment de regles. A les figures 5.10. – 5.13. es mostra una sèrie d'exemples d'informes interns realitzats a partir de diferents softwares de visualització.

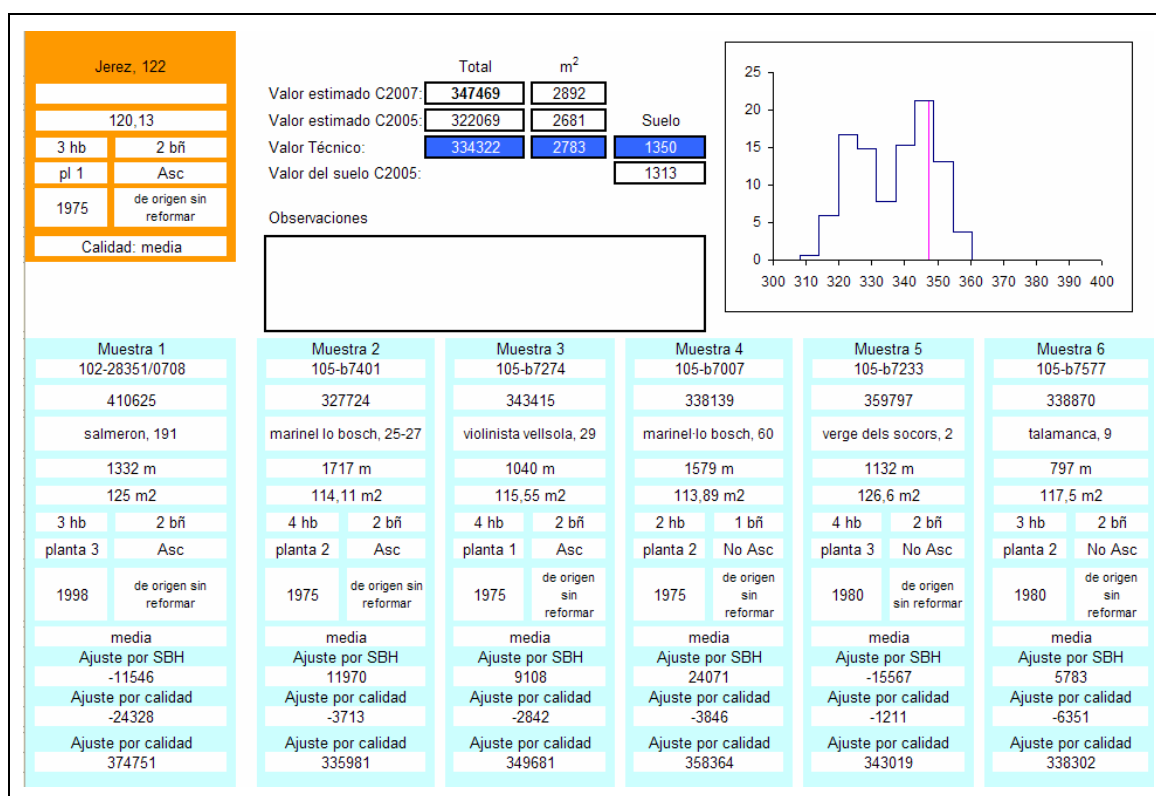


Figura 5.10: Exemple d'informe de comprovació en fase de proves

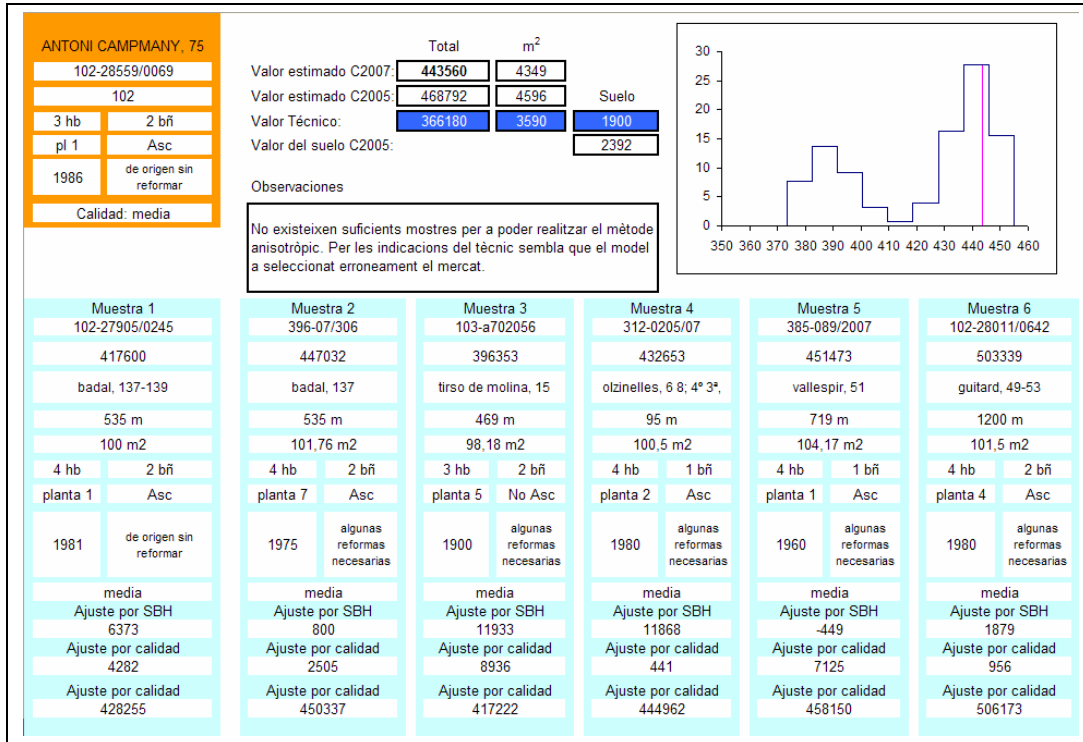


Figura 5.11: Exemple d'informe de comprovació en fase de proves

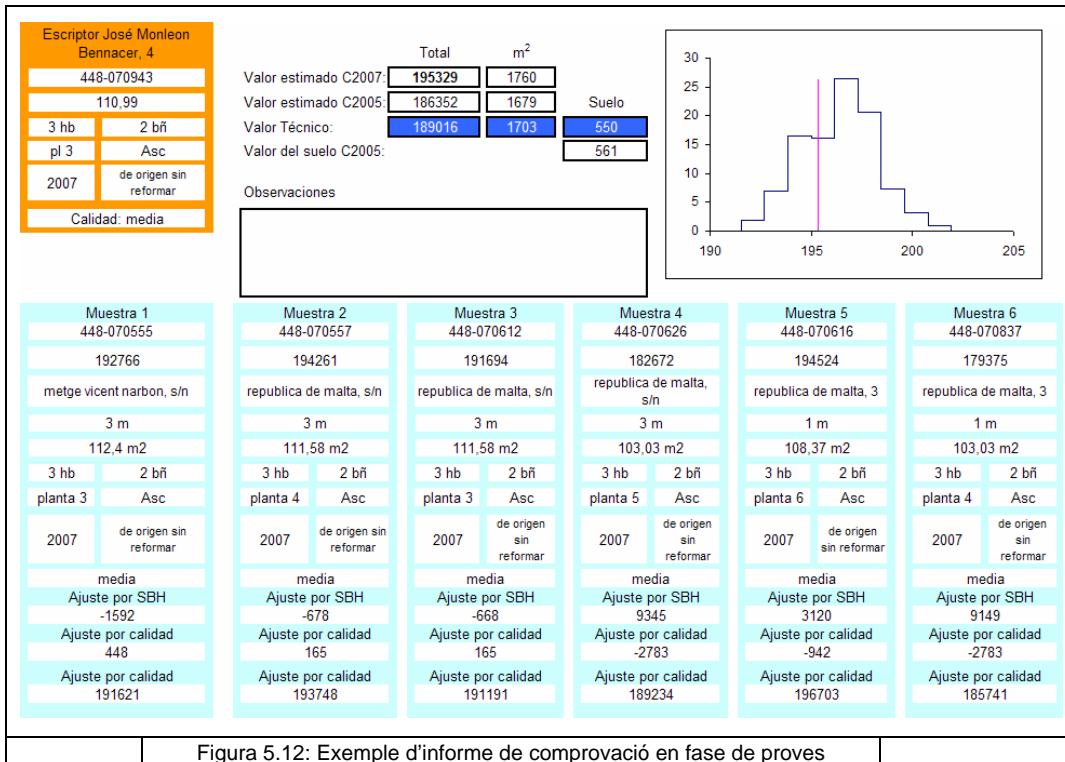


Figura 5.12: Exemple d'informe de comprovació en fase de proves

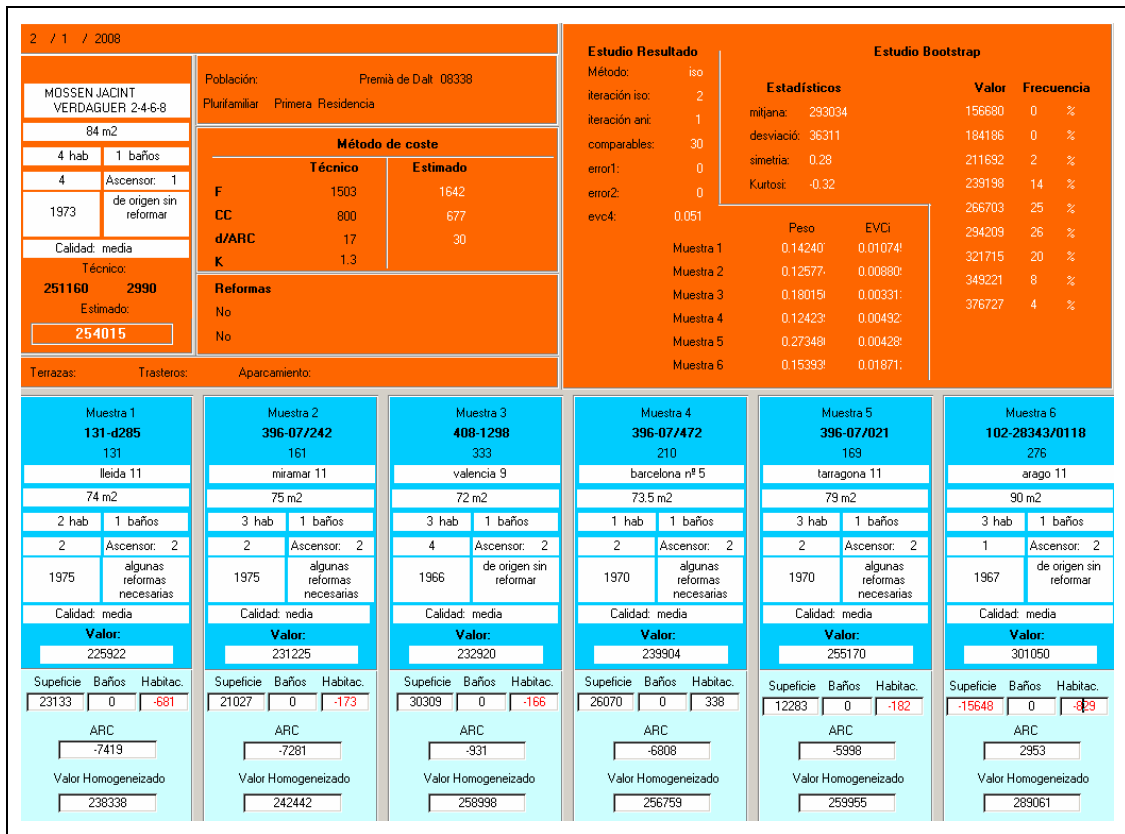


Figura 5.12: Exemple d'informe de comprovació en fase de proves avançades

6. Conclusions i treballs futurs

El primer objectiu de la tesi era l'obtenció d'una metodologia que permetés calcular el valor de taxació pel mètode de comparació de mercat. Per assolir aquest objectiu es va plantejar un procediment per parts. Les parts que defineixen aquest procediment van ser les següents:

- L'aplicació dels procediments tècnics d'homogeneïtzació plantejats pel departament tècnic de CATSA amb l'objectiu d'aconseguir un conjunt de mostres idèntiques a l'objecte a valorar.
- L'elaboració d'una llista de mostres a utilitzar en base a la distància, les característiques constructives, l'estat i la qualitat.
- L'execució del procés iteratiu basat en els mètodes d'estimació i interpolació espacial.
- L'obtenció de la distribució empírica a partir dels mètodes de remostreig *bootstrap*.

Tot el procediment de valoració automatitzat es va plantejar amb l'objectiu de complir tant amb la llei com amb la manera de procedir dels tècnics. Això va suposar haver de crear un procés realista envers l'àmbit de treball i acurat

envers la llei tant amb les necessitats d'informació plantejades com en els resultats i informacions aportats per la solució.

Tot i els problemes de qualitat en la informació disponible, es pot concloure que l'algoritme iteratiu creat per a la realització del procés de valoració funciona amb una precisió suficientment elevada com per ser utilitzat com eina de suport a la valoració.

L'aportació principal de l'algoritme iteratiu radica en l'increment del percentatge de valoracions realitzades sense produir-se pèrdua de precisió en el procés.

Per qüestions de calendari, restarà per realitzar una nova revisió de la qualitat de l'estimació quan el nou aplicatiu de CATSA es trobi en funcionament i no calgui realitzar cap procés de recodificació de la informació. És de suposar que quan la qualitat de les bases sigui l'òptima les estimacions seran més precises i de més qualitat.

El segon objectiu era establir un mètode de selecció entre els diferents estimadors del valor de taxació que permetés minimitzar l'error d'estimació i fos alhora un indicador de la qualitat de l'estimació. Es va aconseguir un sistema de selecció del mètode d'estimació de gran qualitat, mentre que no va ser possible definir un indicador de la qualitat de l'estimador.

En el context de futurs treballs s'ha resolt estudiar les possibilitats dels mètodes *bootstrap* com a indicador de la qualitat de la valoració.

Conceptualment es podria plantejar el mètode de selecció de mostres de la següent manera: la primera solució del procés iteratiu es basa únicament en la distància dins de l'espai de comparables; i que el criteri de 'proximitat' serà substituït a mesura que passen les iteracions pel criteri de 'similitud' donant

més importància a la semblança entre mostra i objecte que a la seva proximitat espacial. Consegüentment, la solució final quedarà establerta mitjançant una combinació de distància i semblança dins de l'espai de comparables.

Una de les aportacions importants de l'algoritme iteratiu ha estat facilitar la detecció i identificació de les mostres amb comportaments no habituals dins del mercat, fet que permet substituir-les per d'altres de noves.

Es pot concloure que el mètode de selecció de mostres que proporciona l'algoritme iteratiu permet valorar correctament un major nombre d'objectes i augmentar el percentatge d'estimacions òptimes tot mantenint la qualitat de la resta de mostres. Tanmateix, es va constatar que en la majoria dels casos la solució adoptada en finalitzar el procediment iteratiu era de major precisió que la solució inicial.

Dels sis mètodes d'estimació plantejats inicialment per treballar amb distàncies es va haver de desestimar el mètode basat en funcions suavitzades bicúbiques (*spline*), ja que aquest mètode no proporciona les ponderacions de cada una de les mostres en la construcció del valor de l'objecte. Sota el requeriment de màxima claredat, aquesta situació no resultava òptima i es va optar per eliminar aquest mètode d'interpolació per passar a treballar únicament amb els mètodes d'estimació per mitjanes mòbils, distància inversa, funcions radials i els *kriging* anisotròpic i isotròpic.

Dels cinc mètodes restants, els mètodes *kriging* són els únics que proporcionen estadístics, ja que la resta són mètodes d'interpolació que no aporten estadístics per a la realització d'estimacions o la creació d'interval de confiança. Amb la introducció de les metodologies *bootstrap* es va superar aquest inconvenient tot aconseguint de manera empírica la funció de distribució del valor en base a l'ECO805/2003.

L'obtenció d'una funció de distribució del valor d'estimació era el tercer objectiu d'aquesta tesi.

Actualment, s'està estudiant la possibilitat d'identificar l'existència de més d'un mercat en una zona a partir de les formes de la funció empírica aportades per la metodologia *bootstrap*. El plantejament futur és estudiar la incorporació de funcions de probabilitat d'aparició de les mostres en funció de la distància dins el procediment *bootstrap*.

Pel que fa als mètodes d'estimació i d'interpolació emprats, cal fer esment de l'elevada precisió proporcionada pels mètodes *kriging* en situacions complexes. Els resultats aportats per la metodologia *kriging* són més que un simple estimador en funció de la distància, ja que incorporen l'estudi de la dependència espacial a gran escala, via l'estudi de la tendència global de la mostra i la interdependència entre mostres a partir de l'anàlisi estructural. El desavantatge dels estimadors basats en la metodologia *kriging* és que per poder garantir-ne una certa qualitat, i que la informació proporcionada pel variograma aportí un valor afegit a la qualitat de l'estimador s'han de basar en un tamany mostral que no es requereix en l'àmbit de les valoracions amb caràcter hipotecari. Això suposa que aquesta metodologia només pugui ser aplicada ocasionalment, donat que per a moltes localitzacions els mercats de comparables no són suficients.

El variograma ha estat utilitzat per a la realització dels mètodes *kriging* i també com a mecanisme per a definir l'espai de comparables en base al rang, indicador dels límits de la dependència espacial.

En l'apartat 5.4. es va presentar un exemple que va permetre crear un mapa de valor de mercat.

La variació d'una de les característiques permet crear diferents mapes de valors de mercat i presentar-los com un sistema de capes que possibilita estudiar tant el comportament general com el comportament en una localització determinada per al seu conjunt de possibles característiques.

Durant el procés de creació de mapes van sorgir diverses preguntes que van quedar sense resposta.

Un projecte futur té com a objectiu desenvolupar tècniques que permetin definir polígons espacials no estimables a partir del conjunt mostral existent, per a la qual cosa caldrà estudiar:

- Si és possible l'extrapolació i quines pèrdues de qualitat i precisió pot comportar.
- Una definició més acurada de l'espai estimable dins d'un polígon mostral.
- Les possibilitats del rang; determinar com pot ser utilitzada la informació que proporcionen els variogrames direccionals.

Es va evidenciar que la interpolació de valors en espais on el mercat és desconegut o inexistent pot induir a mapes de valor erronis.

A més, es va fer palès que quan el mapa de mostres era incapaç de produir una imatge d'una determinada situació espacial, difícilment les estimacions que es poguessin realitzar d'aquell conjunt mostral podrien ser correctes.

En relació a la ubicació de les mostres es van presentar dos problemes.

El primer problema afectava a la possibilitat d'utilitzar mostres localitzades en la mateixa localització de l'objecte; el segon estava relacionat amb la possibilitat d'utilitzar mostres localitzades en el mateix punt.

En aquesta situació, el mètode dels radials és l'únic mètode capaç d'obtenir una estimació del valor més o menys encertada. Mentre que els mètodes de la distància inversa i de les mitjanes mòbils no aporten cap solució, el mètode *kriging* -tal com ha estat plantejat en aquesta tesi- dona un pes proper a 1 a la mostra a distància 0; una situació no molt adient pel que fa a l'ECO805/2003, doncs s'estaria estimant el valor de l'objecte en funció d'una única mostra.

Com a plantejament de futur caldrà estudiar la manera d'introduir adientment les mostres localitzades en la mateixa ubicació que l'objecte a taxar. És possible que la metodologia *kriging* arribi a estimar aquest comportament com l'efecte llavor (*nugget effect*) del procés, però perquè això es produeixi ha d'existir un conjunt de mostres prou ampli per tal de poder realitzar l'estimació d'aquest efecte.

Tanmateix, caldrà definir la manera d'utilitzar mostres situades en una mateixa localització per tal de poder emprar la metodologia *kriging* en casos en què ho permeti el nombre de mostres disponibles, ja que els procediments matricials plantejats en la metodologia *kriging* no admeten aquesta situació.

Els processos de prova sobre els models presentats van començar a funcionar el dia 15 d'octubre del 2007.

La forma d'estudiar la qualitat d'aquests processos, identificar possibles errors, situacions de desviació, etc. serà a partir de l'avaluació per part dels tècnics validadors dels resultats d'aquests processos per tal d'incorporar-hi en cas necessari les regles adients per a millorar el model expert.

S'espera que l'aplicatiu de CATSA es trobi funcionant de forma òptima al gener del 2009.

7. Bibliografía

Akaike, H. (1974). *A new look at the statistical model identification*. IEEE Transactions on Automatic Control 19 (6): 716–723

Anselin L. (1988): *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.

Anselin, L. (1990): *Some Robust Approaches to Testing and Estimation in Spatial Econometrics*. Regional Science and Urban Economics, 20, 141-163.

Aznar, J. i Guijarro, F. (2006). *Nuevos métodos de valoración*. Biblioteca UPV.

Ballesteros, E. i Rodríguez, J. A. (1999). *El precio de los inmuebles Urbanos*. CIE Inversiones Editoriales Dossat 2000.

Basu, S. i Thibodeau, T.G. (1998). *Analysis of Spatial Autocorrelation in House Prices*. The Journal of Real Estate Finance and Economics, Vol.17, Num. 1, pp. 61-85.

Bennett, R.J. i Haining R.P. (1985): *Spatial structure and spatial interaction: Modeling Approaches to the statistical analysis of geographical data*. Journal of the royal statistical society. Series A, vol. 148. No 1, 1-36.

Bivand, R. (2001): *More on spatial data*. R news, 1(3):13-17, september 2001.

Brown, J. N. I Rosen, H. S. (1982). *On the Estimation of Structural Hedonic Price Models*, *Econometrica* 50(3):765-778.

Bruce, R.W., Sundell, D.J. (1977): *Multiple Regression Analysis History and Applications in the Appraisal profession*, *The Real State Appraiser*, 43, 1.

Can, A. (1992): *Specification and Estimation of Hedonic Housing Price Models*. *Regional Science and Urban Economics*, 22, 453-474.

Can, A., Megbolugbe, I. (1997): *Spatial Dependence and House Price Index Construction* *Journal of Real Estate Finance and Economics*, 14: 203–222

Cano Guervós, R. (1999): *Aproximación al valor de la vivienda. Aplicación a la Ciudad de Granada*. Biblioteca de Económicas y Empresariales. Editorial Universidad de Granada.

Cassel, D.K., O. Wendroth i D.R. Nielsen. (2000). *Assessing spatial variability in an Agricultural Experiment Station Field: Opportunities arising from spatial dependence*. *Agronomy Journal* 92: 706-714.

Ceular N. i Caridad J.M. (2001): *Un análisis del mercado de la vivienda a través de redes neuronales artificiales*. *Estudios de economía aplicada*. 18, 2, 41-66.

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

Chica Olmo, J.M. (1994): *Teoría de las variables regionalizadas : aplicación en economía espacial y valoración inmobiliaria*. Editorial Universidad de Granada

Chica Olmo, J.M, Cano Guervós, R. (2003): *Aproximación a la variabilidad espacial de las características y del precio de la vivienda*. *Anales de economía aplicada*, 2003, ISBN 84-607-7655-7

Chica Olmo, J.M, Cano Guervós, R. (2004): *Una metodología Geo-Econométrica para la valoración inmobiliaria*. *Ciudad y territorio: Estudios territoriales*, ISSN 1133-4762, N° 139, 2004, pags. 135-154

Clapp, J.M.(2004): *A Semiparametric Method for Estimating Local House Price Indices*, Real Estate Economics Volume 32, Issue 1, Page 127-160, Mar 2004

Cressie, N. (1991): *Statistics for Spatial Data*. John Wiley and Sons, New York.

Christensen, O. F., Ribeiro, P.J. (2002): *Georglm: A package for generalised linear spatial models*. R News, 2(2):26-28, June 2002.

Dubin, R.A. i Hsing, C. (1990) *Specification of hedonic regressions: Non-tested test on measures of neighbourhood quality*. Journal of Urban Economics 27:97-110.

Dubin, R.A. (1992): *Spatial Autocorrelation and Neighborhood Quality*. Regional Science and Urban Economics, 22, 433-452.

Dubin, R. A. (1998): *Predicting House Prices Using Multiple Listings Data*. Journal of Real Estate Finance and Economics , 17(1), 35-59.

Duda, R. O., Hart, P. E. (1973): *Pattern Classification and Scene Analysis*. Wiley-Interscience, New York.

Efron, B., Tibshirani, R.J. (1993): *An introduction to the Bootstrap*. Chapman & Hall. New York

Fernández Pirla, S. (1992): *Valoración de bienes inmuebles*. Ed. Consejo Sup. Arquitectos de España.

Fik, T. J., Ling, D.C., Mulligan, G.F. (2003): *Modelling Spatial Variation in Housing Prices: A Variable Interaction Approach*. Real Estate Economics Volume 31, Issue 4, Page 623-646, Dec 2003

Fox, J. (1997); *Applied regression analysis, linear models, and related methods*. Sage Publications, Inc

Freeman, J. A., Skapura, D. M. (1991). *Neural Networks Algorithms, Applications and Programming Techniques*, pages 89-128. Addison-Wesley, first edition, 1991.

Gallego Mora-Esperanza, J. (2004): *La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid*. Catastro, 2004

Gámez, M. i Montero, J.M. (1999): *A spatial methodology for the estimation of housing price in albacete*.

Garcia A. i DiPasqua, H. (2007): *Documents de treball de CATSA*.

Garcia-Almirall, P. (1998): *La valoració urbana en base a les noves tècniques de SIG. L'exemple de L'Hospitalet de Llobregat*. Tesi doctoral

Garcia-Almirall, P. et Al. (2004): *Projecte CATSA: La elaboración de un modelo de valoración de viviendas en el ámbito de Cataluña*. Garcia, A. i

Gringarten, E. i Deutsch, C.V (2001): *Teacher's Aide Variogram Interpretation and modeling*. Mathematical Geology, Vol. 33, No. 4.

Guimet, J. i Urrutia, G. (2000): *Expresión del valor catastral mediante ecuaciones de regresión: su aplicación a la difusión pública de un valor de referencia*.

Johnston, J. (1987): *Métodos de econometría*, Vicens Vives

Kain, J.F i Quigley, M.J., (1970): *Measuring the Value of Housing Quality*. Journal Of the American Statistical Association 65, pp. 532-548.

Krige, D. G. (1951). *A statistical approach to some basic mine valuation problems on the Witwatersrand*. J. Of the Chem., Metal. And Mining Soc. Of South Africa 52 (6): 119-139.

Martín A. (2002). *Geoestadística aplicada*. Instituto de Geofísica, UNAM, Instituto de Geofísica y Astronomía, CITMA, Cuba.

MacQueen, J. B. (1967): *Some Methods for classification and Analysis of Multivariate Observations*. Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281-297

Matheron, G. (1962). *Traité de géostatistique appliquée*. Editions Technip..

Matheron, G. (1965): *Les Variables regionalisées et leur estimation*. Masson, Paris.

Matheron G. (1970): *La Théorie des variables regionasées et ses applications*. Centre de Géostatistique et de Morphologie Mathématique, Fas.1. Ecole Nationale Supérieure de Mines de Paris.

Medsker, L.R., and Jain, L.C. (2000): *Recurrent neural networks: design and applications*, Boca Raton, FL: CRC Press,

Meese, R., and Wallace, N. (1991): *Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices*. Journal of the American Real Estate and Urban Economics Association, 19, 308-322.

Meese, R., and Wallace, N. (1997): *The Construction of Residential Housing Price Indices: A comparison of Repeat-Sales, Hedonic-Regression and Hybrid Approaches*. Journal of Real Estate Finance and Economics, 14, 51-73.

Montero Lorenzo, J.M, (2004): *El precio medio del metro cuadrado de la vivienda libre: Una aproximación metodológica desde la perspectiva de la Geoestadística*. Revista Estudios de Economía Aplicada.

Montero J.M., Larraz, B. (2006): *Estimación espacial del precio de la vivienda mediante métodos de Krigeado*. Estadística española, ISSN 0014-1151, Vol. 48, Nº 162, 2006, pags. 201-240.

Fik, T., Ling, D. C. i mulligan, G. F. (2003): *Modeling Spatial Variation in Housing Prices: A Variable Interaction Approach*. Real Estate Economics Vol. 31, Num. 4, pp. 623-646.

Pace, R. K., Barry, R. i Sirmans, C. F. (1998): *Spatial Statistics and Real Estate*. Journal of Real Estate Finance and Economics, 17 (1), 5-13.

Rathbun, S. (1994): *Classnotes environmental statistics*. Univ. of Georgia, Athens, Juny 1994.

Ribeiro P.J., Jr. Diggle, P.J. (2001): *geoR: A package for geostatistical analysis*. R News, 1(2):15-18, June 2001.

Ripley, B.D. (2001): *Spatial Statistics in R*. R News, 1(2):14-15, June 2001.

Roca Cladera, J. (1982): *Tesi doctoral: Vers una interpretació espacial dels valors del sòl a Barcelona*.

Roca Cladera, J. (1986): *Manual de valoraciones inmobiliarias*. Ariel Economía, Barcelona.

Roca Cladrea, J (1988): *La estructura de los valores urbanos: un análisis teórico empírico*. Madrid. Instituto de Estudios de la administración local.

Rodon, O. (2006): *Curso de Krigeado*. CESMA group. Venezuela

Rosen, S. (1974): *Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition*. Journal of Political Economy, University of Chicago Press, vol. 82(1), pages 34-55.

Rossiter, D.G. (2006): *An introduction to applied geostatistics. Department of earth systems analysis*.

Stein A., Corsten, L. C. A. (1991): *Universal Kriging and Cokriging as a Regression Procedure*. Biometrics, Vol. 47, No. 2 (Jun., 1991), pp. 575-587

Santillana del Barrio, A. (1972): *Análisis económico del problema de la vivienda*. Ed. Ariel, Barcelona.

Schlather, M. (2001): *Simulation and analysis of random fields*. R News, 1(2):18-20, June 2001.

Stoet C. i Snepvangersm J. (2005): *Mapping Curvilinear Structures with Local Anisotropy Kriging*. Mathematical Geology, Vol. 37, No. 6, August 2005 (C_ 2005)

Wilkowski, W., Budzyński, T. (2006): *Application of Artificial Neural Networks for Real Estate Valuation*. Shaping the Change XXIII FIG Congress Munich, Germany, October 8-13.

Witte, A.D, Sumka, H.J, Erekson, H. (1979): *An estimate of a structural hedonic Price model of the housing market: An application of Rosen´s theory of implicit Markets*. Econometrica 47, pp.1151-1173.

Yang C. et al. (2005): *Twelve different interpolation methods: a case study of surfer 8.0*.

Zárate, L.E. i Neto, A.P. (2003): *Valuation of Urban Real Estate Through Artificial Neural Networks*. Artificial Intelligence and Applications.