

TESIS DOCTORAL

Sistemas de información
geográfica y algoritmos de
optimización aplicados a
problemas geoespaciales en la
Administración de Tierras

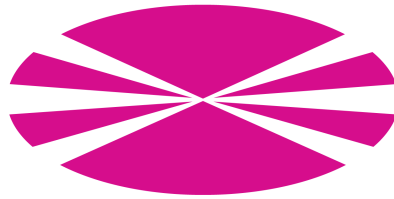
Juan Porta Trinidad

2013



Departamento de Electrónica y Sistemas
Universidad de A Coruña

Departamento de Electrónica y Sistemas
Universidad de A Coruña



TESIS DOCTORAL

**Sistemas de información geográfica
y algoritmos de optimización
aplicados a problemas
geoespaciales en la Administración
de Tierras**

Juan Porta Trinidad

Mayo 2013

Dirigida por:
Ramón Doallo Biempica

Dr. Ramón Doallo Biempica
Catedrático de Universidad
Dpto. de Electrónica y Sistemas
Universidade da Coruña

CERTIFICA

Que la memoria titulada *Sistemas de información geográfica y algoritmos de optimización aplicados a problemas geoespaciales en la Administración de Tierras* ha sido realizada por **Juan Porta Trinidad** bajo mi dirección en el Grupo de Arquitectura de Computadores del Departamento de Electrónica y Sistemas de la Universidade da Coruña, y constituye la Tesis que presenta para optar al grado de Doctor en Ingeniería Informática.

En A Coruña, mayo de 2013

Fdo.: Ramón Doallo Biempica
Director de la Tesis Doctoral

Fdo.: Juan Porta Trinidad
Doctorando

El tribunal abajo firmante certifica que esta es la versión aprobada de la Tesis de
D. Juan Porta Trinidad titulada:

**Sistemas de información geográfica y algoritmos de
optimización aplicados a problemas geoespaciales en la
Administración de Tierras**

Tribunal:

Presidente,

Miembro/a,

Miembro/a,

Miembro/a,

Secretario/a.

Resumen

La tierra es un factor estratégico de suma importancia en el desarrollo de una determinada región. La tierra ha de ser considerada como un bien capaz de generar crecimiento económico y de la que hay que intentar sacar el máximo provecho. Esta tesis aplica las nuevas tecnologías a procesos de la Administración de Tierras que hasta ahora o bien no se hacían, o bien se hacían de forma rudimentaria.

Uno de estos procesos es la ordenación territorial que, por su propia naturaleza, involucra gran cantidad de elementos con información geográfica (parcelas, edificaciones, entorno, ...) con los que se realizan operaciones espaciales computacionalmente muy costosas. Para estas tareas se han desarrollado algoritmos de búsqueda heurística que hacen uso de técnicas de computación de altas prestaciones para reducir tiempos de ejecución.

También se ha diseñado una arquitectura Web-SIG genérica pero fácilmente adaptable para la resolución de problemas concretos como pueden ser la dinamización de los mercados de tierras y la optimización de la gestión común de parcelas, dos aspectos importantes para el desarrollo rural.

Las soluciones desarrolladas hacen uso de estándares y herramientas libres para favorecer la compatibilidad entre sistemas, abaratar costes, promover el e-Government y seguir las directivas de la Unión Europea en temas de información geoespacial.

Resumo

A terra é un factor estratéxico de suma importancia no desenvolvemento dunha determinada rexión. A terra ten que ser considerada como un ben capaz de xerar crecemento económico e da que se ten que sacar o máximo proveito. Nesta tese aplícanse as novas tecnoloxías a procesos da Administración de Terras que hasta agora ou ben non se facía, ou ben facíanse de forma rudimentaria.

Un destes procesos é a ordenación territorial que, pola súa propia natureza, involucra gran cantidade de elementos con información xeográfica (parcelas, edificacións, entorno, ...) cos que realízanse operacións espaciais computacionalmente moi custosas. Para estas tarefas desenvóléronse algoritmos de busca heurística que fan uso de técnicas de computación de altas prestacións para reducir tempos de execución.

Tamén foi deseñada unha arquitectura Web-SIX xenérica pero doadamente adaptable para a resolución de problemas concretos coma a dinamización dos mercados de terras e a optimización da xestión común de parcelas, dous aspectos importantes para o desenvolvemento rural.

As solucións desenvolvidas fan uso de estándares e ferramentas libres para favorecer a compatibilidade entre sistemas, abaratar custos, promover o e-Government e seguir as directivas da Unión Europea en temas de información xeoespacial.

Summary

Land is a very important strategic factor for the rural development. Land must be considered as a good able to generate economic growth and thus try to take the most of it. This master thesis applies new technologies to Land Administration processes which until now had been done in a rudimentary way or even not done at all.

One of these processes is the spatial planning that, by its very nature, involves high number of elements with geographic information (plots, buildings, environment, ...) and carry out computationally expensive spatial operations. Heuristic search algorithms have been developed for these issues and they use high performance computing techniques for reducing execution times.

A generic architecture of an Web-GIS system has been also designed. It can be easily adapted to specific problems such as the invigoration of land markets and the optimization of common management of plots, two important issues for the rural development.

The developed solutions use standards and license-free tools for encourage system compatibilities, reduce costs, support e-Government and follow the European Union directives in geospatial information issues.

Agradecimientos

En primer lugar me gustaría mostrar mi agradecimiento a Ramón Doallo por la ayuda prestada durante la realización de esta tesis. Junto a él, quisiera añadir también a este agradecimiento a Jorge Parapar y a Juan Touriño por su labor y apoyo durante todos mis años trabajando en el Grupo de Arquitectura de Computadores (GAC) de la Universidad de A Coruña (UDC). De igual manera, agradecer a Rafael Crecente del Laboratorio del Territorio (LaboraTe) de la Universidad de Santiago de Compostela (USC), su aportación e implicación en todos los proyectos.

Quiero también dar las gracias a todos los/as compañeros/as y ex-compañeros/as del GAC y del Departamento de Electrónica y Sistemas por crear el clima necesario para el trabajo, la investigación, el esparcimiento y la amistad; y a la gente del LaboraTe por toda la ayuda prestada.

Agradecer también a todos los organismos que han financiado los proyectos que forman esta tesis: a la Universidad de A Coruña; a la Consellería do Medio Rural e do Mar y a la Xunta de Galicia por los proyectos 08SIN011291PR, Consolidación de Grupos de Investigación Competitivos 2010/06 y 2010/28, Red de Tecnologías de la Información Geoespacial para la Gestión Territorial (CN2012/323) y Grupos de Referencia Competitiva (CN2012/211); a la Agencia Gallega de Desarrollo Rural (AGADER); y al Fondo Europeo de Desarrollo Regional (FEDER). Especial agradecimiento para el Grupo de Arquitectura de Computadores de la USC y para el Centro de Investigación TIC (CITIC).

Aprovecho estas líneas para mostrar también mi cariño a todos los que me rodean y que durante estos años han sido una maravillosa vía de escape. Son muchos nombres los que quiero poner aquí, y eso ya es lo que realmente me alegra. No obstante quisiera nombrar a Jose (cuyo agradecimiento tiene carácter retroactivo), a

Gloria y a Ricardo (en representación de los amigos de A Coruña), a Javi y a Miguel (en representación de los amigos de A Guarda, Fornelos y alrededores), a Xurxo y a Paula (en representación de los amigos del trabajo), a María y a Aldana (en representación de los *otros* amigos del trabajo), a Ben y a Gomez (en representación de las bandas que me han acompañado estos años), a Alex, a Dani y a Victor (en representación de sus respectivas madres y padres), y a todos los que cuando sepan que he presentado esta tesis, sonrían. Y especial mención merecen Trini y Pepe a los que, en representación del resto de mi adorable familia, quiero agradecerles... todo.

Juan Porta Trinidad

Índice general

Índice general	xv
Índice de tablas	xxiii
Índice de figuras	xxv
Introducción	1
I El valor de la tierra y herramientas para su gestión	5
1. La tierra como factor socio-económico	7
1.1. Entendiendo la situación actual del medio rural español y gallego . . .	8
1.2. Características del medio rural actual gallego	10
1.2.1. Alta fragmentación	11
1.2.2. Abandono del campo y envejecimiento de la población rural .	12
1.2.3. Emigración	13
1.2.4. Intereses sociales	15
1.2.5. Costes de información	16
1.3. Resumen	16

2. Las TIC como medio para paliar la problemática del medio rural	19
2.1. Principales problemas del suelo rural	20
2.2. Sistemas de información geográfica	22
2.2.1. Los SIG aplicados a la Administración de Tierras	23
2.2.2. Ventajas de las herramientas Web-SIG	25
2.3. Estándares, servicios y herramientas SIG	26
2.3.1. Infraestructuras de Datos Espaciales. Iniciativa INSPIRE . . .	27
2.3.2. Servicios OGC	28
2.3.3. Software SIG libre	29
2.3.4. La ausencia de herramientas SIG en la Administración de Tierras	32
2.4. Computación de altas prestaciones	33
2.5. Resumen	36
 II Sistemas de información para el desarrollo rural	 37
 3. Sistema de información Web-SIG para la Administración de Tie- rras	 39
3.1. Arquitectura del sistema	40
3.2. Componentes del sistema	42
3.2.1. Aplicación web Java	42
3.2.2. Visor geográfico	42
3.2.3. Servidor geográfico	44
3.2.4. Almacenamiento	44
3.2.5. Integración de componentes	45

3.3. Fuentes de datos geográficos	46
3.4. Detalles de implementación de las funcionalidades SIG	46
3.4.1. Control de acceso a la información geográfica	46
3.4.2. Localización de parcelas y obtención de sus datos mediante protocolos estándares	47
3.4.3. Modificación de geometrías	48
3.4.4. Centrado y resaltado de parcelas	50
3.4.5. Agrupamiento de capas geográficas	52
3.5. Resumen	53
 4. SITEGAL: sistema de información Web-SIG para la gestión del Banco de Tierras de Galicia	 55
4.1. Impedimentos para la movilidad de tierras en Galicia	56
4.2. El Banco de Tierras de Galicia y SITEGAL	57
4.3. Sistema de roles y funcionalidades del sistema	59
4.3.1. Usuarios no autenticados	61
4.3.2. Usuarios autenticados	61
4.3.3. Gestores	62
4.4. Evaluación del sistema	63
4.4.1. Evaluación de rendimiento	63
4.4.2. Evaluación de la utilización	65
4.5. Sistema de información Web-SIG para la gestión común de parcelas forestales	66
4.6. Resumen	68

III Metodologías y algoritmos para los procesos de ordenación y planificación del territorio 71

5. Algoritmos de búsqueda heurística aplicados a problemas de asignación espacial 73

- 5.1. Contexto de la investigación sobre algoritmos para asignación espacial 74
- 5.2. Algoritmos genéticos en problemas de asignación espacial 76
 - 5.2.1. Funcionamiento de un algoritmo genético 77
- 5.3. Algoritmo *greedy* iterativo basado en población 78
 - 5.3.1. Funcionamiento de un PBIG 79
- 5.4. Características comunes en los algoritmos implementados 79
 - 5.4.1. Codificación de los individuos 80
 - 5.4.2. Relaciones de vecindad entre parcelas 80
- 5.5. Tecnologías utilizadas 83
- 5.6. Resumen 83

6. Algoritmo PBIG para la delimitación de núcleos rurales 85

- 6.1. Criterios para la demarcación de núcleos rurales 87
 - 6.1.1. Marco legal 88
 - 6.1.2. Nueva metodología propuesta 93
- 6.2. Etapa de preprocesado 96
 - 6.2.1. Datos de entrada de la etapa de preprocesado 97
 - 6.2.2. Datos de salida de la etapa de preprocesado 97
 - 6.2.3. Fases de la etapa de preprocesado 99
- 6.3. Algoritmo PBIG para la delimitación de núcleos rurales 104

6.3.1.	Datos de entrada del algoritmo PBIG	105
6.3.2.	Datos de salida del algoritmo PBIG	107
6.3.3.	Descripción de las clases	107
6.3.4.	Cálculo de las aptitudes de las parcelas	110
6.3.5.	Validación de individuos	110
6.3.6.	Evaluación de individuos	111
6.4.	Fases del algoritmo PBIG para la delimitación de núcleos rurales . . .	112
6.4.1.	Validación del Máx-NR	112
6.4.2.	Inicialización de la población	113
6.4.3.	Operador <i>borrar-y-añadir</i> (<i>remove-and-add</i> , <i>R&A</i>)	115
6.4.4.	Obtención de la solución final	116
6.5.	Casos de estudio	118
6.5.1.	Núcleo rural de Viladonga	119
6.5.2.	Núcleo rural de Ferreira	120
6.5.3.	Núcleo rural de Saa	121
6.5.4.	Análisis del comportamiento del algoritmo PBIG en casos reales	125
6.5.5.	Cálculo del entorno inmediato de un núcleo rural	125
6.6.	Resumen	129
7.	Algoritmo genético paralelo para la ordenación del territorio	133
7.1.	Definición del problema de ordenación territorial	135
7.2.	Etapas de preprocesado	137
7.2.1.	Datos de entrada de la etapa de preprocesado	137
7.2.2.	Datos de salida del preprocesado	138

7.3.	Algoritmo genético para la asignación de usos del suelo	138
7.3.1.	Datos de entrada del algoritmo genético	138
7.3.2.	Datos de salida del algoritmo genético	139
7.3.3.	Fases del algoritmo genético	139
7.3.4.	La función de fitness	143
7.4.	Paralelización del algoritmo genético para la asignación de usos del suelo	145
7.4.1.	Algoritmo genético paralelo en sistemas multi-núcleo	145
7.4.2.	Algoritmo genético paralelo en sistemas clúster	146
7.4.3.	Algoritmo genético paralelo en sistemas clúster con nodos multi-núcleo	148
7.5.	Casos de estudio	149
7.5.1.	Influencia del tamaño de la población	151
7.5.2.	Influencia de la tasa de mutación	151
7.5.3.	Comparación entre las funciones de fitness	153
7.5.4.	Rendimiento del algoritmo paralelo	158
7.5.5.	Simulación de un caso real sobre el municipio de Guitiriz	166
7.6.	Funcionalidades adicionales	167
7.6.1.	Checkpointing	169
7.6.2.	Interfaces de usuario	170
7.7.	Resumen	172
8.	Aportaciones y trabajo futuro	173
8.1.	Gestión de elementos con información geográfica	174
8.2.	Procesos de demarcación y ordenación territorial	175

8.3. Epílogo	177
A. Publicaciones de la tesis	179
A.1. Artículos en revistas	179
A.2. Congresos nacionales e internacionales	180
B. Encuesta a técnicos de los posibles impedimentos para la movilidad de tierras	183
Bibliografía	185

Índice de tablas

1.1. Número de explotaciones gallegas y su superficie en distintos años, según el INE.	13
3.1. Herramientas utilizadas en el desarrollo de la aplicación web Java. . .	43
4.1. Principales obstáculos en el arrendamiento de parcelas.	57
4.2. Tiempo medio de petición de una ejecución (en segundos).	64
4.3. Parcelas en el fondo de tierras del Banco y arrendamientos realizados a través de SITEGAL en mayo de 2013.	67
6.1. Resumen de las siglas utilizadas en el Capítulo 6.	89
6.2. Datos de entrada de la etapa de preprocesado para la delimitación de núcleos rurales.	98
6.3. Datos de salida de la etapa de preprocesado para la delimitación de núcleos rurales.	98
6.4. Datos de entrada del algoritmo PBIG de delimitación de núcleos rurales.	107
6.5. Datos de salida del algoritmo PBIG para la delimitación de núcleos rurales	107
6.6. Significados de las variables y funciones del pseudocódigo del algoritmo PBIG para la delimitación de núcleos rurales.	113

6.7. Configuración de las ejecuciones en los casos de prueba del PBIG para la demarcación de núcleos rurales.	120
7.1. Datos de entrada de la etapa de preprocesado para la delimitación de núcleos rurales.	140
7.2. Comparación de los valores de los distintos componentes de la función objetivo utilizando distribución aleatoria y compacta (I).	156
7.3. Comparación de los valores de los distintos componentes de la función objetivo utilizando distribución aleatoria y compacta (y II).	156
B.1. Puntuaciones (1-5) de las restricciones que dificultan la movilidad geográficas (RF: Restricciones Físicas, CT: Costes de Transacción, RPO: Reparos de la Parte Ofertante).	184

Índice de figuras

1. Porcentajes sobre la ocupación, población, VAB y empleo del medio rural diseccionado por países de la Unión Europea.	3
1.1. Número de explotaciones agrarias (A) y porcentaje (B) clasificados por superficie en Galicia en el año 2009, según el INE.	12
1.2. Pirámide de población gallega.	14
1.3. Distribución de la población gallega según edad y número de habitantes de la localidad donde residen.	15
3.1. Arquitectura del sistema.	41
3.2. Patrones de diseño utilizados en la aplicación Java.	44
3.3. Diagrama de secuencia para el control de acceso a las capas de información.	48
3.4. Diagrama de secuencia de las comunicaciones WFS con el servidor local (A) y con el servidor de Catastro (B), para obtener la capa de usos del suelo	50
3.5. Visor geográfico en modo edición con la geometría original (línea amarilla).	51
3.6. Diagrama de secuencias para el centrado y resaltado de una parcela identificada mediante sus códigos catastrales.	52

3.7. Diagramas de secuencia comparando la utilización del WMS en cascada (A) y no utilizandolo (B).	53
4.1. Lista de parcelas y el visor geográfico de SITEGAL.	60
5.1. Arrays utilizados para almacenar la información sobre los vecinos. . .	82
6.1. Buffers alrededor de las edificaciones seleccionadas y la delimitación final.	103
6.2. Diagrama de clases utilizadas en el algoritmo PBIG de la delimitación de núcleos rurales.	109
6.4. Operador R&A aplicado a un individuo (dos ejecuciones consecutivas).	117
6.5. Valores de fitness de las ejecuciones del algoritmo PBIG sobre el asentamiento rural de Viladonega.	121
6.6. Demarcación del asentamiento rural de Viladonga obtenido por la solución acorde a la ley y por B1.	122
6.7. Valores de fitness de las ejecuciones del algoritmo PBIG sobre el asentamiento rural de Ferreira.	123
6.8. Demarcación del asentamiento rural de Ferreira obtenido por la solución acorde a la ley y por A2.	124
6.9. Valores de fitness de las ejecuciones del algoritmo PBIG sobre el asentamiento rural de Saa.	126
6.10. Demarcación del asentamiento rural de Saa obtenido por la solución acorde a la ley y por A1.	127
6.11. Evolución del fitness a lo largo de las iteraciones.	128
6.12. Delimitación del núcleo rural de Donis (Lugo).	130
6.13. Delimitación y entorno inmediato del núcleo rural de Donis (Lugo). .	130

7.2. Balanceo de carga entre threads esclavos en el algoritmo paralelo multi-núcleo.	146
7.3. Flujo de trabajo del algoritmo paralelo multi-núcleo.	147
7.4. Flujo de trabajo del algoritmo genético paralelo para sistemas clueter	149
7.5. Flujo de trabajo del algoritmo genético paralelo híbrido.	150
7.6. Valor de la función de fitness frente al tamaño de la población.	152
7.7. Valor de la función de fitness frente a tasas de mutación estáticas y adaptativas.	153
7.8. Diagramas de caja con los valores de fitness obtenidos en las pruebas del tamaño de la población (A) y la tasa de mutación (B).	154
7.9. Parcelario con distribución aleatoria (A) y distribución compacta (B).	155
7.10. Solución obtenida para la configuración <i>Apt</i> : distribución aleatoria (A) y distribución compacta (B).	159
7.11. Solución obtenida para la configuración <i>CC</i> : distribución aleatoria (A) y distribución compacta (B).	159
7.12. Solución obtenida para la configuración <i>Apt:CC 25 %-75 %</i> : distribución aleatoria (A) y distribución compacta (B).	160
7.13. Solución obtenida para la configuración <i>Apt:CC 50 %-50 %</i> : distribución aleatoria (A) y distribución compacta (B).	160
7.14. Solución obtenida para la configuración <i>Apt:CC 75 %-25 %</i> : distribución aleatoria (A) y distribución compacta (B).	161
7.15. Solución obtenida para la configuración <i>CM</i> : distribución aleatoria (A) y distribución compacta (B).	161
7.16. Solución obtenida para la configuración <i>Apt:CM 25 %-75 %</i> : distribución aleatoria (A) y distribución compacta (B).	162
7.17. Solución obtenida para la configuración <i>Apt:CM 50 %-50 %</i> : distribución aleatoria (A) y distribución compacta (B).	162

7.18. Solución obtenida para la configuración <i>Apt:CM 75 %-25 %</i> : distribución aleatoria (A) y distribución compacta (B).	163
7.19. Solución obtenida para la configuración <i>CC+CM:50 %-50 %</i> : distribución aleatoria (A) y distribución compacta (B).	163
7.20. Speed-up para las versiones paralelas multi-núcleo con respecto a la versión secuencial.	165
7.21. Speed-up (A) y eficiencia (B) en la iteración 20.000 de la versión paralela multi-núcleo con respecto a la versión secuencial.	165
7.22. Valor de la función de fitness para las versiones paralelas multi-núcleo e híbrida del algoritmo genético.	167
7.23. Resultado propuesto por el algoritmo en un caso real.	168
7.24. Comparación entre la solución propuesta por el algoritmo (A) y la solución propuesta por expertos (B).	169
7.25. Interfaz gráfica de gvSIG con SEXTANTE.	171

Introducción

“El 80 % de los efectos vienen del 20 % de las causas”. Principio de Pareto.

Este postulado enunciado por Vilfredo Pareto¹ a finales de 1800 después de observar cómo el 80 % de la tierra italiana estaba en posesión del 20 % de la población y de comprobar que el 20 % de las vainas que plantaba en su jardín contenían el 80 % de los guisantes que consumía, se conoce como *la regla del 80-20* o *el principio de Pareto*.

Sus aplicaciones son numerosas y con ellas sus formulaciones, desde las muy genéricas a las centradas en un determinado asunto: el 20 % de la población tiene el 80 % de riqueza mundial, el 80 % de las ganancias vienen del 20 % de los clientes, el 80 % del esfuerzo en el desarrollo de un software informático genera el 20 % del código, o el 20 % de los riesgos en el lugar de trabajo genera el 80 % de las bajas laborales [79]. Por supuesto este principio trabaja con datos aproximados y lo que refleja es la idea de que unos pocos generan mucho, o viceversa.

El mundo rural no es ajeno a este principio, y el que consta a continuación adquiere mucha importancia por sus implicaciones.

“El 20 % de la población vive en el medio rural y este ocupa el 80 % del territorio mundial” [59].

Estudiando ya datos reales, la Figura 1 muestra porcentajes de ocupación del territorio, población, Valor Añadido Bruto (VAB) y empleo que genera el medio rural en la Unión Europea, diseccionados por países². En la figura PR significa

¹Vilfredo Pareto fue un economista, sociólogo y filósofo italiano que vivió entre mediados del siglo XIX y principios del XX

²Los datos han sido obtenidos de <http://ec.europa.eu/agriculture/statistics/rural-development/2012/>

Principalmente Rural, IR Rural Intermedio, PU Principalmente Urbano, y EU-12, EU-15 y EU-27 indican los países que se han tenido en cuenta³.

Estos datos muestran como, por ejemplo, en la Europa de los 15 (EU-15) compuesta por los países más desarrollados, en las zonas principalmente urbanas vive el 46,2 % de la población, se genera más de la mitad del VAT y casi la mitad del empleo, y todo ello ocupando solamente el 10,1 % del territorio. Mientras los flujos migratorios avanzan hacia las ciudades, las zonas rurales se van despoblando poco a poco con el consecuente abandono de explotaciones, granjas, fincas, etc. El objetivo, por tanto, es intentar sacar más provecho del 89,9 % del territorio restante mediante una óptima gestión de sus recursos.

La tierra es un importante factor de producción que no debe ser utilizado sólo para el auto-suministro o para la producción de bienes, sino que ha de ser considerado como un bien social y ambiental que sea aprovechable por toda la sociedad, y así convertirlo en un motor económico capaz de aportar riqueza a su entorno. Es por esto que se hace indispensable la utilización de nuevos mecanismos, herramientas y protocolos para conseguir una buena gestión del campo y aprovechar todo su potencial.

Objetivos de la tesis

Es pues objetivo de esta tesis crear y aplicar nuevas soluciones capaces de manejar datos e información espacial que ayuden a crear, dinamizar, optimizar y promover acciones sobre el mundo rural, generar riqueza a partir de la tierra, ayudar a su gestión a través de la Administración de Tierras y promover la utilización de las nuevas tecnologías en la comunicación bidireccional entre ciudadanos y administración pública.

Durante la lectura de la tesis se podrán encontrar soluciones a problemas actuales que tienen relación con los ambientes rurales. La ordenación territorial, la

³EU-12: Bélgica, Grecia, Luxemburgo, Dinamarca, España, Holanda, Alemania, Francia, Portugal, Irlanda, Italia, Reino Unido.

EU-15: EU-12 + Austria, Finlandia, Suecia.

EU-25: EU-15 + Polonia, República Checa, Chipre, Letonia, Lituania, Eslovenia, Estonia, Eslovaquia, Hungría, Malta.

EU-27: EU-25 + Bulgaria, Rumanía.

Country	Context 2 - Importance of rural areas - NUTS 3											
	% Territory			% Population			% GVA			% Employment		
	2008			2008			2008			2008		
	% PR	% IR	% PU	% PR	% IR	% PU	% PR	% IR	% PU	% PR	% IR	% PU
Belgium	33.8	31.8	34.4	8.6	23.8	67.5	5.5	19.1	75.5	6.7	20.6	72.6
Bulgaria	53.6	45.1	1.2	38.8	44.9	16.3	25.5	37.0	37.5	33.3	42.9	23.7
Czech Republic	48.3	37.1	14.6	33.2	43.4	23.4	27.1	37.0	35.9	31.9	40.2	27.9
Denmark	71.8	27.0	1.2	42.8	36.0	21.3	39.0	31.3	29.7	40.3	32.7	27.0
Germany	39.8	48.4	11.8	17.4	40.0	42.6	14.6	35.9	49.5	15.8	38.3	45.9
Estonia	82.3	17.7		48.2	51.8		32.3	67.7		42.9	57.1	
Ireland	98.7		1.3	72.6		27.4	60.0		40.0	67.9		32.1
Greece	82.2	12.1	5.6	43.0	10.5	46.5	36.6	10.0	53.4	40.8	10.8	48.4
Spain	46.1	39.5	14.4	13.2	38.3	48.5	10.8	35.6	53.6	12.0	36.4	51.6
France	64.6	27.3	8.1	28.7	35.7	35.6	22.4	31.5	46.1	26.4	34.0	39.5
Italy	45.5	42.3	12.3	20.5	44.0	35.5	18.6	42.6	38.8 ²⁰⁰⁷	19.4	43.5	37.2 ²⁰⁰⁷
Cyprus		100.0			100.0			100.0			100.0	
Latvia	62.8	21.1	16.1	38.2	13.4	48.4	22.7	10.4	66.9	35.9	12.8	51.3
Lithuania	65.0	19.9	15.0	43.5	31.3	25.3	30.1	30.7	39.2	41.2	31.0	27.8
Luxembourg		100.0			100.0			100.0			100.0	
Hungary	66.3	33.1	0.6	47.3	35.7	17.0	34.4	27.9	37.7	43.3	31.7	24.9
Malta			100.0			100.0			100.0			100.0
Netherlands	2.2	51.5	46.3	0.7	28.2	71.2	0.8	25.9	73.3	0.6	26.0	73.4
Austria	72.2	18.9	8.8	39.3	26.5	34.2	30.4	28.9	40.7	34.8	29.6	35.6
Poland	55.6	34.5	9.9 ^{GISCO}	37.9	33.8	28.3	27.5	30.8	41.7	35.4	32.0	32.7
Portugal	84.1	8.7	7.3	36.2	15.2	48.5	30.1	11.4	58.5	35.1	14.7	50.1
Romania	59.3	39.9	0.8	45.8	43.8	10.4	32.1	42.6	25.3	41.7	46.7	11.6
Slovenia	61.0	39.0		43.2	56.8		36.3	63.7		40.1	59.9	
Slovakia	59.0	36.8	4.2	50.4	38.3	11.4	40.8	33.1	26.2	44.5	36.3	19.2
Finland	83.3	14.6	2.1	43.0	30.7	26.3	36.5	28.0	35.5	39.6	29.1	31.3
Sweden	52.6	45.8	1.6	22.5	56.1	21.3	19.8	51.1	29.1	21.6	54.4	24.1
United Kingdom	27.4	47.0	25.6	2.9	25.9	71.2 ²⁰⁰⁷	1.9	21.9	76.1	3.0	26.2	70.8
EU-27	56.6	34.3	9.2	23.6	35.5	40.9	16.9	32.0	51.1	21.6	34.5	43.9
EU-15	56.0	33.9	10.1	19.2	34.6	46.2	15.9	31.7	52.5	17.6	33.6	48.8
EU-12	58.4	35.3	6.3	40.7	38.6	20.7	29.4	35.9	34.7	37.2	38.0	24.7

Figura 1: Porcentajes sobre la ocupación, población, VAB y empleo del medio rural diseccionado por países de la Unión Europea.

delimitación de los núcleos rurales y la dinamización de los mercados de tierras son los principales temas que se abordan.

Ámbito socio-económico de la tesis

El lector podrá notar que la problemática y los casos de estudio explicados en esta tesis se enmarcan en la situación del medio rural de Galicia. No obstante todas las soluciones implementadas se han diseñado de forma que sean fácilmente exportables a otros territorios. Como se verá, muchos de los problemas del medio rural gallego son iguales a los que están empezando a sufrir otras regiones como las del centro y este de Europa. Es por ello que algunas de las soluciones aportadas hayan recibido muestras de interés por parte de las administraciones de otros países, pudiéndose realizar su adaptación de manera sencilla.

Estructura de la tesis

La tesis consta de cuatro partes bien diferenciadas: la Parte I detalla la situación actual y los problemas que tiene el medio rural e introduce las herramientas SIG y de computación de altas prestaciones que se utilizarán en esta tesis; la Parte II se centra en el diseño e implementación de herramientas Web-SIG que aplicados a problemas concretos del campo pueden suponer una reactivación económica del entorno; la Parte III explica el desarrollo de algoritmos de búsqueda heurística que serán utilizados en procesos de planificación y demarcación del territorio; el Capítulo 8 reflexiona sobre todo el trabajo realizado en la tesis; y por último se incluye el Apéndice A, donde se enumeran las publicaciones que ha dado lugar todo el trabajo que ha dado la tesis, y el Apéndice B, que aporta información más concreta sobre uno de los temas tratados en la Parte II.

Parte I

El valor de la tierra y herramientas para su gestión

Capítulo 1

La tierra como factor socio-económico

La llegada de la Revolución Industrial en la segunda mitad del siglo XVIII supuso una gran transformación a todos los niveles (social, económico, tecnológico, cultural) de la sociedad tal y como se conocía. El trabajo, hasta ahora manual, se mecanizó y dio lugar a una economía basada en la industria, con una mayor capacidad de producción y eficiencia. Cuando este proceso de mecanización llegó al mundo rural, la agricultura sufrió la llamada Revolución Agrícola, haciendo más eficiente la producción de alimentos con menos mano de obra. La población rural, al tener cada vez más difícil el encontrar un trabajo agrícola, vio una oportunidad de encontrar empleo en las ciudades gracias al auge de la industria, dando así lugar al llamado *éxodo rural*.

Una de las consecuencias más palpables de este éxodo hacia las ciudades es el abandono de las tierras por parte de los agricultores. Esto afecta principalmente a la disminución de la actividad económica agraria y también a la propiedad de la tierra: las administraciones tienen dificultades en saber a quién pertenecen las parcelas, potenciales compradores no saben a quién dirigirse, la gestión común de la tierra por parte de la administración se complica y la movilidad de tierras se estanca.

Otro aspecto asociado al mundo rural es la dispersión de la información, y no sólo sobre la propiedad, sino también sobre las infraestructuras. Los datos no suelen estar centralizados ni tampoco ser muy precisos, lo que dificulta la inclusión de nueva

información, su utilización, su actualización o su corrección.

Esto también afecta a los procesos de planeamiento y ordenación territorial que, al ser muy laboriosos por involucrar muchas parcelas y muchos intereses, suelen consumir muchos recursos de las administraciones públicas. Elaborar planes que satisfagan a todas las partes no suele ser tarea sencilla. Además, los propietarios de las tierras suelen ser reacios a cualquier cambio que afecte a sus propiedades.

La información con la que se trabaja en estos procesos tienen una gran componente geográfica, por lo que la aplicación de herramientas con capacidades espaciales al medio rural puede ayudar a paliar estas situaciones que poco a poco van menguando la actividad económica del campo.

Para entender la situación actual del medio rural español, y concretamente el gallego, es aconsejable realizar una revisión histórica. Es por esto que la Sección 1.1 está dedicada al relato de los acontecimientos más importantes de los últimos siglos relacionados con el medio rural, con el campo, con sus tierras y su propiedad. Por su parte, la Sección 1.2 se centra en los problemas del medio rural gallego por ser este el ámbito de aplicación de las herramientas desarrolladas.

1.1. Entendiendo la situación actual del medio rural español y gallego

La tierra siempre ha sido un bien muy valioso para el ser humano. Su organización y explotación tienen consecuencias directas sobre el modelo de vida de una sociedad. Desde los nómadas en su incansable búsqueda de tierras más fértiles, hasta el sedentarismo iniciado hace 10 mil años gracias a la habilidad aprendida de cultivar los campos, el tratamiento de la tierra ha influido de manera determinante en todas las culturas que han habitado el planeta. En los siguientes párrafos se destacan de forma resumida principales acontecimientos que han afectado al suelo español, y más concretamente al gallego, desde que Godoy comenzó la llamada desamortización hasta las reformas de las Políticas Agrarias Comunes que en la actualidad se están llevando a cabo a nivel europeo.

Considerada por los economistas clásicos como uno de los tres factores de pro-

ducción junto al trabajo y al capital, la tierra y en concreto su propiedad, ha sido fuente de muchas disputas a lo largo de la historia. La *Revolución liberal*, que cierra la Edad Moderna y da lugar a la Edad Contemporánea, tuvo como clave las desavenencias entre los campesinos y los señores feudales por la propiedad de las tierras que los primeros cultivaban y los segundos amortizaban.

Por esa misma época, a finales del siglo XVIII, comenzó la *desamortización española* [131], un largo proceso que duró hasta la segunda década del siglo XX, donde se ponían en el mercado mediante subasta pública las tierras pertenecientes a la nobleza y a la Iglesia Católica, y los bienes de propios y comunales (tierras propiedad del Estado y municipios). Godoy (que inició el proceso en 1798), Mendizábal (que realizó otra desamortización en 1835) y Madoz (que hizo lo propio en 1855) son los tres nombres más destacados en la desamortización española y cuyas consecuencias no fueron todo lo buenas que se deseaban ya que, según [70], se acabó creando una nueva clase de propietarios formada por emprendedores en el sector de la agricultura y del comercio, lo que supuso que las tierras comunales que servían principalmente para pasto pasaran a manos privadas.

En Galicia, desde la Edad Media, la mayor parte de las tierras estaban bajo el control de la Iglesia. El sistema agrario utilizado era el *foro*, un contrato a largo plazo por el cual el titular de la tierra se la cedía a un campesino que la cultivaba y este le daba una parte de lo producido, los *diezmos*. Debido al pequeño tamaño de las explotaciones gallegas y a esos diezmos, la agricultura en Galicia era básicamente de subsistencia, por lo que la desamortización en Galicia fue algo distinta al resto de España: se subastaron las tierras que eran rentas forales, no las propiedades de la tierra; y las pocas propiedades de los ayuntamientos disminuyeron el impacto de la desamortización municipal.

Los distintos procesos de desamortización fracasaron ya que ninguno de los objetivos fundamentales fue cumplido. El principal, que era la creación de pequeños y medianos propietarios, era una idea utópica ya que un campesino con un nivel económico de subsistencia no podría comprar los terrenos desamortizados [123]. El problema de la propiedad de la tierra, lejos de desaparecer, se convierte en una situación conflictiva.

Nada más proclamarse la II República en 1931 comienzan a promulgarse decretos

con el objetivo de preparar una reforma agraria y evitar una posible sublevación en el campo. En 1932 se promulga la Ley de Reforma Agraria de España con la que se expropiaba *con indemnización justa* las tierras y con la que se buscaban tres objetivos principales: evitar el paro obrero en el campo, distribuir la tierra y racionalizar la economía agraria. Pero esta reforma tenía dos graves problemas: su coste (por las indemnizaciones) y la tardanza en la entrega de las tierras, lo que causó un malestar general entre los jornaleros del campo que acabó en una aguda agitación social. Finalmente, la Ley de Reforma Agraria fue derogada durante la Guerra Civil en aquellas zonas bajo control de los sublevados, y en todo el territorio español cuando acabó con la victoria del bando nacional.

En los años 40 la economía agraria se caracterizó por la actitud intervencionista del Estado [29]. Esto, sumado al aislamiento sometido por la comunidad internacional al régimen franquista, la política comercial exterior española sumió en una profunda crisis al campo español. En los años 50 y viendo el fracaso de la política intervencionista, se reorientó el sector agrícola hacia su modernización, lo que supuso un nuevo esplendor de la agricultura tradicional. Es a finales de los años 70 y principios de los 80 cuando se produjo un verdadero cambio en la política agraria española gracias a una gestión más eficiente de los recursos del país. En la actualidad, la agricultura española está en un nivel de modernización similar al de las agriculturas europeas más avanzadas. Sin embargo, la reforma de las Políticas Agrarias Comunes (PAC) y la crisis económica actual (que se ve reflejada en las ayudas y subvenciones al sector) pondrán una vez más a prueba a la agricultura española.

1.2. Características del medio rural actual gallego

Galicia es una comunidad autónoma española situada al noroeste de la Península Ibérica. Tiene una superficie de 29.574 km², cuenta con 2.797.653 habitantes¹, y consta de 3.778 parroquias distribuidas en 315 municipios. A su vez, 11.525.007 de las 12.965.481 parcelas que hay en Galicia están en suelo rural y 1.440.474 en suelo urbano².

¹Según el censo del Instituto Nacional de Estadística (INE) del 2010

²Según datos de Catastro de 2009. URL: http://www.catastro.meh.es/esp/estadisticas_6.asp

Estos datos concluyen que por cada habitante existen casi 5 parcelas, lo que da una idea de la alta fragmentación del suelo gallego, suelo que se caracteriza también por su minifundismo, su cada vez más escasa densidad de población y por un mercado de tierras deficiente.

1.2.1. Alta fragmentación

El paisaje gallego tiene forma de un complejo mosaico, a diferencia del resto de España que suele ser más regular. La causa es un elevado grado de subdivisión parcelaria debido principalmente a que en Galicia existe una tradición histórica de dividir las parcelas entre todos los herederos en los procesos de herencia. Este hecho es contrario al llamado mayorazgo habitual en otras regiones, mediante el cual los bienes pasaban a un sólo hijo, normalmente al primogénito, y así el patrimonio de la familia no se diseminaba.

Esta gran fragmentación de la tierra hace que los agricultores necesiten a menudo varias parcelas para reunir suficiente superficie como para construir explotaciones rentables económicamente.

A su vez, la excesiva cantidad de parcelas hace que el estado de la propiedad en Galicia también se vea afectado y esté definido por la existencia de muchas propiedades pero demasiado pequeñas. Esto implica que de media cada parcela tenga unos 2500 m² y que cada titular posea entre 6 y 8 parcelas de media [34]. Hay incluso numerosos municipios con más titulares de parcelas rústicas que habitantes. Estas circunstancias complican la cesión y el arrendamiento de las tierras abandonadas, lo que hace que algunas nunca se lleguen a alquilar a agricultores que las necesitan para ampliar sus explotaciones.

Las explotaciones agrarias gallegas tienen por lo general poca extensión. La Figura 1.1 muestra el número de explotaciones agrupadas por su superficie (A) y su porcentaje sobre el total (B). Podemos observar que más de un tercio de las explotaciones tienen menos de 3 Ha., y de entre estas, las parcelas de entre una y dos hectáreas son las más habituales en la comunidad gallega.

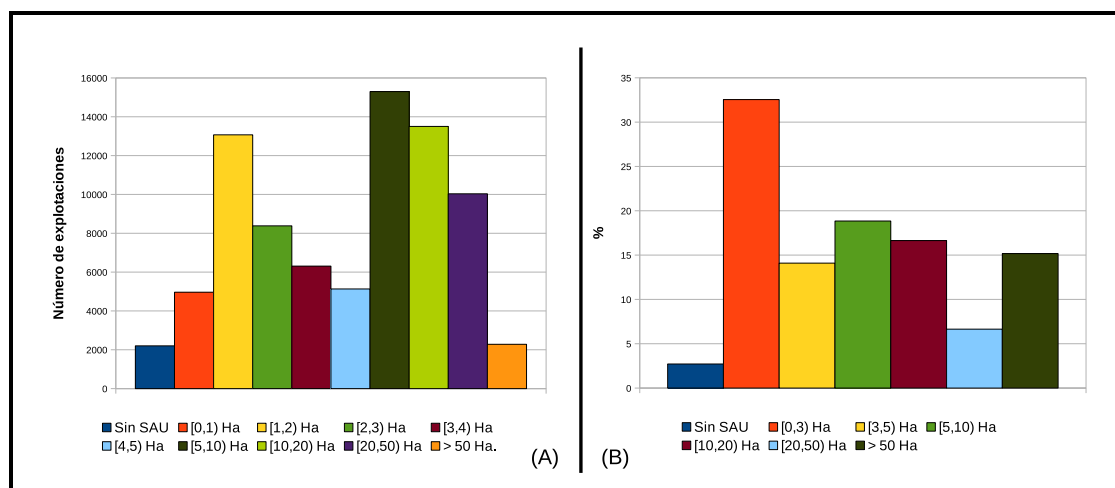


Figura 1.1: Número de explotaciones agrarias (A) y porcentaje (B) clasificados por superficie en Galicia en el año 2009, según el INE.

1.2.2. Abandono del campo y envejecimiento de la población rural

Aunque el saldo migratorio de *lo rural* a *lo urbano* que afectó a Galicia entre los años 50 y 70 se ha estabilizado e incluso invertido, sigue habiendo una continua migración interior en cascada de los pequeños núcleos de población a las cabeceras municipales y de las capitales del interior a las zonas costeras [115].

Los datos mostrados por la Tabla 1.1 confirman este hecho ya que desde el año 2003 hasta el 2009, el número de explotaciones agrarias en Galicia ha disminuido en un 21,48 % y su superficie en un 19,1 %. Consecuencia de esto es que el número de explotaciones sin SAU, aún siendo pocas (sólo un 0,54 % del número total), han crecido en un 188 % en ese mismo periodo de tiempo, y su superficie ha aumentado un 535 %, pasando de las 538 Ha. de 2003 a las 2.877 Ha. del año 2009. SAU, que significa Superficie Agraria Útil, se refiere a las tierras labradas o con pastos permanentes excluyendo las superficies ocupadas por especies forestales o matorrales entre otras.

Estos datos indican que cada vez más la actividad rural va perdiendo fuerza. La gente se traslada a las ciudades y las explotaciones cierran o ven reducida su extensión. Y aunque en algunos casos las superficies liberadas por parte de los agricultores cesantes son transferidas normalmente a otros agricultores, en ocasiones esto no así

Año	Explotac.	Superficie	Explotac. sin SAU	Superficie sin SAU
2009	79.909	914.853 Ha.	428 (0,54 %)	2.877 Ha. (0,31 %)
2007	88.012	1.102.844 Ha.	230 (0,26 %)	507 Ha. (0,05 %)
2003	101.764	1.130.922 Ha.	228 (0,22 %)	538 Ha. (0,05 %)

Tabla 1.1: Número de explotaciones gallegas y su superficie en distintos años, según el INE.

así, y numerosas parcelas cambian a un uso no agroganadero, o directamente se abandonan. Estas dificultades, que ocurren también en algunos países de Europa Central y del Este, se acentúan en Galicia debido a la compleja estructura de la propiedad, caracterizada por la ya comentada elevada fragmentación.

La edad de la población gallega también es un problema ya que el envejecimiento se viene acentuando sobre todo en las poblaciones rurales. En la Figura 1.2 podemos ver la pirámide de población de Galicia según los datos del censo del INE a día 1 de Octubre de 2011. La forma constrictiva de la pirámide indica que la población es generalmente vieja, causa principalmente de una gran esperanza de vida y de una baja tasa de natalidad. La inmigración de los últimos años también influye en la forma de la pirámide.

La Figura 1.3 muestra la distribución de la población gallega según la edad y el tamaño, por número de habitantes, de la ciudad en la que residen. Los datos corresponden al año 2011 según el IGE (Instituto Galego de Estatística). En la gráfica podemos observar cómo el 26,72 % de la población de más de 65 años vive en localidades con menos de 5.000 habitantes. Si nos quedamos sólo con la población rural, este porcentaje aumenta hasta el 35 %³.

1.2.3. Emigración

La emigración también juega un papel fundamental a la hora de explicar la situación actual del suelo gallego. Desde la segunda mitad del siglo XIX y hasta el tercer cuarto del siglo XX, Galicia vivió un movimiento migratorio hacia Latinoamérica en una primera época, sobre todo hacia Argentina, y más tarde hacia países

³Según datos obtenidos en http://www.elpais.com/articulo/sociedad/busca/joven/salvar/campo/elpepusoc/20110202elpepisoc_1/Tes

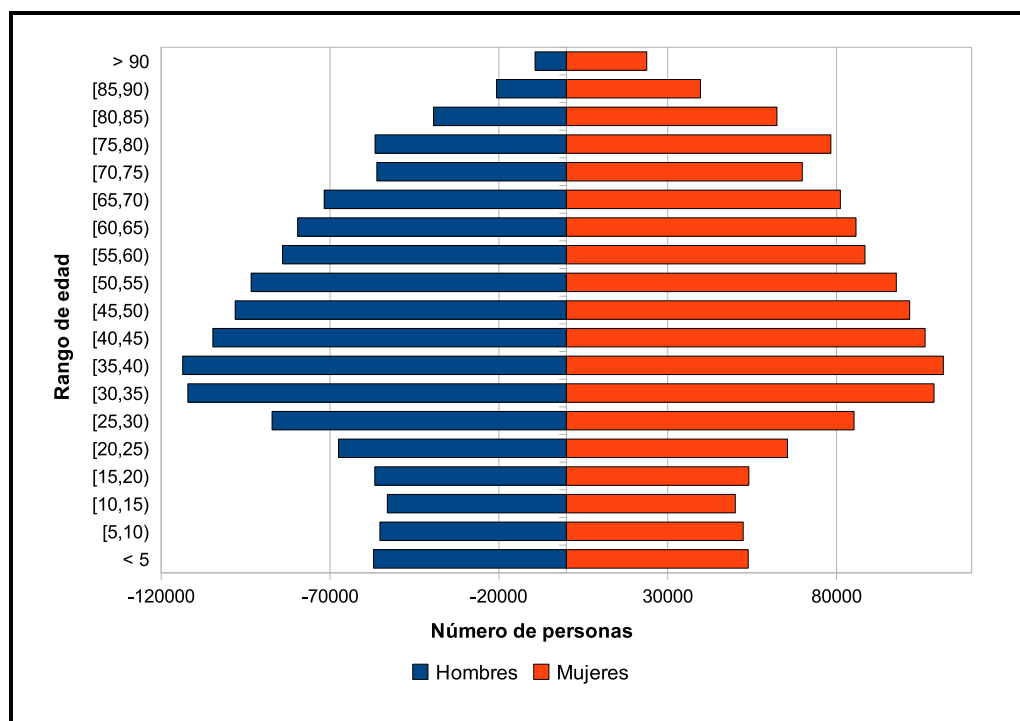


Figura 1.2: Pirámide de población gallega.

europeos como Francia, Alemania o Suiza. Como ya se ha comentado, la tradición de dividir las tierras para repartirlas entre todos los hijos en los procesos de herencia hizo que muchas parcelas fuesen propiedad de gente que había emigrado. Hoy en día hay ocasiones en las que es muy difícil saber a quien pertenece una parcela debido a que esta fue pasando de padres a hijos en el extranjero, dificultando el seguimiento de la propiedad.

Esta también es una de las causas de las dificultades que tiene el mercado gallego de tierras en despegar y conseguir funcionar de manera eficiente. Los agricultores que necesitan ampliar sus explotaciones a menudo tienen dificultades para saber a quien pertenecen las tierras que necesitan, y aunque lo sepan, si los propietarios viven fuera de Galicia puede ser muy complicado contactar con ellos. Y aún así, si lo consiguen, puede ocurrir que ni siquiera los herederos de las tierras sepan que les pertenecen.

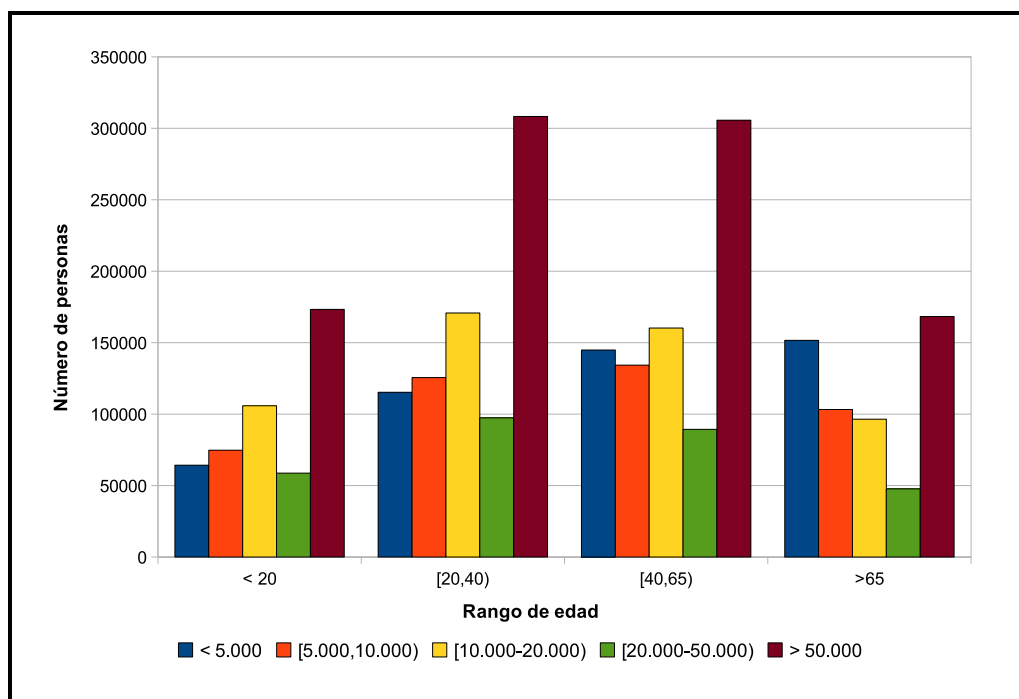


Figura 1.3: Distribución de la población gallega según edad y número de habitantes de la localidad donde residen.

1.2.4. Intereses sociales

Desde siempre la tierra ha sido un motivo de conflicto social y el ámbito rural y agrario gallego no es ajeno a ello. Es por esto que cualquier medida o acción relacionada con la propiedad de la tierra, los límites de las parcelas o la gestión de las mismas, es visto con recelo y con una cierta preocupación por parte de los propietarios. Durante el desarrollo de los proyectos que se explican en los siguientes capítulos de esta tesis, se han encontrado varios ejemplos relacionados con esto.

En los temas relacionados con la dinamización del mercado de tierras mediante la cesión temporal de parcelas y con la gestión común de tierras (ver Parte II de esta tesis), se ha visto que los propietarios tienen la preocupación de que las tierras cedidas no vayan a ser devueltas en el mismo estado en las que se cedieron o que puedan ser estropeadas. Hay propietarios que incluso temen que sus parcelas nunca les sean devueltas o pierdan la titularidad sobre ellas.

En temas referentes a la ordenación y planificación territorial (ver Parte III de

esta tesis), los temores de los propietarios suelen estar relacionados con la pérdida de valor de sus tierras. Por ejemplo, que una parcela quede dentro o fuera de un núcleo rural puede suponer la diferencia entre poder construir edificaciones en ella o no, de que esas edificaciones puedan superar una determinada altura o puedan ser construidas con unos determinados materiales. Otro supuesto es que si una parcela es asignada a una categoría del suelo determinada, puede implicar algunas obligaciones o restricciones sobre las acciones que se quieran efectuar sobre dicha parcela.

1.2.5. Costes de información

Los *costes de información* son aquellos en los que se incurre en determinar si el bien necesitado está en el mercado, quien tiene el mejor precio, etc. Generalmente están incluidos en los conocidos *estudios de mercado*. En lo que se refiere a los problemas planteados del medio rural, estos costes pueden ser muy elevados debido sobre todo a los problemas ya planteados en los puntos anteriores (abandono del campo, la alta fragmentación, emigración) que añaden bastante dificultad a la hora de, primero, encontrar las tierras con las características adecuadas, y segundo, encontrar a los dueños de esas tierras.

También existen otros costes asociados (de negociación, de transacción, etc.) que ya fueron introducidos por Ronald Coase en 1937 [31] y que en el ámbito rural, son muy elevados [143, 11].

1.3. Resumen

En este capítulo se ha descrito la situación actual del medio rural gallego y sus problemas asociados. A través de una revisión histórica se han repasado los acontecimientos más importantes que han afectado al estado de la propiedad de las tierras españolas y gallegas, desde el inicio del periodo de desamortización a finales del siglo XVIII hasta la actualidad.

Consecuencia de todas estas circunstancias históricas es que, a día de hoy, el suelo gallego tiene varias características que dificultan la movilidad de sus tierras y su gestión, lo que fomenta el abandono de parcelas y el cese de la actividad económica

de muchas explotaciones agrarias. La enorme fragmentación también es un obstáculo que dificulta los procesos de planificación rural.

Aunque este capítulo ha centrado el análisis en Galicia por ser la zona de aplicación de las herramientas desarrolladas en esta tesis, esta es comparable con la de otros países como Portugal o las regiones del centro y este de Europa. Por tanto, soluciones tomadas en Galicia para paliar problemas concretos gallegos, son perfectamente exportables a otros entornos ya que los desarrollos realizados han sido abordados de una manera genérica y con capacidades de adaptación a las características concretas de las regiones en las que vayan a ser empleados. Prueba de esto es que algunas de estas herramientas han despertado el interés de organismos internacionales y de países con una problemática similar.

Capítulo 2

Las TIC como medio para paliar la problemática del medio rural

En la actual era de la información en la que el acceso y la gestión de los datos adquieren suma importancia, la implantación y utilización de Tecnologías de la Información y Comunicación (TIC) especializadas, se convierte en fundamental para cualquier órgano de gobierno. Concretamente para el medio rural, su uso puede significar una mejora en la gestión y un mejor aprovechamiento de sus recursos, lo que podría suponer una reactivación económica.

El 70 % de la información de la que disponen los gobiernos es información geoespacial [78], es decir, es información que se puede situar en un mapa. Esta proporción es aún mayor cuando se tratan situaciones relacionadas con la Administración de Tierras que, definida por la UNECE¹ como *el proceso de registrar y diseminar información sobre la propiedad, el valor y el uso de la tierra y sus recursos asociados* [130], se convierte ahora en algo fundamental a la hora de desarrollar políticas relacionadas con la gestión de recursos.

Las instituciones, sobre todo públicas, están cada día más conciencadas en publicar los datos que generan para que puedan ser utilizadas por terceros. Las iniciativas *Open Data*, que trabajan bajo la idea de que los datos generados por organismos públicos han de ser de acceso público, han colaborado enormemente en la liberalización

¹Comisión Económica de las Naciones Unidas para Europa, por sus siglas en inglés.

de este tipo de información. A nivel europeo se han promulgado leyes que regulan la publicación y acceso de los datos públicos. La iniciativa INSPIRE, explicada con más detalles en la Sección 2.3.1, es un claro ejemplo de esta política.

Este capítulo introduce los Sistemas de Información Geográfica (SIG) y argumenta cuales son las ventajas que tiene su utilización dentro de la Administración de Tierras. La Sección 2.1 analiza los problemas más destacados que tiene el medio rural. La Sección 2.2 se centra en los SIG y en las ventajas de su inserción en los procesos de decisión de las instituciones. Para su interacción e integración se suelen seguir estándares que definen los servicios que utilizarán las aplicaciones SIG, que serán explicados en la Sección 2.3. Además, dado que operar con información espacial es muy costoso computacionalmente, suele ser necesario utilizar tecnologías de computación de altas prestaciones relacionadas con el reparto de tareas entre diferentes recursos de computación para poder obtener resultados en tiempos razonables y así estar en disposición de ofrecer productos útiles para el usuario final. La Sección 2.4 introducirá algunas de estas tecnologías de altas prestaciones utilizadas por las herramientas desarrolladas.

2.1. Principales problemas del suelo rural

Tanto el abandono de suelo rural como la paralización de los mercados de tierras son características identificables en regiones de toda Europa, siendo el centro y el este las más representativas de este retraimiento económico [86, 3, 67] junto con España y Portugal [110]. Estos dos conceptos están relacionadas entre sí ya que se espera que los mercados de tierras permitan el ajuste de las estructuras rurales cambiando las condiciones económicas. Esto es, las explotaciones más dinámicas y eficientes podrían aprovecharse de los mercados de tierras para adquirir otras explotaciones más pequeñas y menos eficientes. Sin embargo si estos mercados no funcionan de forma adecuada, puede ocurrir que se desaproveche suelo agrícola al no ser utilizado ni por sus propietarios ni por aquellos agricultores interesados en trabajarlo.

En Galicia ocurre exactamente lo descrito: tiene suelo suficiente como para ampliar la productividad de las explotaciones existentes o para la creación de nuevas actividades económicas agrícolas, pero al mismo tiempo cuenta con una gran canti-

dad de parcelas de alto potencial agrario abandonadas, infra-utilizadas, o dedicadas a otros usos, normalmente forestales. Por esta razón, entre otras, en [85] se concluye que el elemento básico que dificulta la reactivación económica de las explotaciones agrarias es la falta de movilidad de la tierra. Como muestra de esta circunstancia decir que en los últimos 20 años 145.000 hectáreas han sido abandonadas en Galicia [32] aumentando en un 66 % en la última década², mientras que una de las principales quejas del sector agrícola es la necesidad de ampliar sus explotaciones.

Otro impedimento para la movilidad de tierras es el elevado coste de transacción [143, 11], que ya han sido introducidos en el Capítulo 1. Según Coase [31], los costes de transacción están compuestos por los costes de negociación, garantía e información, pero son estos últimos los que adquieren gran importancia precisamente cuando la falta de información es notable [36], ya que aumenta a su vez el consumo de recursos por parte de la persona u organismo interesado. Así pues, los costes de información suponen un freno a todos los actores involucrados en estas transacciones: al agricultor ya no sólo le resulta difícil encontrar una parcela con las características deseadas, si no que cuando la encuentra le puede resultar aún más complicado el ponerse en contacto con sus propietarios; y a los propietarios les faltan herramientas o medios que faciliten la publicación de sus tierras y actúen como un canal comunicador entre propietario y agricultor. Ambos tienen que gastar recursos (sobre todo tiempo y/o dinero) para reunir toda la información necesaria para llevar a cabo la transacción deseada.

Normalmente a mayor fragmentación del suelo (esto es, un alto número de parcelas y propietarios por unidad de superficie), el coste de información aumenta. Lo mismo ocurre cuando la separación geográfica entre las parcelas, los potenciales propietarios y los potenciales agricultores es relativamente grande, algo que pasa a menudo en las zonas rurales debido a los movimientos migratorios [56]. Así pues la reducción de los costes de transacciones es crucial para la mejora de los distintos procesos de la Administración de Tierras, especialmente en aquellas zonas con una alta fragmentación [100, 38].

Esta misma problemática afecta también a la planificación y asignación de categorías al suelo. Y además, en estos procesos también influye la forma de las parcelas:

²Según datos obtenidos en <http://www.farodevigo.es/galicia/2011/08/29/superficie-agraria-abandonada-incremento-66-ultima-decada/575323.html>

cuanta más similitud a una cuadrícula tenga un parcelario, más sencillos serán. En cambio, si el parcelario tiene forma de mosaico irregular, como pasa en Galicia, los procesos se complican.

2.2. Sistemas de información geográfica

Se define un Sistema de Información como *el conjunto de elementos (tecnología, datos, procesos, personas, etc.) que interactúan entre sí con el objetivo de presentar la información de una manera útil y manejable para el usuario final.*

Por su parte, un Sistema de Información Geográfica o SIG³ se puede definir como *un conjunto de equipos electrónicos, aplicaciones y datos geográficos que se utilizan para capturar, gestionar, analizar y mostrar, la información geográficamente referenciada.* Y en un sentido más genérico del término, se puede considerar a un SIG como *la herramienta que permite a los usuarios crear peticiones interactivas y búsquedas personalizadas, analizar información espacial, editar datos y mapas y presentar los resultados de todas esas operaciones* [102].

Los datos con los que trabajan los SIG tienen dos partes diferenciadas de información: sus atributos temáticos, que es información alfanumérica; y una geometría asociada que permite la identificación del objeto y que este sea localizable en un mapa. De esta manera, señalando un objeto en un visor de mapas se pueden conocer sus atributos e, inversamente, preguntando por un dato se puede saber su localización.

Para la creación de información geográfica es necesario un proceso de abstracción de los datos para poder pasar de la complejidad del mundo real a una representación simplificada asequible a los lenguajes de programación. Este proceso comienza con la concepción de la estructura de la información, que normalmente será en capas. Así, según la utilización que le vayamos a dar y de los distintos datos que queramos mostrar, iremos añadiendo capas con la información temática de la zona geográfica en la que estamos interesados. En todo este proceso aparecen dificultades que tienen que ver con la abstracción realizada sobre los datos. El primero de ellos es que, para que las máquinas puedan trabajar con datos geográficos y temáticos, estos han de ser reducidos a formas simples como puntos, líneas o polígonos. El segundo problema

³También conocidos como *GIS* por sus siglas en inglés: *Geographical Information System*.

viene dado por las relaciones existentes entre los objetos geográficos, lo que se conoce como *topología* ya que son muy complejas a nivel computacional.

Existen fundamentalmente dos tipos de almacenamientos de datos: *vectoriales* y *raster*. Las unidades básicas de información geográfica en los datos vectoriales son puntos, líneas y polígonos. Cada una de estas unidades se compone de uno o más pares de coordenadas (relativas a algún sistema de coordenadas). Por otro lado, el modelo raster consiste en dividir cada superficie a representar, en filas y columnas formando una malla o rejilla regular en la que cada celda ha de ser rectangular aunque no necesariamente cuadrada. Cada celda, llamada *píxel*⁴, guarda tanto las coordenadas de su localización como un valor temático. A diferencia de los datos vectoriales la localización de cada píxel es implícita, depende directamente del orden que ocupa en la rejilla.

2.2.1. Los SIG aplicados a la Administración de Tierras

Por definición, la Administración de Tierras ha de ser el elemento central donde se encuentra toda la información relacionada con las tierras, su propiedad y todos los elementos naturales (ríos, relieves, etc.) o artificiales (edificaciones, viales, raíles, etc.), de su entorno. Una buena gestión de esa información tiene impacto directo sobre los niveles micro y macroeconómicos [91, 92] y facilita la implementación de las políticas agrarias y los mercados de tierras [134, 23].

Normalmente los instrumentos que tiene la Administración para reunir la información sobre la tierra es el Catastro y el Registro de la Propiedad. Existe no obstante otra entidad pública que puede jugar un papel importante en las transacciones que involucran tierras. Son los llamados Bancos de Tierras, que pueden ser vistos como un nuevo tipo de actor en los mercados de tierras, una especie de *propietario* que no está interesado en su crecimiento si no en la distribución de sus propiedades a agricultores que necesitan tierras para aumentar su capacidad productiva [41]. De esta manera, el bloqueo causado por factores no económicos y los costes de transacción debería evitarse y así poder crear un mercado de tierras dinámico. En esencia, un Banco de Tierras actúa como intermediario (más o menos activo) que conecta oferta y demanda, fomentando la transacción entre propietarios y agricultores,

⁴Píxel es la abreviatura inglesa de la expresión *picture element*

proporcionando información acerca del uso y del valor de la tierra, y apoyando las transacciones desde una perspectiva legal y administrativa. Los Bancos de Tierra ya están operativos en varias regiones del oeste de Europa mientras que en países del centro y del este está siendo analizada su posible aplicación [41]. El Capítulo 4 trata sobre el diseño e implementación de la herramienta de gestión del Banco de Tierras de Galicia.

Otro objetivo de la Administración de Tierras es la puesta en valor del suelo. Debido a que las tierras con usos agrarios son más susceptibles de ser trabajadas y son más activas en los mercados de tierras, es necesario gestionar de forma adecuada las parcelas forestales. Si una administración es capaz de obtener rendimiento del monte obtendrá notables réditos económicos y reducirá el riesgo de incendio al estar los montes más cuidados. Con esta intención nacen las UXFOR y el sistema de información que las gestiona, SIXUXFOR, que se explicará en el capítulo 4.5.

Para poder sacar un rendimiento al campo es indispensable realizar ocupación racional del territorio, que se consigue mediante la aplicación de reglas y normas que permitan o prohíban unos determinados usos de la tierra. Es lo que se conoce como *ordenación del territorio*. La asignación de categorías al suelo es un laborioso proceso que forma parte de la ordenación. Y es laborioso porque suele involucrar un número elevado de parcelas (dependiendo de la región y el nivel administrativo del territorio que se ordene), un número considerable de restricciones y un número incuantificable de intereses particulares y comunes. La automatización mediante algoritmos heurísticos de este proceso supone una mejora en cuanto a la eficiencia del proceso y un ahorro en cuanto a tiempo. En el Capítulo 7 se explica en detalle un algoritmo genético para procesos de asignación de categorías del suelo, que además es paralelizado para obtener el máximo rendimiento.

Y lo mismo ocurre en los procesos de delimitación de núcleos rurales, una tarea que en muchas ocasiones aún se realiza sobre programas de CAD⁵ o incluso a mano. Y aunque el número de parcelas con las que se trabaja suele ser menor que en los procesos de asignación de categorías debido a la extensión del territorio involucrado, no sólo es conveniente la utilización de programas especializado como los SIG, sino que también resulta importante la aplicación de algoritmos diseñados específicamente para dichos procesos. El Capítulo 7 detalla en profundidad un algoritmo *greedy*

⁵Programas de diseño asistido por computador, por sus siglas en inglés Computer-Aided Design

cuyo objetivo es obtener una buena delimitación para un núcleo rural dado.

Para que la Administración de Tierras pueda funcionar eficientemente, la utilización de los SIG se vuelve crucial. Muchas de las capacidades de los sistemas de administración de tierras modernos recaen sobre un componente catastral, esto es, en información detallada a nivel de parcela [50, 133], y las parcelas tienen una serie de peculiaridades que las convierten en elementos diferentes a otros recursos, por lo que requieren un trato especial. El primer aspecto a tener en cuenta es que una parcela es definida por sus límites, es decir, por información espacial. No existirá una *entidad parcela* (llamésmole así) si sus límites no están claros. Además, no hay dos parcelas iguales en la Tierra ya que para empezar, cada una estaría en una localización diferente. Es por esto que la utilización de herramientas SIG para el manejo y la gestión de información relacionada con parcelas (información de carácter espacial), es clave para obtener un aprovechamiento óptimo de sus datos. Hay que tener en cuenta también que la precisión y la fiabilidad de la información espacial define la identificación de los objetos y sus características. Las herramientas SIG proporcionan además una representación visual de esta información que es de suma importancia para su análisis y la toma de decisiones en base a ella. Esto es mucho más significativo que un simple mapa en el sentido de que conecta espacios identificados con elementos de sobra conocidos como calles, propiedades, lugares de interés, etc. [133].

2.2.2. Ventajas de las herramientas Web-SIG

Combinando los SIG y las tecnologías web se obtiene la infraestructura necesaria, denominada Web-SIG, para compartir y acceder a la información espacial de una manera eficiente. Este hecho es fundamental para alcanzar objetivos impulsores de la economía rural ya que, como se ha comentado anteriormente, todavía existe mucha descentralización de la información. Lo que suponen las herramientas Web-SIG al mercado de tierras puede tener su analogía a lo que supuso internet al comercio convencional. Con el comercio electrónico se redujeron el coste de transacción [122] y además, por parte del cliente, el coste de información y búsqueda [8]. Enemark enfatiza en [49] cómo la separación tradicional de la información geográfica de la alfanumérica supone una considerable devaluación del valor del propio recurso y de

lo importante que puede llegar a ser la utilización de las nuevas tecnologías a la hora de facilitar el acceso a dicha información.

Algo importante a tener en cuenta es que la aplicación de los sistemas Web-SIG a los entornos rurales beneficiaría el mercado de tierras y también tendría efectos positivos sobre el planeamiento urbano y rural o sobre la gestión eficiente de recursos [135]. Por tanto, el poner a disponibilidad de la sociedad toda la información geográfica que se pueda, ayudará a mejorar la toma de decisiones en los temas relacionados con la tierra [93].

Aprovechando la conectividad actual y al hecho de que cada vez hay más información publicada, los SIG son capaces de integrar distintas fuentes de datos (bases de datos externas, servicios web estándares, etc.) para que los usuarios puedan acceder a una gran cantidad de información desde un único punto.

Esta integración de la información también proporciona soluciones con una buena relación beneficio/coste. Por ejemplo, y con respecto al mercado de tierras, según [132] el ahorro viene debido a que se facilitará el acceso y la gestión de todo lo relacionado con la propiedad y los derechos de los propietarios de parcelas. De hecho, la mayoría de los sistemas de administración modernos están basados en información catastral, esto es, en información detallada de las parcelas y sus propietarios [133].

Una vez se estén utilizando las herramientas Web-SIG en el ámbito de la Administración de Tierras, el siguiente paso es su integración con el Gobierno Electrónico, conocido también como *e-Government*, y la Democracia Electrónica, o *e-Democracy*. Estos términos se refieren a la utilización de las TIC tanto en la comunicación bidireccional entre gobierno y ciudadanos/empresas, como en la participación electrónica de los ciudadanos en los procesos democráticos de un país, respectivamente.

2.3. Estándares, servicios y herramientas SIG

Es importante que las fuentes de datos puedan distribuir su información de una manera estándar para que todos los potenciales interesados puedan acceder a ella de una manera tal que, si se decidiera cambiar una fuente de información, sólo se tenga que modificar la dirección de acceso a ella, pero que toda la lectura y procesamiento

de los datos pudiera mantenerse.

Por tanto, lo que se pretende es definir una serie de estándares para que los organismos proveedores de la información geográfica puedan implementar servicios, que serán los utilizados por las herramientas SIG, pudiendo acceder a ellos independientemente de la fuente de datos.

2.3.1. Infraestructuras de Datos Espaciales. Iniciativa INSPIRE

Una Infraestructura de Datos Espaciales (IDE) se puede definir como *una red descentralizada de servidores que incluye: datos y atributos geográficos; metadatos; métodos de búsqueda, visualización y valoración de los datos (cartografía en red y catálogos); y algún mecanismo para proporcionar acceso a los datos espaciales*. Otra definición, aunque esta de un ámbito más organizativo, es la de *un conjunto básico de tecnologías, políticas y acuerdos institucionales destinados a facilitar la disponibilidad y el acceso a la información espacial* [73, 44].

Normalmente, los dueños y/o responsables de datos geográficos suelen ser entidades tanto públicas como privadas. La iniciativa INSPIRE es una propuesta legal a nivel Comunidad Europea que aspira a implementar, de una manera común para todos los estados miembros, infraestructuras con información geográfica que faciliten la compartición de información [52]. La importancia que adquiere la información espacial es tal que los órganos de gobierno de las distintas entidades propietarias están basando su toma de decisiones en estos datos, además de invertir en infraestructura adecuada para su consulta, mantenimiento y publicación [87].

Basado en INSPIRE, España sacó en 2010 la *Ley 14/2010, de 5 de julio, sobre las infraestructuras y los servicios de información geográfica en España (LISIGE)*. Bajo esta ley está el proyecto IDEE⁶, que integra las Infraestructuras de Datos Espaciales (IIDDEE) de todos los productores oficiales de datos a nivel tanto estatal como autonómico y local, y todo tipo de infraestructuras sectoriales y privadas. En el ámbito regional diversas autonomías han desarrollado también sus IIDDEE, siendo

⁶Infraestructura de Datos Espaciales de España. <http://www.idee.es/>

la IDEG⁷ la correspondiente a Galicia.

2.3.2. Servicios OGC

Para favorecer la interoperabilidad entre todos los programas con funcionalidades SIG y el intercambio de información entre ellos, es necesario la definición de estándares que regulen estas acciones. De ello se encarga el Open Geospatial Consortium (OGC), una organización internacional sin ánimo de lucro que estandariza servicios basados en localizaciones e información geoespacial. Creada en 1994 con el nombre de Open GIS Consortium, actualmente reúne a más de 481⁸ organizaciones públicas y privadas que colaboran en la creación de las especificaciones de los estándares [102].

Las especificaciones son documentos accesibles sin coste alguno que detallan las interfaces o la codificación que todo sistema SIG debería seguir para ser compatible con el resto. Es tarea de los desarrolladores el implementar estas interfaces en sus productos o servicios. Idealmente se espera que si dos organizaciones crean productos independientes pero siguiendo estas especificaciones, los componentes creados puedan trabajar juntos sin necesidad de realizar ningún cambio en el código.

De entre las más de 50 especificaciones las siguientes son algunas de las más conocidas y que han sido utilizadas en los diferentes proyectos descritos en esta tesis:

- *Web Map Service (WMS)*: es un servicio de generación de mapas. Los mapas son creados por el servidor en formatos de imagen, como PNG, GIF o JPEG, o en elementos gráficos basados en vectores como SVG (Scalable Vector Graphics). Esta especificación estandariza la forma en la cual los mapas son solicitados por los clientes. Las peticiones se hacen mediante el protocolo HTTP a través de URLs donde se indica mediante parámetros, qué información debe ser mostrada en el mapa, así como el sistema de coordenadas de referencia, la altura y anchura la imagen, etc. Los mapas devueltos se pueden solapar facilitando así la creación de una red de servidores distribuidos.

⁷Infraestructura de Datos Espaciales de Galicia. <http://sitga.xunta.es/>

⁸Cifra obtenida en febrero de 2013.

- *Web Feature Service (WFS)*: es un servicio para la descarga de información espacial de tipo vectorial con la que, a parte de la consulta de atributos, se permite la edición de la información. La comunicación se realiza igual que en WMS pero se suele utilizar el estándar GML para el envío de información geográfica.
- *Geography Markup Language (GML)*: lenguaje basado en XML que permite modelar, transportar y almacenar información geográfica. Al igual que XML es extensible por lo que cualquiera puede crear sus propias entidades.
- *Simple Features - SQL (SFS)*: extensión de SQL que permite la consulta y manipulación de objetos geográficos.
- *Styled Layer Descriptor (SLD)*: esquema XML que describe la apariencia de las capas de los mapas obtenidos a través de WMS. Permite al cliente definir cómo quiere que se le presente la información.
- *Filter Encoding*: permite definir filtros en formato XML. Un filtro es un conjunto de restricciones que se aplican sobre uno o varios atributos de un objeto con el propósito de seleccionar un conjunto limitado de instancias sobre las cuales se va a realizar alguna operación.

2.3.3. Software SIG libre

Se dice que el software es *libre* cuando una vez obtenido se puede usar, copiar, modificar, estudiar y redistribuir sin ninguna restricción, o con las restricciones necesarias para asegurarse de que el usuario final tiene las mismas libertades sobre el producto que los autores originales. Para que sea libre, el software debe ir acompañado de licencias⁹ que especifiquen qué es lo que se permite hacer, y además el código fuente debe estar disponible (o en su defecto debe estar construido en base a unas especificaciones públicas). El software libre suele poder obtenerse a coste cero o a coste de distribución, aunque esto no es estrictamente necesario, ya que puede ser vendido a terceros sin que pierda su carácter libre.

⁹Hay multitud de *licencias de software libre* de las que cabe destacar la GNU GPL, GNU LGPL, BSD License, Apache License, Mozilla Public License (MPL) o MIT License

El software SIG libre toma cada vez mayor relevancia y aunque, en general, no cuenta con todas las funcionalidades que los productos propietarios ofrecen, sí es verdad que se pueden realizar trabajos fiables y con entidad únicamente con la ayuda de programas SIG libres.

En los siguientes apartados se dará una breve descripción del principal software libre que existe en el mundo SIG y que han sido utilizados o probados a lo largo del desarrollo de las herramientas descritas en esta tesis.

Bases de datos espaciales

Para que tanto PostgreSQL como MySQL, los dos sistemas gestores de bases de datos más comunes en el software libre, tengan capacidades espaciales, hay que añadirles extensiones como las siguientes:

- *PostGIS*: extensión espacial para PostgreSQL con licencia BSD. No forma parte del núcleo de PostgreSQL sino que constituye una especie de módulo que se añade a cada base de datos a la que se quiere dotar de capacidades espaciales.
- *MySQLSpatial*: extensión espacial para MySQL con licencia GPL. En este caso es el propio equipo de desarrollo de MySQL quien lo está llevando a cabo y también se basa en la especificación SFS.

Los datos raster se suele almacenar en archivos y para su manejo se suele utilizar, dentro del software libre, GDAL, que es una librería para escritura y lectura de datos con formato raster, que tiene licencia X/MIT.

Servidores geográficos

Los servidores geográficos son los encargados de gestionar la información geográfica, de publicarla y de generar los mapas requeridos por los clientes. Para ello interactúan con el software de almacenamiento recogiendo los datos necesarios. Destacan como servidores libres:

- *MapServer*: está escrito en C y se ejecuta como un CGI (Common Gateway Interface). Soporta algunos estándares del OGC y los formatos SIG más importantes. Trabaja con PostGIS y con GDAL y soporta multitud de lenguajes de *scripting* como pueden ser PHP, Java o Perl entre muchos otros. Licencia BSD.
- *GeoServer*: está escrito en Java y soporta muchos estándares del OGC. Entre ellos se encuentra el WFS-T (Web Feature Service - Transactional), que permite al usuario insertar, actualizar y borrar elementos a través de un visor web o de aplicaciones de escritorio preparadas para ello. Tiene una interfaz gráfica para su configuración y permite al usuario compartir, de manera sencilla y rápida, toda su información geográfica a través de la web. Licencia GPL.

Aplicaciones de escritorio

Las aplicaciones de escritorio nos permiten realizar consultas y visualizar los datos y mapas accediendo directamente a las bases de datos sin necesidad de realizarlas vía web. Algunas de las mas completas son:

- *GRASS*: es sin duda la aplicación libre más importante y completa del mundo SIG. Funciona en varias plataformas a través de una interfaz gráfica y mediante consola de comandos. Cuenta con numerosas funciones de análisis espacial. Licencia GPL.
- *Quantum GIS*: también llamado QGIS, es una aplicación que permite crear y visualizar mapas. Es bastante ligera, ideal para entornos en los que los recursos de memoria y procesador son escasos. El potencial de esta herramienta reside en la posibilidad de personalizarla y extenderla mediante scripts y plugins. Es multiplataforma y puede funcionar como interfaz gráfica de GRASS. Soporta datos raster y vectoriales, y los formatos de archivos espaciales más comunes (ShapeFile, GeoTIFF, ...). Licencia GPL.
- *uDig*: aplicación realizada en Java basándose en la plataforma Eclipse. Aparte de ser una aplicación SIG, también es una plataforma a través de la cual los desarrolladores pueden crear nuevas aplicaciones. Soporta algunos estándares de la OGC como cliente. Licencia LGPL.

- *gvSIG*: aplicación SIG española, desarrollada en la Comunidad Valenciana, multiplataforma, escrita en Java y que soporta estándares de la OGC. Admite datos tanto en formato vectorial como en formato raster. Licencia GPL.

Visores geográficos web

Los visores web nos permiten ver los mapas generados por los servidores geográficos y realizar diversas operaciones a través de un navegador web. *OpenLayers* es de los mas conocidos y utilizados actualmente, lo que implica que tenga una comunidad bastante activa a su alrededor. Basado en JavaScript y en AJAX para la comunicación asíncrona con el servidor, es independiente de la fuente de datos y soporta los estándares WMS, WFS, WFS-T entre otros muchos. Es una herramienta libre que está dentro de la Open Source Geospatial Foundation.

2.3.4. La ausencia de herramientas SIG en la Administración de Tierras

Hemos visto como el disponer de un entorno SIG para mostrar y gestionar datos espaciales es crucial para poder tener una información completa de los objetos consultados, ya sean parcelas, casas, fincas, etc. Pues bien, aún son mayoría las aplicaciones web que teniendo como objetivo la compra-venta/alquiler de propiedades rurales, no tienen algún tipo de herramienta SIG. Las aplicaciones web especializadas en publicaciones de anuncios relacionados con propiedades de inmuebles [62, 27, 72] y con propiedades rurales [55, 54, 80] no tienen un entorno SIG de base. Algunos sitios delegan la funcionalidad SIG a servicios web de terceros como GoogleMaps, que aunque son útiles para propósitos de localización, no ofrecen funcionalidades ni para conocer u operar con datos espaciales propios del objeto buscado ni de su entorno. El alto coste que pueden tener las capas de información y la complejidad de integración con procesos de compra-venta/alquiler, pueden ser algunas de las causas que expliquen la casi total ausencia de sistemas SIG en este tipo de sitios web.

Por otro lado, en los procesos de ordenación del suelo es habitual encontrar a gente trabajando con programas de CAD¹⁰, los cuales no están pensados para esas

¹⁰Programas de diseño asistido por ordenador, de sus siglas en inglés Computer-Aided Design

labores. Es más, hay administraciones donde aún se trazan líneas a mano con la ayuda de reglas, cartabones, escuadras y compases, para comprobar que las restricciones impuestas sobre los procesos de delimitación o asignación se cumplen. Por tanto, la utilización de herramientas SIG específicas para estos casos supondría ganar en eficiencia, en rapidez y en la calidad de los resultados obtenidos.

2.4. Computación de altas prestaciones

La computación de altas prestaciones (más conocida por sus siglas en inglés HPC, de *High Performance Computing*) abarca las tecnologías orientadas al desarrollo de aplicaciones que requieran el procesamiento de grandes volúmenes de datos a través de la utilización de sistemas paralelos o supercomputadores, junto con la explotación de comunicaciones entre los procesos (o threads) involucrados en aplicaciones paralelas y/o distribuidas a través de redes de interconexión de muy baja latencia y gran ancho de banda.

El área de computación y comunicaciones de altas prestaciones abarca un abanico muy amplio de tecnologías para la supercomputación y la realización de transacciones entre procesos de muy alta velocidad. Entre estas tecnologías están el diseño de microarquitecturas del procesador, diseño de sistemas de supercomputación, compiladores automáticos, librerías y lenguajes de programación paralela, programación de propósito general de procesadores gráficos (GPGPU), o redes y protocolos de comunicación para sistemas multinúcleo o clusters.

El procesamiento paralelo está convirtiéndose en algo habitual gracias sobre todo a la presencia en prácticamente la totalidad de los ordenadores personales y dispositivos móviles de procesadores multinúcleo y tarjetas gráficas dotadas de GPUs (*Graphics Processing Unit*), que a su vez son pequeños multiprocesadores masivamente paralelos. Por esta razón esta área de trabajo ha dejado de ser exclusivamente objeto de estudio en los laboratorios de investigación para pasar a ser considerada de sumo interés para el sector industrial de las TIC.

Las GPUs han aparecido en los últimos años como una plataforma potente para la computación de altas prestaciones, y no solamente para aplicaciones en gráficos por computador sino también para aplicaciones de propósito general. El diseño

hardware especializado de las GPUs modernas, optimizado para la ejecución de tareas gráficas, puede funcionar mucho más rápidamente que una CPU normal. Así, las GPUs están especialmente optimizadas para el procesamiento paralelo de una función o kernel sobre un dominio de elementos, siendo capaces de ejecutar eficientemente un número muy elevado de tareas de forma transparente. Las GPUs presentan un gran número de elementos de computación en comparación con las CPUs, permitiendo la asignación de múltiples tareas a cada núcleo a fin de explotar su gran poder computacional.

Sin embargo, desde el punto de vista de la programación, las CPUs tienen muchas ventajas sobre las GPUs, ya que pueden programarse mediante lenguajes estándar ampliamente conocidos como C++ o Java, y disponen de herramientas muy potentes para el desarrollo y depuración de software así como de APIs de programación paralela ampliamente extendidas como OpenMP. Los primeros lenguajes generales que aparecieron para GPUs, por ejemplo, Cg o Microsoft HLSL, eran APIs orientadas a gráficos que requerían una codificación de muy bajo nivel. El interés reciente en la utilización de las GPUs para realizar computación de propósito general ha fomentado la mejora de su programabilidad y la aparición de diversos lenguajes de programación tales como CUDA de Nvidia o Brook+ de ATI, los cuales con frecuencia exponen las características o las limitaciones del hardware. Esto restringe la flexibilidad de los programas para GPUs y obliga al programador a tener ciertos conocimientos sobre el hardware a fin de explotar eficientemente el rendimiento de las GPUs. Esfuerzos recientes para estandarizar la programación de las GPUs modernas han llevado a la creación de DirectCompute y OpenCL, un lenguaje de programación estándar para computación heterogénea. Dado que los lenguajes de alto nivel para GPUs son muy recientes, actualmente hay una escasez de herramientas de programación y librerías especializadas. Esto representa un problema para el desarrollo de nuevas aplicaciones porque obliga a escribir en primer lugar todos los módulos auxiliares necesarios desde cero. Lo que es más, no es posible portar las librerías existentes directamente a la GPU porque su arquitectura normalmente requiere algoritmos especiales o modificados. Además, aunque la utilización de nuevos entornos de programación facilita la programabilidad de las GPUs, una implementación directa de los algoritmos usando estos lenguajes no siempre produce el rendimiento esperado.

Estas últimas razones indican que la programación para GPUs aún no ha alcanzado el suficiente grado de madurez como para convertirse en una opción mayoritaria a la hora de abordar el diseño de aplicaciones o algoritmos paralelos. Además, su gran dependencia al hardware dificulta la implementación de aplicaciones multiplataforma, lo que limita enormemente la posterior difusión de las herramientas desarrolladas. Es por estas razones por las que en esta tesis no se ha optado todavía por la utilización de GPUs.

Como ya se ha comentado, los procesadores multinúcleo han penetrado visiblemente en el mercado de los ordenadores personales, servidores y sistemas empujados, no viéndose su uso solamente limitado al tradicional mercado de la supercomputación. Así pues, el uso del paralelismo se convertirá en el principal método para mejorar el rendimiento de las aplicaciones, satisfaciendo de esta manera la siempre creciente demanda de recursos computacionales con objeto de soportar nuevas funcionalidades y analizar conjuntos de datos cada vez más grandes. Desde el punto de vista del desarrollo de software, la tendencia a aumentar el número de núcleos de procesamiento como fórmula para aumentar la potencia de los computadores tiene un gran impacto en el trabajo y en la productividad de los programadores.

Los proyectos desarrollados y descritos en esta tesis están pensados para aprovechar al máximo las capacidades de los procesadores multinúcleo y los clusters. Como los problemas planteados son de carácter espacial, existe una gran cantidad de datos geográficos, lo que implica que las operaciones efectuadas sobre ellos sean complejas y muy costosas computacionalmente. La unión de geometrías, su intersección, diferencia y demás cálculos consumen muchos recursos y mucho tiempo, por lo que ejecutarlas en ordenadores mononúcleo puede resultar inviable. En algunos casos se ha tenido que recurrir a la paralelización de los algoritmos para reducir su tiempo de ejecución y obtener así herramientas utilizables por los usuarios. Dependiendo de si la máquina donde se ejecuten los algoritmos utiliza memoria compartida (sistemas multinúcleo) o memoria distribuida (sistemas cluster), habrá que utilizar diferentes estrategias de paralelización, tal y como se explicará con detenimiento en el Capítulo 7 de esta tesis.

En este punto es conveniente destacar la importancia que tiene que el lenguaje de programación aproveche bien las capacidades del multiprocesador. En los proyectos desarrollados que necesitan de programación paralela se ha utilizado Java, ya que

está siendo una alternativa real frente a lenguajes como C o Fortran gracias a la aparición de numerosas utilidades para el desarrollo de aplicaciones en sistemas clusters, particularmente en lo que respecta a las redes de comunicación, al soporte multi-núcleo y a la mejora continua de la JVM. Además, desde hace un tiempo a esta parte, han aparecido numerosas librerías que facilitan, y a la vez promueven, el desarrollo de este tipo de aplicaciones en Java [125].

Por poner un ejemplo concreto, para los algoritmos desarrollados se ha hecho uso de una librería llamada MPJ Express [116], la cual facilita la implementación de algoritmos paralelos que utilicen el paradigma de paso de mensajes en entornos cluster. Una alternativa a MPJ Express sería la librería Fast-MPJ [126] pero en el momento en el que se desarrolló este trabajo aún no había alcanzado el suficiente grado de madurez.

2.5. Resumen

En este capítulo hemos comenzado comentando algunos de los problemas actuales del medio rural y de cómo la aplicación de las TIC, y concretamente de los SIG, a los procesos de la Administración de Tierras, pueden ayudar a solventar o por lo menos paliar, algunos de ellos. También se ha destacado la importancia de disponer de información de calidad y de que esta esté accesible de una manera estándar para que cualquier programa de cualquier organización pueda acceder a ella sin importar cómo se guarde o se gestione internamente. Iniciativas como INSPIRE están ayudando a que los organismos públicos publiquen la información georreferenciada de la que disponen para que pueda ser utilizada por parte de los ciudadanos.

Para publicar u obtener dichos datos podemos hacer uso de las IDE y de los servicios OGC. Para gestionarlos, la utilización de software SIG libre es una buena opción ya que aporta calidad, robustez y madurez todo ello con un bajo coste.

Ya en la última parte, se han introducido las tecnologías de computación de altas prestaciones. Su utilización se hace necesaria cuando se trabaja con muchos datos geográficos, ya que las operaciones espaciales suelen ser muy costosas computacionalmente.

Parte II

Sistemas de información para el desarrollo rural

Capítulo 3

Sistema de información Web-SIG para la Administración de Tierras

En el capítulo anterior se han comentado los beneficios de utilizar las TIC, y concretamente los SIG, en la gestión de recursos que tienen que ver con la Administración de Tierras. En este capítulo se lleva a la práctica lo comentado hasta ahora y se define e implementa una arquitectura que sirve de soporte para la construcción de sistemas Web-SIG aplicables al tratamiento y resolución de algunos de los problemas de las regiones rurales. En este caso, sobre este sistema base se han desarrollado dos sistemas de información para la dinamización de los mercados de tierras, SITEGAL, y para la gestión común de parcelas forestales, SIUXFOR.

La arquitectura propuesta ha sido desarrollada utilizando exclusivamente software libre. Esto, a parte de reducir los costes de desarrollo de una manera muy importante, es posible debido al avance que las herramientas SIG libres han tenido en cuanto a funcionalidades, fiabilidad, eficiencia y soporte [66]. Otra ventaja es la disponibilidad del código fuente de estas herramientas, con lo que se podrían realizar modificaciones que se adapten a necesidades específicas del proyecto a desarrollar.

Para facilitar la comunicación con servicios externos de entidades que proveen capas de información, se hace uso de los servicios estándar del OGC. De esta manera, se pueden cambiar las fuentes de datos utilizadas de manera directa, simplemente cambiando la URL a la que se tiene que acceder.

El resto del capítulo se estructura de la siguiente manera: la Sección 3.1 define detalladamente la arquitectura del sistema de información propuesto y las tecnologías utilizadas. Las secciones 3.2 y 3.3 enumeran los componentes que forman parte del sistema y las diferentes fuentes de información a las que se accede, respectivamente. Finalmente, la Sección 3.4 estudia en profundidad algunas de las principales funcionalidades SIG del sistema.

3.1. Arquitectura del sistema

La Figura 3.1 muestra la arquitectura del sistema. El núcleo se basa en una aplicación web Java que implementa toda la lógica de negocio, el almacenamiento de la información en bases de datos y la generación de páginas web de manera dinámica. La aplicación Java también es la encargada de acceder a los datos geográficos. Estos son gestionados por un servidor geográfico, un proceso independiente que actúa como proxy entre la información geográfica y los usuarios. Este servidor geográfico también es el responsable de la generación de los mapas y de la obtención de información de los distintos objetos geográficos existentes (como ríos, carreteras, vías de tren, etc.) utilizando para ello fuentes de datos locales y/o externas. Utilizando esta arquitectura, la aplicación web y las fuentes de datos son completamente independientes.

El visor geográfico es el encargado de interceptar las acciones del usuario y transformarlas en peticiones que serán enviadas directamente al servidor geográfico utilizando AJAX (Asynchronous Javascript And XML). El servidor de mapas recibe estas peticiones, genera la respuesta y la devuelve al visor geográfico que recargará sólo las partes de la página web necesarias para mostrar la información recibida.

La información que maneja el sistema, que se guarda en dos bases de datos con capacidades espaciales, se puede clasificar en dos grupos: la información relacionada con los usuarios y las parcelas¹ (incluida su geometría), la cual es gestionada por la aplicación Java; y la información relacionada con los objetos geográficos (edificacio-

¹Nos referimos a las parcelas por este el objeto geográfico más común que se utilizará en los sistemas que se basen de esta arquitectura. No obstante, en vez de parcelas se puede utilizar cualquier otro objeto. Normalmente este va a tener una componente espacial y pueden ser desde edificaciones hasta puntos en un mapa, ríos, carreteras, etc.

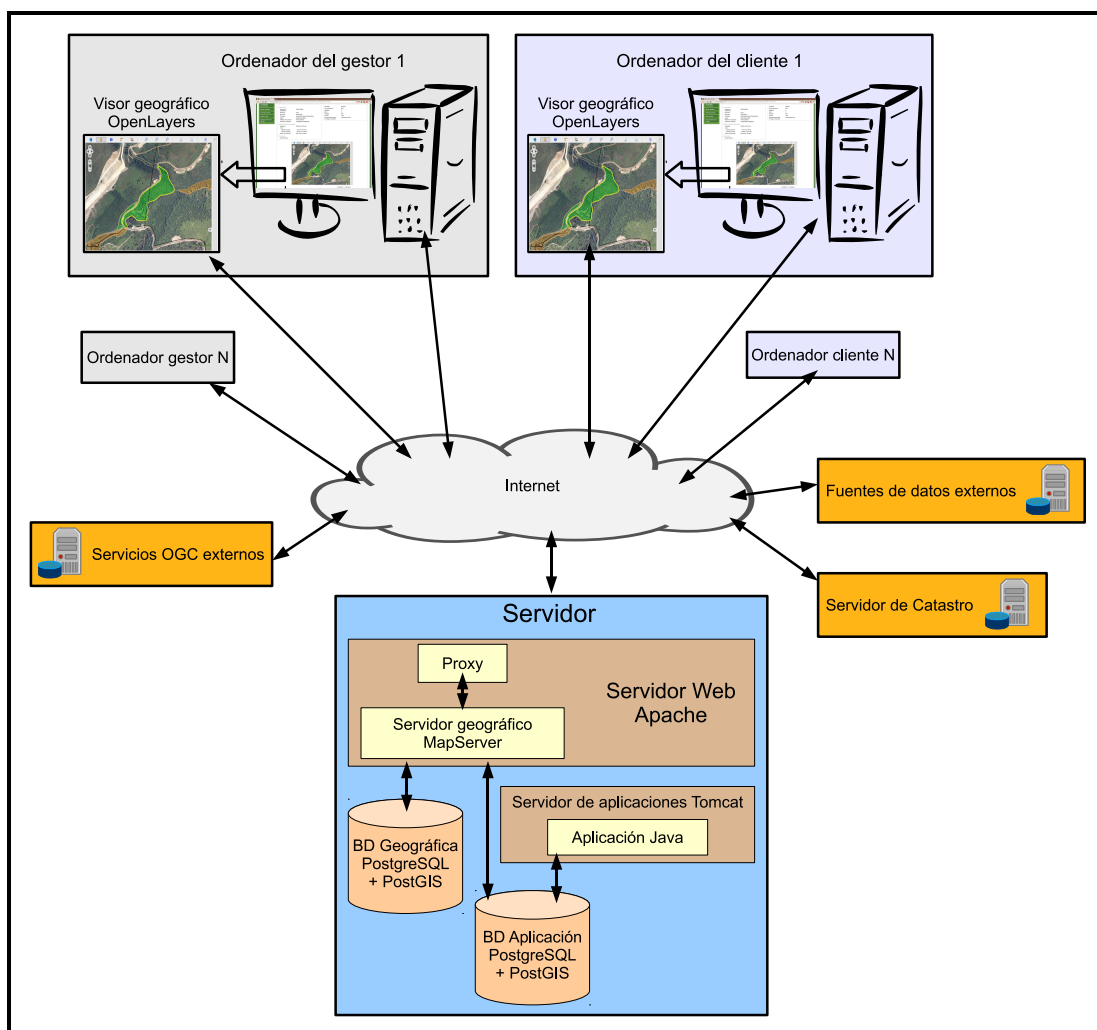


Figura 3.1: Arquitectura del sistema.

nes, redes de abastecimiento, todas las parcelas de una zona, etc.), que es gestionada por el servidor geográfico. Por comodidad, de aquí en adelante nos vamos a referir a la primera base de datos como *base de datos de la aplicación* y a la segunda como *base de datos geográfica*. Hay que destacar también que se utilizan otras fuentes de información geográfica accedidas mediante los estándares WMS y WFS, tal y como se explica en las siguientes secciones.

3.2. Componentes del sistema

En lo siguientes párrafos se realiza una breve descripción de los distintos componentes que forman el sistema de información implementado, así como las librerías y herramientas, todas ellas libres, que se han utilizado y de cómo se integran en un único sistema.

3.2.1. Aplicación web Java

La aplicación web ha sido desarrollada con la tecnología J2EE haciendo uso de librerías de terceros (resumidas en la Tabla 3.1) y está basada en patrones de diseño ampliamente conocidos [76] mostrados en la Figura 3.2. En cuanto a la arquitectura, se ha utilizado el patrón Mode-View-Controller (MVC) que nos permite mantener separado el núcleo de la aplicación de lo que es la vista.

3.2.2. Visor geográfico

El visor geográfico está basado en OpenLayers, un visor web que actualmente es de los más populares. Hecho en JavaScript, utiliza AJAX para la comunicación con las distintas fuentes de datos y, entre otras cosas, permite: acceso a datos mediante servicios estándares (WMS, WFS, WMTS, KML, GML, etc.) o no estándares (Google Maps, Yahoo Maps, OpenStreetMaps, Bing Maps, etc.); navegación sobre los mapas, zooms, medición de longitudes y áreas; petición de información sobre objetos visibles; selección y centrado de objetos, *permalinks* (enlaces para recuperar el estado del visor); digitalización y modificación de geometrías; etc.

Nombre	Capa	Descripción
Hibernate	Modelo	Herramienta Java de mapeo objeto-relacional utilizada en la comunicación con la base de datos. Implementación del patrón Data Access Object (Figura 3.2).
Spring	Modelo	Framework Java para el desarrollo de aplicaciones. Implementa gestión de transacciones con base de datos, inversión de control, autenticación y autorización con Spring Security, etc.
Apache Struts	Vista/Controlador	Framework Java utilizado como implementación del patrón Front Controller (Figura 3.2).
Apache Wicket	Vista/Controlador	Framework Java basado en componentes. Sustituye a Struts en las funcionalidades más recientes de la aplicación
JavaServer Pages (JSP)	View	Tecnología Java para el desarrollo de páginas web dinámicas. Implementación del patrón View Helper (Figura 3.2). Utilizada junto a Struts.
AjaxTags	Vista	Librería JSP utilizada para la implementación de funcionalidades AJAX. Utilizada junto a Struts.
DisplayTag	Vista	Librería JSP utilizada para la creación de tablas HTML dinámicas, utilizando para ello el patrón Value List Handler (Figura 3.2). Utilizada junto a Struts.
JCaptcha	Vista	Implementación de mecanismos CAPTCHA.
JODReports	Controlador	Generación de documentos en varios formatos basándose en plantillas ODT (formato de OpenOffice).
Quartz Scheduler	Modelo	Librería Java utilizada para la planificación de tareas asíncronas.
ClamAV	Model	Motor anti-virus que puede ser utilizada en aplicaciones Java a través de su API.
Java Topology Suite (JTS)	Modelo	Librería de operaciones espaciales bi-dimensionales: intersección, unión, diferencia, buffers, etc.
GeoTools	Modelo	Librería con varias funcionalidades para el manejo de datos espaciales.

Tabla 3.1: Herramientas utilizadas en el desarrollo de la aplicación web Java.

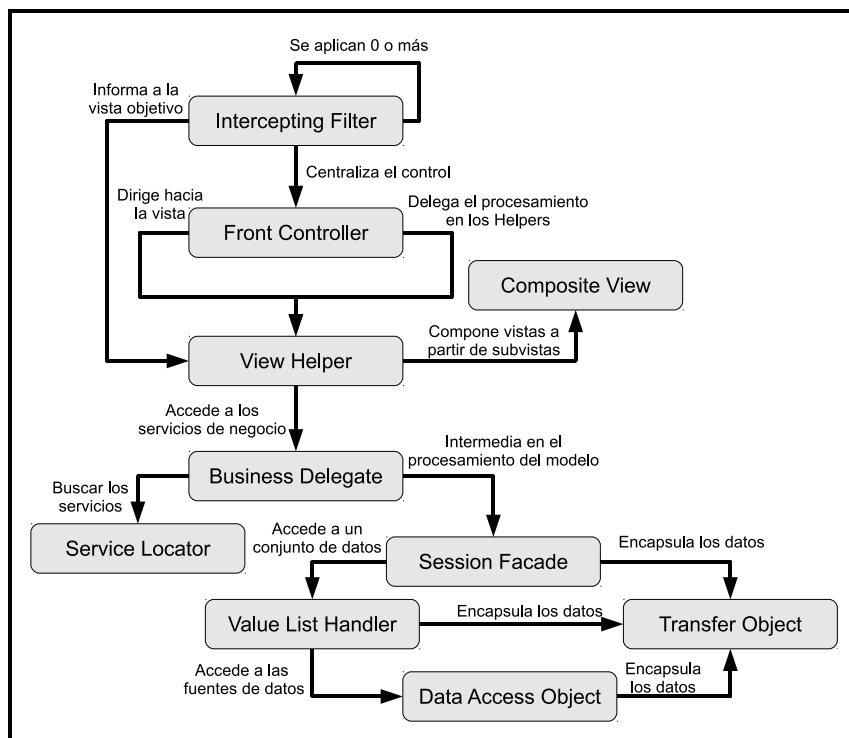


Figura 3.2: Patrones de diseño utilizados en la aplicación Java.

3.2.3. Servidor geográfico

Como servidor geográfico se utiliza MapServer, que es una de las herramientas SIG más veteranas, más populares y con más soporte. Está desarrollado en C y trabaja como CGI en un servidor web, aunque también ofrece APIs para poder utilizar otros lenguajes. Su rendimiento es bueno y permite el acceso a otras fuentes de datos mediante servicios estándares. En particular, permite el acceso en cascada a otros servidores WMS y WFS.

3.2.4. Almacenamiento

Como sistema gestor de base de datos se utiliza PostgreSQL con su extensión espacial PostGIS. PostGIS tiene soporte para los estándares de la OGC, como por ejemplo el SFS ². Tiene una buena indexación espacial y es soportado por casi todos

²Simple Feature Specification for SQL. URL: www.opengeospatial.org/standards/sfs

los servidores geográficos.

3.2.5. Integración de componentes

La conexión entre la aplicación Java y la base de datos de la aplicación es hecha a través de Hibernate e Hibernate Spatial, que soporta tanto PostgreSQL como PostGIS. Cualquier otra fuente de datos soportada por Hibernate también puede ser utilizada con sólo cambiar ficheros de configuración, ya que la aplicación web es transparente a cómo y a dónde se accede a la información. Por su parte, el sistema gestor de base de datos geográfica también puede ser cambiado ya que MapServer soporta varias bases de datos espaciales encargándose de esta conexión.

La comunicación del servidor geográfico con el visor y la aplicación Java es la parte que ha requerido más trabajo. Aunque la visualización de capas en el visor se hace de forma sencilla, hay otros casos en los que se requiere una comunicación más complicada y que hay que hacer en varios pasos. Por ejemplo, para la inserción de una parcela en el sistema, el usuario debe realizar lo siguiente:

- Localizar en el mapa la zona donde está la parcela (petición *GetMap* de WMS para mostrar el mapa).
- Hacer click sobre la parcela para que esta sea resaltada mostrando sus bordes en otro color (petición *GetFeatureInfo* de WMS para identificar la parcela y petición *GetMap* de WMS para obtener el mapa con la parcela resaltada).
- Confirmar que la parcela resaltada es la que se quiere introducir en el sistema y si lo es, guardarla en base de datos (petición *GetFeature* de WFS para obtener su geometría en formato vectorial y sus atributos alfa-numéricos).

Con esta estructura, los distintos componentes que forman el sistema son independientes entre sí y cualquier cambio sobre uno de ellos no afecta al resto. Por ejemplo, el servidor geográfico puede ser reemplazado por otro servidor que siga los estándares y esto será completamente transparente al resto del sistema.

3.3. Fuentes de datos geográficos

Se necesitan al menos tres capas básicas de información geográfica para que se puedan realizar las acciones más comunes en los procesos llevados a cabo en la Administración de Tierras sobre un visor de una manera eficaz: ortofotografías del terreno, una capa de usos del suelo, y un mapa catastral. No obstante, añadir otras fuentes de datos al visor es muy simple gracias a la estructura modular del sistema y a la utilización de estándares.

A menudo el coste de tener y mantener actualizada toda la información geográfica es demasiado alto como para siquiera contemplarlo. Es útil por tanto que el sistema sea capaz de conectarse a servicios de terceros y extraer de ellos la información necesaria. Además, el sistema aquí descrito sigue las recomendaciones marcadas por la iniciativa INSPIRE y todos los servicios utilizados y proporcionados son estándares.

3.4. Detalles de implementación de las funcionalidades SIG

En esta sección se comentan las funcionalidades SIG más interesantes que han sido implementadas en el sistema, dando a su vez una explicación de cómo se han resuelto.

3.4.1. Control de acceso a la información geográfica

No todos los usuarios pueden consultar toda la información disponible. Los permisos de los usuarios están basados en roles: cada usuario tendrá asignado un rol que le permitirá o no el acceso a ciertas zonas de la aplicación, a determinada información y determinadas acciones. El sistema por tanto dispone de dos tipos de acceso a los datos: acceso a toda una capa de información o acceso a sólo determinados objetos (o sólo a algunas características de objetos) de la capa de información.

El control de acceso a la base de datos de la aplicación está integrada con el

mecanismo de autenticación de la aplicación web pero, para ser capaces de utilizar el control de acceso independiente de las fuentes de datos, así como de servicios externos del OGC o incluso del servidor geográfico, se utiliza un filtro SLD acoplado a las peticiones. El encargado de acoplar este filtro es un proxy escrito en Python y ejecutado como CGI en el servidor web. Como se puede observar en la Figura 3.3, el visor dirige todas las peticiones hacia dicho proxy que comprueba los permisos del usuario: si este tiene permiso para consultar la información pedida, el proxy reenvía la petición (con el filtro ya acoplado) al servidor geográfico que genera y devuelve la respuesta.

3.4.2. Localización de parcelas y obtención de sus datos mediante protocolos estándares

Cuando un usuario realiza una consulta sobre una parcela y la selecciona en el visor geográfico, el sistema tiene que pedir los datos de la parcela al servidor geográfico en base a su localización y geometría. La Figura 3.4 muestra la secuencia de acciones para obtener los usos de una parcela. Primero la aplicación web realiza una petición *GetFeature* de WFS al servidor geográfico especificando el *Bounding Box* (BB) de la parcela, esto es, el rectángulo con menor área posible para que todos los puntos de la parcela queden dentro. Por su parte, el servidor geográfico devuelve una lista de elementos (o *features*) que estén dentro de ese BB. Estos elementos tienen información sobre los usos del suelo y sus geometrías, así la aplicación Java descarta aquellos elementos que no intersequen con la parcela originalmente seleccionada. En este diagrama, así como en los sucesivos, el proxy comentado en el apartado anterior será ignorado con el fin de simplificar las figuras y hacerlas más comprensibles.

De un modo similar, para localizar una parcela en el visor geográfico y conseguir toda la información sobre ella, también se utiliza su BB en una petición *GetFeature* de WFS. Esta petición devuelve una lista de elementos con información sobre las parcelas dentro del BB, y el sistema es el encargado de escoger el elemento con más área de intersección con la parcela seleccionada por el usuario.

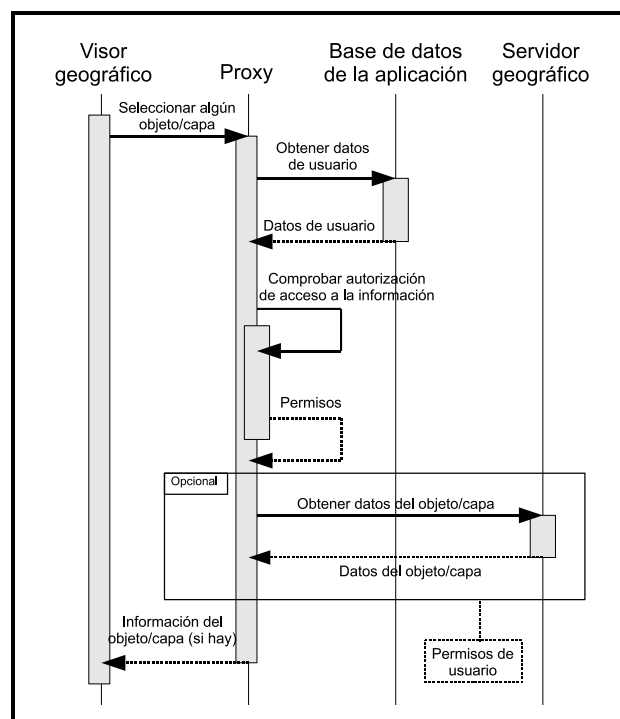


Figura 3.3: Diagrama de secuencia para el control de acceso a las capas de información.

3.4.3. Modificación de geometrías

Como la capa de parcelas puede contener errores, el sistema permite la modificación de las geometrías de las parcelas. Esto se puede realizar de tres maneras diferentes: a través del servidor WFS del Catastro, mediante archivos de formato espacial, o directamente modificando de forma visual la geometría de la parcela mediante el visor geográfico.

Esta funcionalidad es muy interesante ya que el sistema puede mantener una comunicación con la fuente de datos que proporciona la capa de parcelas y corregir posibles errores en sus datos. Señalar que si la modificación es realizada por un usuario registrado, un gestor debe revisarlo para asegurar la veracidad de dichos datos.

A continuación se explican tres maneras con las que un usuario puede visualizar geometrías de servicios externos y corregir las que tiene el sistema.

Integración con el servicio WFS del Catastro

La Figura 3.4 (B) muestra el diagrama de secuencia entre los distintos componentes que forman parte en la comunicación entre el sistema y un servicio WFS, en este caso el del Catastro. Al inicio del procedimiento el usuario debe introducir el código catastral de la parcela a la que quiere modificar su geometría en la aplicación web que es la encargada de la conexión con el servicio WFS. Como la información requerida es privada, el usuario inserta su nombre de usuario y contraseña que previamente ha de haber conseguido en las oficinas de Catastro. A continuación, el servicio WFS devuelve el elemento que coincida con los códigos introducidos y la aplicación web guardará la geometría del elemento devuelto reemplazando la existente. Después, la aplicación creará dinámicamente una página con el visor geográfico centrado en la nueva geometría para que el usuario pueda visualizarla.

Modificación de geometrías mediante archivos espaciales

Los usuario también pueden optar por la opción de subir a la aplicación archivos de formato espacial, opción válida para cuando es el propio usuario quien tiene la geometría correcta. La aplicación, que acepta los formatos de archivo Shape y GML, leerá la nueva geometría con GeoTools y, después de comprobar que es válida, la asignará a la parcela en cuestión. Además, la aplicación también es capaz de leer geometrías en formato WKT (Well-Known Text), un formato de marcado regulado por la OGC utilizado para representar objetos geométricos. En este caso, la aplicación utiliza la herramienta JTS para leer la geometría.

Modificación de geometrías sobre el visor geográfico

OpenLayers admite dibujar geometrías en el propio visor geográfico. Esta funcionalidad es aprovechada para que el usuario, basándose en la geometría original de la parcela, pueda mover sus vértices corrigiéndola. Como muestra la Figura 3.5, para poder llevar a cabo este proceso en el visor se activan unos controles de edición. Para mostrar la geometría, la aplicación Java la recupera de la fuente de datos oportuna y la traduce a WKT, formato soportado por OpenLayers. El usuario puede entonces modificar los vértices de la geometría mostrada modificándola hasta que coincida

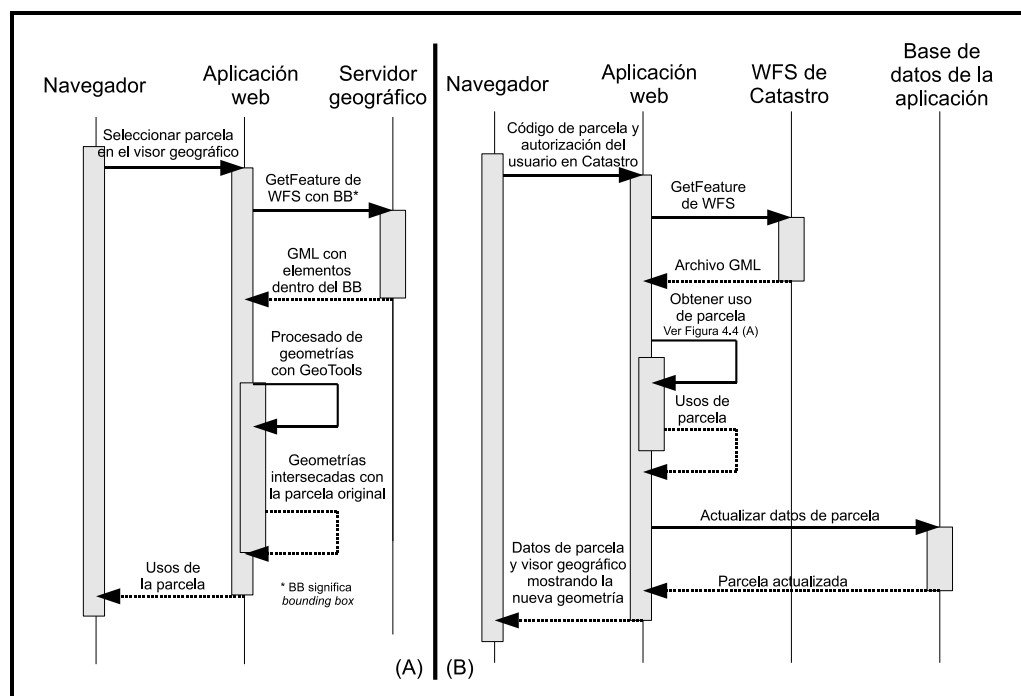


Figura 3.4: Diagrama de secuencia de las comunicaciones WFS con el servidor local (A) y con el servidor de Catastro (B), para obtener la capa de usos del suelo

con la geometría real. Cuando acabe, la aplicación Java realizará el proceso contrario al anterior y traducirá la nueva geometría de WKT al objeto *Geometry*, que es con el que trabaja la librería JTS. Es entonces cuando la nueva geometría es guardada en base de datos y asignada a la parcela correspondiente.

3.4.4. Centrado y resaltado de parcelas

Teniendo siempre en cuenta la utilización de estándares, el centrado y resaltado de parcelas es realizado mediante una combinación de SLD (para el resaltado) y filtros (para la identificación de la parcela) en varios pasos. También es necesaria la utilización de un BB para ajustar el visor a las dimensiones de la parcela. La Figura 3.6 muestra la secuencia de acciones y comunicaciones que se han de realizar cuando un usuario proporciona un código catastral. Primero el visor lanza una petición *GetFeature* de WFS con un filtro especificando el código de la parcela. A continuación el servidor geográfico consulta la base de datos geográfica, consigue la información sobre dicha parcela y construye con ellos un archivo GML que envía al visor. Es

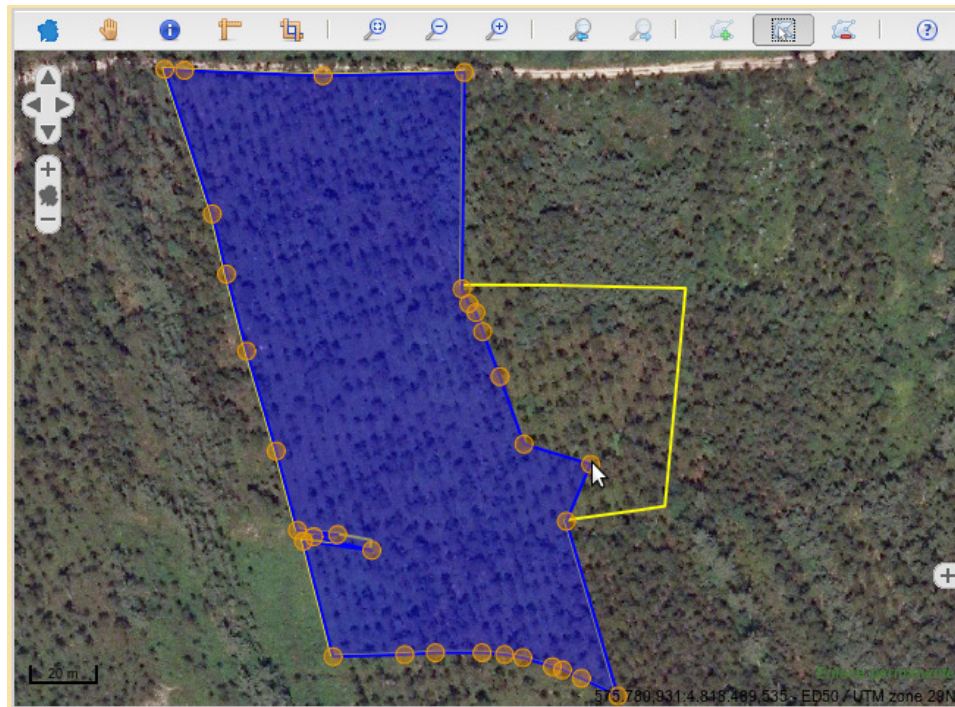


Figura 3.5: Visor geográfico en modo edición con la geometría original (línea amarilla).

entonces cuando el visor geográfico puede extraer todos los datos necesarios del GML y construir en base a ellos una petición *GetMap* de WMS con el correspondiente BB que envía al servidor geográfico. De esta manera el servidor responde con una imagen en la que la parcela aparecerá centrada y con sus bordes resaltados.

Lo hasta ahora explicado ocurre cuando el usuario inserta los códigos de la parcela en la aplicación web, pero cuando el usuario selecciona la parcela mediante un click en el visor, el proceso cambia ligeramente. La primera petición no es *GetFeature* de WFS ya que el visor no tendrá la identificación de la parcela por desconocer sus códigos, y dado que sólo conocen las coordenadas donde el usuario hizo click, la petición que genera es una *GetFeatureInfo* de WMS. El servidor recibe dicha petición y genera un GML con los atributos de la parcela seleccionada. Es entonces cuando el visor ya tiene identificada la parcela y puede centrarla y resaltarla como se ha explicado previamente.

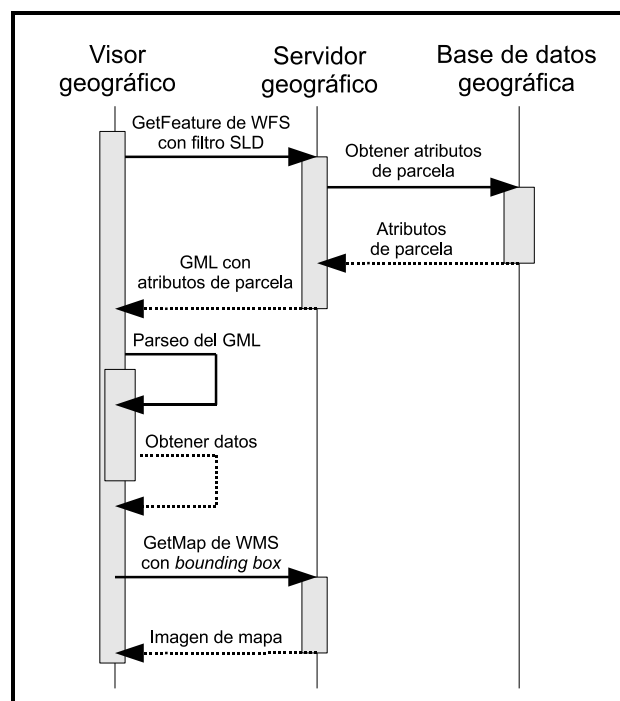


Figura 3.6: Diagrama de secuencias para el centrado y resaltado de una parcela identificada mediante sus códigos catastrales.

3.4.5. Agrupamiento de capas geográficas

El sistema permite la consulta de un gran número de capas de información obtenida de diferentes fuentes. Con el objetivo de reducir las peticiones enviadas al servidor y por ende el número de imágenes devueltas por él, las capas de información son agrupadas en grupos lógicos para que puedan ser consultadas a la vez en la misma petición WMS. Otra forma de conseguir este agrupamiento es utilizando WMS en cascada, que consiste en que el servidor geográfico consigue la información de otras fuentes pero de manera transparente al cliente original, ya que él solo realizará una petición a *su* servidor geográfico. De esta manera se pueden obtener múltiples imágenes de otros tantos servicios externos, como muestra la Figura 3.7, reduciendo así el tráfico de datos.

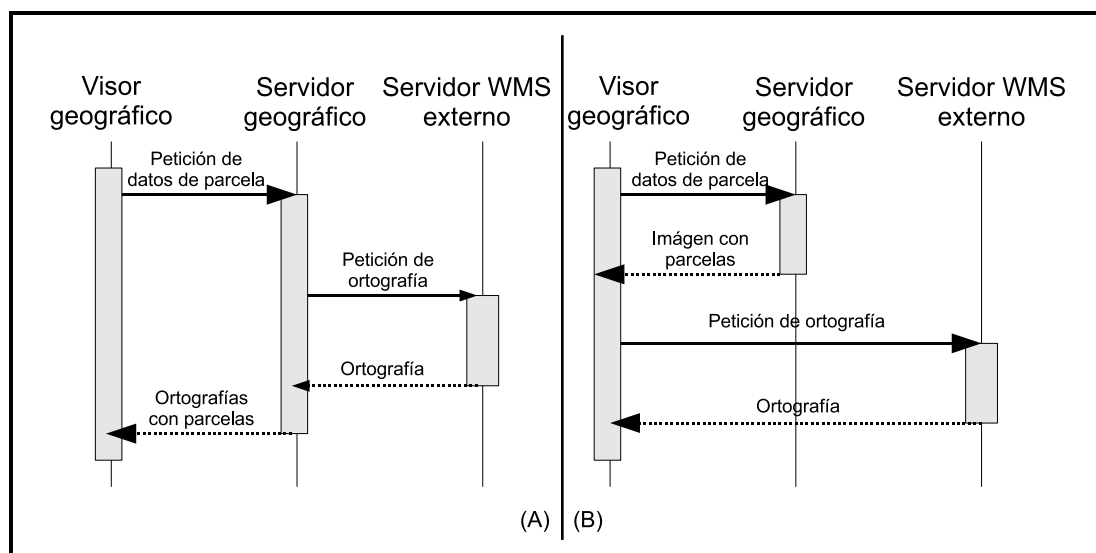


Figura 3.7: Diagramas de secuencia comparando la utilización del WMS en cascada (A) y no utilizándolo (B).

3.5. Resumen

En este capítulo se ha descrito la arquitectura de un sistema de información basado en tecnologías web y SIG cuya función es servir de base para la implementación de herramientas más específicas que traten de resolver problemas concretos y que se basen en elementos geográficos, como por ejemplo, las parcelas.

El sistema, basado en roles, permite la gestión de la información tanto a los empleados de la Administración de Tierras como a los ciudadanos, con el objetivo de promover el e-Government. Dado la naturaleza espacial de la información con la que se trata, el sistema tiene amplias capacidades SIG que permiten consultar y visualizar datos de forma eficiente.

También se ha descrito cómo se han resuelto funcionalidades complejas buscando soluciones siempre soluciones basadas en la utilización de estándares y siguiendo las recomendaciones de iniciativas como la INSPIRE.

El diseño e implementación de estos sistemas se han desarrollado íntegramente con software libre. Las herramientas SIG de código abierto han demostrado en general que son una alternativa real debido a su alto rendimiento. El futuro en este campo de estas herramientas es prometedor ya que tienen el apoyo de importantes com-

pañías, instituciones y organizaciones como la Open Source Geospatial Foundation (OSGeo) o el Open Geospatial Consortium (OGC). Además, la utilización de estándares hace más fácil la utilización de fuentes externas de información geográfica, lo que libera a la organización que utilice el sistema de la obtención y mantenimiento de los datos.

Capítulo 4

SITEGAL: sistema de información Web-SIG para la gestión del Banco de Tierras de Galicia

Este capítulo explica cómo, basándose en la arquitectura base explicada en el capítulo anterior, se ha construido una aplicación Web innovadora con capacidades SIG completas, cuyo principal objetivo es la dinamización del mercado de tierras, en este caso gallego, facilitando la comunicación entre propietarios de tierras y agricultores interesados en aumentar su capacidad productiva. Dicha herramienta, llamada SITEGAL (www.bantegal.com/sitegal), es la pieza clave en el funcionamiento del Banco de Tierras de Galicia, y es utilizada tanto por los gestores del Banco como por los clientes del mismo (propietarios, agricultores y público en general). SITEGAL lidera el progreso hacia el e-Government de la Administración de Tierras en Galicia.

En la comunicación entre propietario de tierras y agricultor existe un intermediario que generalmente es, por lo apropiado, un Banco de Tierras. De esta manera los propietarios podrán ceder las tierras al Banco y este se encargará de publicitarlas y buscarles un agricultor que quiera arrendar dichas tierras para expandir su negocio¹.

Este capítulo se estructura de la siguiente manera: la Sección 4.1 explica el por

¹El sistema se explicará en base a procesos de cesión/arrendamiento y no de compra/venta, ya que los Bancos de Tierras utilizan, por norma general y aún habiendo excepciones, estos procedimientos.

qué el mercado de tierras en Galicia se ha estancado en los últimos años. La Sección 4.2 comenta cómo nació el Banco de Tierras de Galicia y por ende, SITEGAL. Por su parte, la Sección 4.3 resume el funcionamiento de la herramienta y las acciones que pueden realizar los distintos tipos de usuarios que tiene, y la Sección 4.4 expone los resultados obtenidos en las pruebas de evaluación que han sido aplicadas a la herramienta. Por su parte, la Sección 4.5 explica brevemente SIUXFOR, otra adaptación de la arquitectura base pero en este caso con el objetivo de dar soporte a las Unidades de Gestión Forestal. SIUXFOR, incluido en el proyecto *Sistema de Información para la tramitación de las Unidades de Gestión Forestal (UXFOR)*, intenta solucionar otro problema del medio rural gallego: la ineficiencia de la gestión común de parcelas, en este caso, forestales. Tanto este proyecto como el de SITEGAL han sido financiados por la Xunta de Galicia a través de contratos de I+D.

4.1. Impedimentos para la movilidad de tierras en Galicia

A lo largo de esta tesis se ha comentado cuales son los principales problemas que tienen actualmente las áreas rurales en general, y concretamente las gallegas. A modo de recordatorio comentar que se ha identificado que uno de los principales impedimentos para la reactivación económica del sector agrario gallego es el alto grado de abandono de las tierras y el estancamiento de la actividad en los mercados de tierras. En una reciente encuesta llevada a cabo por el Banco de Tierras de Galicia entre los técnicos de las Oficinas Agrarias Comarcales², se han valorado de 1 a 5 (significando 1 muy poca repercusión y 5 mucha repercusión) algunas de las posibles causas por las cuales estos técnicos creen que la movilidad de tierras se ve perjudicada. En los resultados de la encuesta expuestos en el Apéndice B se puede observar cómo las causas más valoradas tienen que ver con las características del parcelario gallego (parcelas pequeñas y de difícil acceso) y con la desconfianza que tienen los propietarios a la hora de ceder sus tierras. La Tabla 4.1 resume los principales obstáculos para el arrendamiento de tierras desde dos puntos de vista: la parcela y la otra parte.

²Sucursales locales que la Consellería do Medio Rural e do Mar de la Xunta de Galicia tiene distribuidas por las zonas rurales gallegas.

Parte	Relativo a la otra parte	Relativo a la parcela
Agricultor	¿Quién es el propietario?	¿Cuáles son sus límites?
	¿Cómo puedo contactar con él/ella?	¿Hay alguna parcela disponible en esta zona?
		¿Cuál su tamaño, su actual uso del suelo, su orientación, su pendiente...?
		¿Cómo son los alrededores: accesos, carreteras, distancia desde mi terreno?
		¿Cuál es el precio?
Propietario	¿Cómo puedo contactar con él/ella?	¿Cuál será un buen precio?
	¿Hay alguien interesado en mi parcela?	¿Dónde está la parcela y cuáles son sus límites?
Transacción	¿Cuáles debería ser las cláusulas y las condiciones del contrato?	

Tabla 4.1: Principales obstáculos en el arrendamiento de parcelas.

4.2. El Banco de Tierras de Galicia y SITEGAL

Aunque los bancos de tierras ya están funcionando en varias regiones del mundo, no se ha encontrado ninguna herramienta con capacidades SIG que gestionen dichos bancos. En el caso de España, el Banco de Tierras de Asturias dispone de una herramienta Web-SIG aunque sólo es utilizada para la gestión interna. Por su parte, Bizcaia también ha creado recientemente un Fondo de Suelo Agrario aunque todavía no dispone de ninguna herramienta pública con las capacidades comentadas. A nivel europeo, diferentes Bancos de Tierras han estado operando en varias regiones especialmente desde la década de los 60 [40, 47]. Una de los ejemplos más significativos es el SAFER (Sociétés d'aménagement foncier et d'établissement rural) francés, que sólo en 2010 estuvo activamente involucrado en la transacción de 76.000 propiedades agrícolas [83]. Basado en la compra-venta, los clientes pueden consultar las ofertas vía web (www.proprietes-rurales.com) pero sin capacidades SIG.

En 2007 la Xunta de Galicia aprueba la *Ley 7/2007, de 21 de mayo, de medidas administrativas y tributarias para la conservación de la superficie agraria útil y del*

Banco de Tierras de Galicia [141] con la que se constituye un instrumento aplicado a la Administración de Tierras, cuyos principales objetivos son reducir el abandono de las tierras y promover la transferencia de tierras entre propietarios y agricultores por medio de arrendamientos. De esta manera, sólo los derechos de utilización de la tierra son cedidos y el agricultor sólo tendrá que centrarse en sacarle provecho. Otro de los objetivos centrales de este instrumento es la reducción de los costes de información asociados a la movilidad de tierras, facilitando para ello el contacto entre propietarios que no utilizan sus tierras y agricultores dispuestos a ampliar sus explotaciones. Así pues, se persiguen tres objetivos principales: dar una alternativa a la compra-venta de tierras, mitigar la desconfianza inicial de los propietarios, y facilitar los flujos de información entre todos los participantes en los procesos de arrendamiento.

Para la gestión del Banco de Tierras existe una oficina central que tiene el apoyo de las Oficinas Agrarias Comarcales que están distribuidas por toda Galicia. Las dos principales operaciones que realiza el Banco son: la incorporación de parcelas al fondo de tierras del Banco, y el arrendamiento de dichas parcelas. La incorporación de una parcela implica que el propietario permite al Banco arrendarla a terceras partes para uso exclusivamente agrícolas. Mientras no se arriende, el propietario podrá seguir trabajándola. Para realizar la incorporación, el propietario debe presentar una solicitud de cesión, que será aceptada por el Banco una vez se haya comprobado que el solicitante tiene capacidad legal para cederla. Además, la parcela debe cumplir algunos requisitos mínimos que tienen que ver con la adecuación de la parcela a la producción agrícola. El precio de arrendamiento es fijado por el Banco en base a unos precios de referencia establecidos anualmente para cada parroquia gallega y dependiendo de los usos del suelo actuales de la parcela, aunque el propietario también puede especificar qué precio le gustaría obtener. El Banco no garantiza el arrendamiento de una parcela, ya que esto depende de la demanda existente, pero si la parcela es arrendada, el propietario recibirá el total de la renta menos un 1 % que se queda el Banco como comisión por la gestión. Lo que también garantiza es que el propietario recibirá la renta de la parcela y que, cuando finalice el contrato, la parcela se le devolverá en el mismo estado en la que la cedió. El tiempo de permanencia de una parcela en el Banco es ilimitado, es el propietario quien debe pedir su retirada si no está ya en alquiler, ya que si lo está, deberá esperar a que finalice el contrato de arrendamiento. A parte de estas parcelas de propiedad privada, el Banco también

dispone en su fondo de tierras parcelas de propiedad pública, que normalmente se corresponden con los terrenos sobrantes en los procesos de concentración parcelaria.

Con lo que respecta al arrendamiento, cualquier persona mayor de 18 años puede pedir el alquiler de una parcela para darle un uso agrícola. El procedimiento de alquiler implica la presentación de una solicitud de arrendamiento por parte del agricultor. Si resultase que hubiera más de una solicitud de arrendamiento sobre una misma parcela, la ley establece una serie de criterios que tendrían que ser evaluados para poder decidir a quién se le concede finalmente el derecho a utilizar la parcela. Una vez la solicitud es aceptada, el agricultor y el Banco deben firmar el contrato de arrendamiento. La duración mínima del contrato son 5 años, pero pueden ser más si ambas partes están de acuerdo.

Para la gestión de todos estos procedimientos se utiliza SITEGAL (Sistema de Información de Tierras de Galicia), una aplicación Web-SIG que está disponible desde noviembre de 2007 en la dirección web www.bantegal.com/sitegal. El sistema, que está en continua evolución a través de actualizaciones periódicas, es utilizado tanto para la gestión interna del banco como para la realización de los trámites de cesión y arrendamiento por parte de propietarios y agricultores respectivamente. La Figura 4.1 muestra una composición de capturas de pantalla de una lista con parcelas disponibles y del visor geográfico sobreimpresionado.

4.3. Sistema de roles y funcionalidades del sistema

Para la gestión de permisos se utiliza un sistema basado en roles. De esta manera los distintos usuarios utilizan la misma aplicación pero las acciones que pueden realizar y los datos que pueden ver son diferentes según su rol. Para facilitar el acceso a la información y promover así el e-Government, la ejecución de algunas acciones no necesitan autenticación en el sistema. Los roles utilizados y los permisos que estos tienen son comentados en los siguientes párrafos.

Localización	Código	Data introdución	Superficie	Usos do solo	Prezo	Dispoñible
Trazo - A Coruña Castelo (Santa María)	SIXPAC (2007) Zona:0 Polígono:503 Parcela:195	15/11/2007	88.635 m ² (8,86 ha)	Prados e pradeiras: 88.635 m ² (8,86 ha)	1.365 euros anuais	Si
Fonsagrada (A) - Lugo Lamas de Campos (San Roque)	SIXPAC (2007) Zona:1 Polígono:5	28/12/2007	69.800 m ² (6,98 ha)	Pasteiro: 26.266 m ² (2,63 ha) Forestal arborado: 43.534 m ² (4,35 ha)	195 euros anuais	Si

541.925,893;4.763.252,387 - ED50 / UTM zone 29N

Figura 4.1: Lista de parcelas y el visor geográfico de SITEGAL.

4.3.1. Usuarios no autenticados

Los usuarios no autenticados pueden consultar las parcelas disponibles en el fondo de tierras del sistema, ya sea en una lista o a través del visor geográfico, donde la posición y la forma de la parcela pueden ser observadas sobre una ortofotografía de alta resolución.

Mediante la lista se pueden ver más datos de la parcela: su localización, superficie, precio de alquiler, y sus usos actuales (forestal, pastoral, huerta, etc.). El usuario también puede realizar búsquedas a través de un formulario especificando algún criterio (superficie, precio, localización, etc.). Si el usuario no encontrase alguna parcela con las características deseadas, puede introducir en el sistema una petición para que le avisen, mediante correo electrónico o SMS, cuando se inserte en el fondo de tierras alguna parcela que coincida con sus necesidades.

4.3.2. Usuarios autenticados

Si un usuario necesita alquilar una parcela o si un propietario quiere añadir alguna de sus posesiones al fondo de tierras del sistema, deben autenticarse. La información requerida para el registro es la mínima para poder realizar los trámites y así evitar que esto sea un impedimento para la utilización de la aplicación.

En el proceso de inserción de una parcela en el fondo de tierras del sistema, el usuario puede localizar la parcela de dos maneras diferentes: a través del visor geográfico o rellenando un formulario con los códigos catastrales de la parcela para que el sistema se encargue de encontrarla automáticamente. Si los límites de la parcela mostrados no son correctos, el usuario puede modificarlos directamente en el visor de mapas. Los cambios serán guardados y revisados posteriormente por un gestor, y sirve además para corregir errores en los datos catastrales. A continuación se deben introducir los datos de los propietarios. El precio del alquiler es automáticamente calculado en base a la localización de la parcela y a sus usos del suelo. Una vez hecho esto, el sistema genera los documentos pertinentes que el usuario podrá descargar en formato PDF, ODT o DOC.

En lo que respecta a las solicitudes de arrendamiento, cuando los usuarios encuentran una parcela que satisface sus necesidades, pueden realizar sus solicitudes

también a través de la aplicación web.

Todas las solicitudes que un usuario realice podrán ser consultadas cuando lo estime oportuno y podrá ver también en qué estado se encuentran: recibida, en estudio, aceptada, denegada, etc.

4.3.3. Gestores

Los gestores son los encargados de realizar los estudios de las solicitudes y de decidir si se aceptan o se deniegan. Algunas funcionalidades a las que su rol tiene acceso son, entre otras muchas:

- Inserción de eventos relacionados con parcelas (visitas para ver el estado de las parcelas, peticiones para cambios de usos, etc.).
- Generación de estadísticas sobre parcelas, solicitudes y usuarios.
- Modificación y validación de la geometría de una parcela a través de varios métodos.
- Conexión con los servicios web de Catastro para obtener las geometrías de las parcelas actualizadas y los códigos catastrales.

Los gestores también pueden adjuntar archivos con toda clase de información relativa a parcelas y las solicitudes, como contratos escaneados, fotografías, etc. Para preservar la integridad del sistema frente a archivos con contenido malicioso, se toman algunas precauciones como la limitación de su tamaño y de su formato, o escaneo periódico de todos los archivos con un anti-virus libre, entre otras acciones. En el caso de que el archivo sea una imagen se creará una miniatura (*thumbnail* en inglés) para que pueda ser mostrada como previsualización en aquellas páginas que lo requieran.

Otra funcionalidad muy útil para los gestores es la asignación automática de los documentos escaneados a las parcelas o solicitudes correspondientes. Por ejemplo, cuando un contrato es escaneado, se guarda en una carpeta en el servidor con el nombre de archivo que indica a qué parcela o solicitud se refiere. De esta manera el

sistema es capaz de detectar cualquier nuevo documento e incorporarlo a la base de datos.

Los gestores pueden también comunicarse entre ellos a través de tres mecanismos diferentes: noticias, avisos y alertas. Las dos primeras son insertadas por algún gestor y las últimas son generadas automáticamente cuando, por ejemplo, va a finalizar algún contrato o se modifican ciertas características de una parcela.

El sistema implementa también la automatización de varias tareas, entre ellas:

- El cálculo de la puntuación asignada a cada solicitud en base a criterios de preferencia que deben ser definidos por entidades capacitadas.
- El cálculo de los precios de las parcelas en base a su localización y los usos del suelo que tengan.
- La actualización de precios de los arrendamientos anualmente en base a la variación del Índice de Precios al Consumo (IPC).

4.4. Evaluación del sistema

El sistema ha sido evaluado en términos de rendimiento y de utilización. Para lo primero se han analizado los tiempos de respuesta del servidor real donde se ejecuta SITEGAL a algunas peticiones que suponen una complejidad espacial. Para lo segundo, se exponen algunos datos de utilización de la herramienta tanto por los gestores como por los usuarios particulares.

4.4.1. Evaluación de rendimiento

Debido a que todos los componentes son multi-plataforma (PostgreSQL, Java, MapServer, etc.), SITEGAL es capaz de ejecutarse en cualquier sistema operativo. En nuestro caso ha sido instalado en dos máquinas virtuales con 4GB de RAM, dos procesadores de doble núcleo Intel X5450 a 3 Ghz, y ejecutando un SUSE Linux Enterprise Server 10 SP1 como sistema operativo. El servidor (Apache), el servidor de mapas (MapServer) y la base de datos geográfica (PostgreSQL+PostGIS)

	1 petición	5 pet/seg	10 pet/seg	20 pet/seg	50 pet/seg
GetMap	1.10	2.21	3.76	6.38	13.81
GetFeatureInfo	0.64	1.46	2.53	4.31	10.14
GetFeature	1.18	2.12	3.31	5.85	11.81

Tabla 4.2: Tiempo medio de petición de una ejecución (en segundos).

son ejecutados en una de las dos máquinas virtuales, mientras que la aplicación Java (ejecutada sobre un Tomcat) y la base de datos de la aplicación (también con la dupla PostgreSQL+PostGIS) han sido instalados en la otra máquina virtual. Las peticiones a la aplicación las gestiona el servidor web Apache que las redirige al contenedor de aplicaciones Tomcat, el cual es sólo accesible desde la red local. Aunque desde su puesta en funcionamiento en 2007 nunca ha habido problemas de sobrecarga, el sistema, que ya ha sido diseñado de forma modular para lidiar con estas circunstancias, permite la replicación de todos sus componentes. Utilizando balanceadores de carga sería posible desviar las peticiones a aquellos componentes con menos carga de trabajo.

En el caso de SITEGAL, la base de datos geográfica contiene: una copia de la capa de parcelas y de usos del SIXPAC (Sistema de Información de Parcelas Agrarias de Galicia); las ortofotos son obtenidas del PNOA (Plan Nacional de Ortofotografía Aérea) que genera imágenes aéreas geométricamente corregidas de España una vez cada dos años [6] y que están accesibles mediante WMS; y el resto de capas están obtenidas de los servicios WMS y WFS de la Oficina Española de Catastro (ver Sección 3.4.3) y del SITGA (Sistema de Información Territorial de Galicia). Por su parte, la base de datos de la aplicación almacena las parcelas incorporadas al fondo de tierras del Banco junto a toda la información referida a los usuarios, procesos de cesión y procesos de arrendamiento.

Las peticiones que implican procesamiento espacial son costosas computacionalmente en este tipo de sistemas. La Tabla 4.2 muestra la media de los tiempos de ejecución para las peticiones más frecuentes que se solicitan al servidor geográfico. Estos datos han sido medidos con diferentes número de peticiones por segundo y para cada clase, las pruebas han sido repetidas cinco veces para calcular su media. Las peticiones seleccionadas por representativas han sido:

- Petición *GetMap* de WMS a la capa de usos del suelo para una zona de 1.000x1.000 metros y un tamaño de imagen de 1.024x1.024 píxeles.
- Petición *GetFeatureInfo* de WMS a la capa de parcelas para una zona de 1.000x1.000 metros y un tamaño de imagen de 1.024x1.024 píxeles.
- Petición *GetFeature* de WFS a la capa de usos del suelo usando un filtro con un *Bounding Box* para una zona de 1.000x1.000 metros.

La base de datos geográfica consta de más 12.000.000 de registros para la capa de parcelas y más de 16.000.000 para la capa de usos del suelo. La Tabla 4.2 prueba que los tiempos de ejecución son muy aceptables para un número moderado de usuarios. El acceso a los datos externos a través de servicios WMS puede ser un cuello de botella ya que, o bien el servidor puede no estar accesible o fuera de funcionamiento, o puede tardar en responder. Para estos casos hay dos soluciones que pueden ser adoptadas: la primera consiste en tener una copia de los datos externos en una base de datos local con actualizaciones periódicas; y la segunda es la implementación de un mecanismo que actúe de forma que cuando haya dos o más fuentes para la misma información, analice las conexiones y envíe la petición al servicio más rápido.

Hay que destacar que todo el sistema ha sido diseñado e implementado pensando en que, en caso de que las necesidades de rendimiento aumenten, se puedan adoptar estas soluciones aplicando los mínimos cambios posibles.

4.4.2. Evaluación de la utilización

Para analizar los datos de uso del sistema ha de tenerse en cuenta el perfil de los potenciales usuarios: gente del rural, de media o avanzada edad, con escasos conocimientos en nuevas tecnologías y con un acceso a internet a veces de poca calidad. Estas características suelen ser un obstáculo bastante grande para el desarrollo de esta clase de iniciativas orientadas a las zonas rurales [21].

A fecha de 2 de mayo de 2013, casi 5 años y medio después del lanzamiento del Banco de Tierras y de SITEGAL, el sistema cuenta con 2.858 clientes: 2.636 de ellos son personas físicas y 222 son personas jurídicas. Además, 1.122 de estos clientes tienen una cuenta de usuario, lo que les permite acceder al sistema y llevar a cabo los

procedimientos a través de la aplicación web. Adicionalmente, los 265 gestores del Banco de Tierras también tienen acceso al sistema para su gestión interna. Desde su comienzo, el Banco ha gestionado 4.416 solicitudes de arrendamiento y 2.515 solicitudes de cesión. De ellas, 1.426 (el 32,29 %) y 656 (el 26,08 %) respectivamente han sido introducidas por clientes registrados a través de SITEGAL. El resto de las solicitudes han sido entregadas en papel en la Sede Central del Banco o en las Oficinas Agrarias Comarcales, encargándose los gestores de introducirlas en el sistema también a través de SITEGAL.

Actualmente el Banco de Tierras tiene 9.291 parcelas: 7.337 están disponibles para su alquiler (de las que 1.500 están ya alquiladas) y las 1.954 restantes no están disponibles por estar siendo inspeccionadas o por cualquier otro motivo. De acuerdo a la ley, todas las parcelas provenientes de procesos públicos de concentración parcelaria, deben estar disponibles en el fondo de tierras del Banco para su alquiler. Lo malo es que estas parcelas son normalmente pequeñas y suelen estar aisladas, lo cual hace que resulte complicado su arrendamiento ya que los usuarios prefieren parcelas más grandes y bien comunicadas. Esta es la razón por la que en la Tabla 4.3 se muestre que sólo el 32,24 % de las parcelas han sido arrendadas o están en proceso de arrendamiento, y sin embargo esto representa el 66,12 % de la superficie gestionada por el Banco. Por lo que respecta a los procesos de cesión, el Banco de Tierras ha aceptado 941 solicitudes de cesión. Esto significa que el 17,10 % de superficie disponible viene de propietarios privados y el 82,9 % pertenece a la administración pública. En el momento de recoger estos datos, 252 solicitudes de cesión estaban siendo estudiadas para su aprobación y 509 peticiones de avisos habían sido introducidas con las que se informa al cliente que una parcela con las características deseadas ha sido introducido en el sistema.

4.5. Sistema de información Web-SIG para la gestión común de parcelas forestales

Como se puede constatar de los resultados aportados en el proyecto del Banco de Tierras del que se ha hablado en la sección anterior, las parcelas que menos se mueven en el mercado de tierras son las que tienen un uso forestal. Esto es debido a

	Parcelas	Superficie (Ha)
Total disponibles	7.337	3.366,79
Arrendadas	1.500 (20,44 %)	1.478,13 (43,90 %)
Solicitudes de arrendamiento en estudio	1.216	1.356,31
Parcelas únicas en las que hay al menos alguna solicitud de arrendamiento	866 (11,80 %)	748,33 (22,22 %)
Solicitudes de cesión aceptadas	941 (12,82 %)	575,65 (17,10 %)
Solicitudes de cesión en estudio	252 (3,43 %)	140,48 (4,17 %)

Tabla 4.3: Parcelas en el fondo de tierras del Banco y arrendamientos realizados a través de SITEGAL en mayo de 2013.

que su explotación es más complicada y los beneficios económicos más limitados, por lo que en muchas ocasiones acaban siendo abandonadas. Además, es frecuente que las parcelas tengan un difícil acceso tanto para maquinaria como para personas, lo que hace que sean muchos los propietarios que desconocen su localización o incluso su existencia.

Para gestionar este tipo de parcelas la Xunta de Galicia creó las Unidades de Gestión Forestal (UXFOR). Se puede considerar a una UXFOR como *un conjunto de parcelas forestales cuyos propietarios se asocian para realizar una gestión conjunta del terreno*. Por la ley comentada, estas asociaciones tienen entidad legal y se le asignan ciertos beneficios. Para poder constituir una UXFOR, la superficie mínima que deben sumar las parcelas ha de ser de 15 hectáreas forestales y el conjunto todas ellas tiene que ser una superficie cerrada y sin huecos, esto es, no puede haber en medio parcelas que no formen parte de la unidad forestal, salvo que se trate de parcelas de uso no forestal. Una vez constituida la UXFOR, todos los propietarios deciden en común, y mediante un sistema de mayorías, su futuro basándose en criterios económicos, sociales y ambientales. No obstante cada propietario conserva sus propiedades sin que sufran modificación alguna en su superficie o situación, y cuyos dominios recuperará cuando se extinga la UXFOR. La Administración, por su parte, apoyará el cumplimiento de los objetivos y compromisos fijados y velará por el interés de todos los implicados en la UXFOR. Además proporcionará ayudas para incentivar la constitución, la inversión de capital y el funcionamiento normal de las UXFOR.

Para tramitar las solicitudes de creación de las UXFOR se desarrolló una aplicación Web-SIG construida en base a la arquitectura explicada en el Capítulo 3 llamada SIUXFOR. La ley sobre la cual se definen las UXFOR [140] fue promulgada por el Gobierno gallego en 2007. Las elecciones de 2009 supusieron el cambio de dicho Gobierno que trajo consigo diversas reformas en el mundo rural. Una de ellas fue la supresión de las UXFOR y la puesta en marcha de las llamadas Sociedades de Fomento Forestal (SOFOR), lo que supuso que SIUXFOR dejase de utilizarse por la Administración. No obstante, las lecciones aprendidas durante el desarrollo de dicha aplicación y los avances que esta tenía con respecto a SITEGAL han servido, por una parte, para modernizar el propio SITEGAL y, por otra parte, para adquirir una experiencia que ha sido puesta en práctica en sucesivos proyectos.

4.6. Resumen

En este capítulo se ha explicado cómo funciona el Banco de Tierras de Galicia y cómo se ha implementado su herramienta de gestión SITEGAL, una aplicación Web-SIG basada en la arquitectura explicada en el Capítulo 3. Su principal objetivo es la dinamización de los mercados de tierras ofreciendo un sistema con capacidades espaciales y accesible a través de la web. Este sistema facilita la transferencia de tierras entre propietarios que no utilizan las parcelas y agricultores que quieren ampliar sus explotaciones, reduciendo el coste de estas operaciones. Actualmente, los propietarios de las parcelas pueden publicar sus propiedades y los agricultores pueden buscar parcelas que se adapten a sus necesidades. Esto reduce el papeleo y facilita tanto la publicación como el arrendamiento de parcelas, mejorando así la conexión entre ofertantes y demandantes de terrenos, situación muy complicada hasta ahora debido a los altos costes de información que implican. Todos los procedimientos administrativos para llevar a cabo las cesiones y los arrendamientos son hechos por empleados públicos utilizando la herramienta desarrollada, promoviendo así el e-Government.

SITEGAL, accesible desde www.bantegal.com/sitegal, ha sido adoptado por el Banco de Tierras de Galicia y se ha convertido en el eje central del funcionamiento del Banco. El sistema también ha atraído la atención de varias entidades nacionales e internacionales a cargo de políticas sobre la gestión de tierras. Es el caso de la

Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), que ha incorporado la implementación de SITEGAL en el proyecto *Support to the Preparation of a National Land Consolidation Strategy and a Land Consolidation Pilot Project in Albania (TCP/ALB/3301)*³. Este proyecto está enmarcado en las actividades de cooperación de la FAO para la mejora de la gestión de tierras en Europa del Este. Como resultado de esto, SITEGAL está actualmente siendo probado en una zona rural de Albania.

SITEGAL es un proyecto de gran envergadura que lleva casi 6 años en funcionamiento. Durante este tiempo no se han dejado de añadir funcionalidades y de mejorar las ya existentes conforme a las experiencias y necesidades de los usuarios. En 6 años el mundo de las nuevas tecnologías cambia enormemente y se hace necesario adaptar la aplicación a herramientas más modernas y eficientes. Por ejemplo, al principio del desarrollo de SITEGAL se eligió Struts como framework para la implementación de la parte Vista y Controlador. Pasados los años, Struts evolucionó en Struts 2 y más tarde aparecieron otros frameworks como Tapestry o Wicket. En otros proyectos desarrollados por el Grupo de Arquitectura de Computadores se decidió utilizar este último, por lo que fue el elegido para aplicarlo también a SITEGAL. Evidentemente mudar todas las funcionalidades de Struts a Wicket era inviable e innecesario, por lo que se optó por utilizarlo solamente en la implementación de las nuevas funcionalidades y en la parte de autenticación y registro de usuarios, que fueron los casos de uso elegidos para las pruebas. La separación lógica entre las peticiones gestionadas por Struts y gestionadas por Wicket se ha hecho en base a filtros sobre las URLs: las peticiones para Wicket tienen una componente característica que el servidor de aplicaciones detecta y redirige hacia las clases apropiadas. Otro potencial problema era el de mantener la sesión del usuario activa en ambos frameworks, pero queda resuelto delegando la gestión de autenticación y autorización a Spring Security, que está un nivel por debajo.

SITEGAL es la herramienta central del Banco de Tierras. Esto implica que otras aplicaciones propias del Banco pero externas a SITEGAL son gestionadas también por este último. Por ejemplo, con SITEGAL ya en funcionamiento, se creó una aplicación externa utilizada para todo el tema de facturación. Se demandaba que esta

³Más información sobre el proyecto en <http://www.fao.org/europe/activities/land-tenure/land-tenure/en/>

aplicación fuera accesible a través SITEGAL y que se pudiera navegar de una aplicación a otra sin necesidad de autenticarse en cada una. Lo que se hizo fue implementar un mecanismo de gestión común de sesiones por el cual SITEGAL, encargado de la autenticación, escribía en una tabla accesible por las dos aplicaciones, datos sobre los usuarios autenticados para que la aplicación externa pudiera saber si debía permitir el acceso o no. Este sistema tiene todas las medidas de seguridad necesarias para evitar incongruencias entre ambas aplicaciones como son *claves hash* para la identificación única de sesiones y tiempos de caducidad de las mismas.

En este capítulo también se ha hecho una breve referencia sobre SIUXFOR, otro sistema de información Web-SIG cuyo principal objetivo fue la optimización de la gestión comunes de parcelas forestales.

Las herramientas desarrolladas comentadas en este capítulo, junto con el sistema de información Web-SIG base, han dado lugar a varias publicaciones y participaciones en congresos y workshops que están referenciadas en [106, 107, 99, 43, 42].

Parte III

Metodologías y algoritmos para los procesos de ordenación y planificación del territorio

Capítulo 5

Algoritmos de búsqueda heurística aplicados a problemas de asignación espacial

El proyecto de investigación *Sistemas de información geográfica para planeamiento urbanístico y ordenación territorial mediante técnicas de optimización sobre procesadores multinúcleo* financiado por la Xunta de Galicia¹ tuvo como objetivo principal el desarrollo de herramientas que ayuden y faciliten la resolución de problemas de asignación de categorías del suelo en los planes generales de ordenación municipal. Este capítulo explica las cuestiones comunes a las herramientas desarrollados y que se explican en profundidad en los capítulos 6 y 7.

La estructura del capítulo es la siguiente: la Sección 5.1 aporta más detalles sobre los objetivos del citado proyecto; en la sección 5.2 se describen los algoritmos genéticos y su buena adaptación a los problemas de asignación espacial; en la Sección 5.3 se exponen las características de los algoritmos *greedy* iterativos basados en población; en la Sección 5.4 se explican características comunes a ambos algoritmos; y en la Sección 5.5 se enumeran las tecnologías que han sido utilizadas en sus desarrollos.

¹En concreto por el Plan Gallego de Investigación INCITE de la Xunta de Galicia. Código del proyecto: 08SIN011291PR

5.1. Contexto de la investigación sobre algoritmos para asignación espacial

El proyecto *Sistemas de información geográfica para planeamiento urbanístico y ordenación territorial mediante técnicas de optimización sobre procesadores multinúcleo* surge de la necesidad de disponer de un conjunto de herramientas que faciliten y doten de mayor eficiencia a los procesos de planificación urbanística y ordenación territorial. Estos dos términos, planificación y ordenación, pueden ser aplicados a diferentes procesos relacionados con la regulación y mantenimiento de los usos de la tierra. Según la FAO², la planificación urbanística es *la evaluación sistemática de las posibles alternativas para la tierra y el agua basándose en condiciones económicas y sociales para seleccionar y adoptar la mejor opción de asignación entre el suelo y sus usos* [53].

El proceso de desarrollo de un plan exhaustivo de ordenación territorial es largo, laborioso y requiere de un gran esfuerzo por parte de las administraciones públicas y equipos técnicos para lograr una solución satisfactoria. Como consecuencia de esto, existe un incremento en la demanda de nuevas metodologías de planificación y de herramientas que permitan llevarlas a cabo de una forma eficiente. No obstante, las soluciones que proporcionen dichas herramientas han de ajustarse, evidentemente, a las especificaciones legales y a ciertos criterios predefinidos.

Hablando de manera muy genérica se puede decir que las herramientas que se han desarrollado durante el proyecto tienen como objetivo la asignación de categorías del suelo a parcelas de modo que el resultado satisfaga una serie de criterios previamente establecidos. Los procesos de asignación suelen ser de ámbito municipal, esto es, las parcelas involucradas en los planes pertenecerán todas a un mismo municipio. Esta idea dificulta el proceso de ordenación debido a la heterogeneidad de las categorías con las que nos podemos encontrar. Además, las categorías que cubren zonas urbanas o núcleos rurales deberían ser tratadas con especial atención ya que por un lado esas categorías son las que generan mayor controversia social entre los habitantes de la zona, principalmente debido a las restricciones de edificación. Y por otro lado, los condicionantes técnicos del proceso de asignación de categorías suelen ser muy

²Organización de las Naciones Unidas para la Alimentación y la Agricultura (Food and Agriculture Organization of the United Nations, en inglés)

específicos y difíciles de satisfacer.

Es por esta razón por la que se ha decidido separar el proyecto en dos fases: en la primera fase se delimitan los núcleos rurales y se clasifican sus parcelas en categorías; y en la segunda, y teniendo ya los núcleos rurales delimitados, se asignan categorías del suelo a las parcelas del municipio que no formen parte de ningún núcleo rural, elaborando así un plan de ordenación territorial.

Este diseño en dos fases ha dado lugar al desarrollo de dos metodologías: una para la delimitación y clasificación de núcleos urbanos, y otra para la asignación de categorías del suelo a parcelas en procesos de ordenación. Por simplificar, la primera fase se le llamará *delimitación* y a la segunda *ordenación*. Además ambas metodologías están acompañadas de sendos algoritmos que tiene como objetivo obtener soluciones satisfactorias en tiempos lo suficientemente cortos como para conseguir que las herramientas sean utilizables por los expertos en planificación y ordenación.

Antes de entrar en los detalles de los algoritmos, es conveniente distinguir entre municipio y entidad singular de población. Un municipio puede ser definido como *una entidad administrativa claramente definida por un término municipal y que puede agrupar a una o a varias localidades*. Mientras que una entidad singular de población es *cualquier área habitable del término municipal, habitada o excepcionalmente deshabitada, claramente diferenciada dentro del mismo, y que es conocida por una denominación específica que la identifica sin posibilidad de confusión*³. Esta definición nos lleva a otras dos: un área se considera habitable cuando *existen en la misma viviendas habitadas o en condiciones de serlo*, y un área se considera claramente diferenciada cuando *las edificaciones y viviendas pertenecientes a la misma pueden ser perfectamente identificadas sobre el terreno y el conjunto de las mismas es conocido por una denominación*.

Aclarados estos términos, comentar que para este proyecto se supone que cada entidad singular de población tendrá únicamente un núcleo rural.

³Definición del Instituto Nacional de Estadística: <http://www.ine.es/nomen2/Metodologia.do>

5.2. Algoritmos genéticos en problemas de asignación espacial

Los problemas de asignación espacial se pueden clasificar como problemas de optimización combinatoria, caracterizados por tener un gran número de posibles soluciones. La componente espacial implica una gran complejidad computacional y a menudo son necesarias ejecuciones largas para alcanzar soluciones satisfactorias [136], especialmente cuando los objetivos incluyen características geométricas y espaciales como la compacidad o la vecindad entre parcelas [45]. Debido a esto y al hecho de que el número de objetos geométricos involucrados en los problemas de asignación espacial es normalmente elevado, la búsqueda de la solución óptima necesita de algoritmos capaces de encontrar soluciones aceptables en un tiempo razonable [88]. Con estas características, los algoritmos de búsqueda heurística parecen los idóneos para resolver este tipo de problemas. Concretamente, los algoritmos genéticos ya han demostrado ser eficientes y efectivos cuando se trata de resolver problemas geográficos [136, 88, 2, 10, 48, 57, 71, 74, 77, 120, 137, 138, 144].

Un algoritmo genético se puede definir como el conjunto de pasos organizados que describen un proceso para la resolución de problemas de optimización con búsqueda de soluciones basados en los postulados de la evolución biológica [7]. Para alcanzar una solución se mantiene un conjunto de entidades que representan posibles soluciones, las cuales se mezclan y compiten entre sí de tal manera que las más aptas son capaces de prevalecer a lo largo del tiempo, evolucionando hacia mejores soluciones cada vez. Los algoritmos genéticos son un caso particular de los algoritmos evolutivos [64].

No obstante, de los artículos citados hay que destacar dos características de los algoritmos que describen. La primera es que la mayoría operan sobre una cuadrícula, esto es, adaptan el parcelario a un rectángulo dividido en celdas donde cada celda es la unidad mínima con la que trabajan. El inconveniente de este método es que no es realista y adaptar a posteriori las soluciones al parcelario real es normalmente muy complicado, ya que se pueden dar casos en que una parcela tenga asignada a varias categorías a la vez o que un grupo de parcelas muy diferentes entre sí tengan la misma categoría. Además las leyes del suelo actuales especifican que el estudio de planeamiento tiene que estar basado en las parcelas catastrales. Sólo dos

de los estudios citados anteriormente utilizan unidades espaciales irregulares, pero en estos casos el número de unidades es muy pequeño: 155 en [10] y solamente 6 en [138]. La segunda cuestión a destacar es que la aplicación de este tipo de algoritmos a problemas de demarcación y clasificación de núcleos rurales (lo que es la primera fase del proyecto) es muy escasa, ya que no están orientados a esta clase de problemas. No obstante, hay artículos como [61] en los que sí se ha desarrollado un algoritmo específico para la óptima subdivisión de núcleos rurales en Brasil pero en zonas poco parceladas. En nuestro caso se ha considerado como mejor opción la utilización de otro tipo de algoritmo que, aunque siendo basado en búsqueda heurística, no es genético si no que es de tipo *greedy*, como se explica en la Sección 5.3.

5.2.1. Funcionamiento de un algoritmo genético

Un algoritmo genético es un tipo de algoritmo evolutivo que trabaja con una población de individuos [64]. Un individuo es una posible solución al problema codificada mediante genes. La población será sometida a acciones con cierto carácter aleatorio similares a las que ocurren en la evolución biológica.

Un algoritmo genético simple funciona de la siguiente manera.

1. Se crea una población con individuos aleatorios. Estos individuos, a pesar de generarse aleatoriamente deben ser, por lo general, soluciones válidas al problema.
2. De entre todos los individuos de la población, se escogen dos mediante el operador *selección*.
3. Los seleccionados intercambian sus genes a través del operador *cruce* dando lugar a dos nuevos individuos: la descendencia.
4. Con el operador *mutación*, genes escogidos aleatoriamente de cada descendiente cambian su valor de una manera también aleatoria.
5. El operador *elección* escoge el mejor individuo de la descendencia y lo inserta en una nueva población.

6. Si a la nueva población le faltan individuos para completarse, se vuelve a 2. Si ya está completa, se sigue con el punto 7.
7. Si la nueva población converge, esto es, que se ha llegado a una solución aceptable, el algoritmo termina proporcionando como salida el mejor individuo de la población. En caso contrario, el algoritmo vuelve a empezar desde el punto 2 pero utilizando la nueva población.

En la fase de elección de los mejores individuos se puede aplicar una técnica, llamada *elitismo*, que consiste en tener en cuenta también a los padres cuando se selecciona el mejor individuo, en vez de comparar solamente la descendencia. De esta manera se asegura la conservación de los mejores individuos en las nuevas generaciones.

5.3. Algoritmo *greedy* iterativo basado en población

La primera intención cuando se empezó con el diseño de los algoritmos era utilizar un algoritmo genético para cada uno de los dos problemas a resolver: el de delimitación y clasificación de núcleos rurales, y el de la asignación de categorías del suelo a parcelas en procesos de ordenación. Y aunque a este último problema los algoritmos genéticos se amoldan perfectamente, en el caso del primero se llegó a la conclusión de que no era así.

Debido a motivos que se comentarán más detalladamente en la Capítulo 6, el algoritmo para la delimitación de núcleos rurales funciona mejor prescindiendo del operador de cruce, y por ende del operador de selección. Dada entonces una población de individuos y sólo utilizando el operador de mutación, el tipo de algoritmo que más se adapta a estas circunstancias es el clasificado como algoritmo *greedy* iterativo basado en población (*PBIG* por sus siglas en inglés *Population-Based Iterated Greedy*). Un algoritmo *greedy iterativo* (IG) es un algoritmo de búsqueda heurística que utiliza decisiones basadas en óptimos locales en cada iteración [33, 97]. Los algoritmos IG son aplicados a una gran variedad de problemas: problemas en *flow shop scheduling* [114, 101, 111], en reconstrucción de estructuras proteicas [128, 22], en

emplazamiento de nodos de fibras ópticas [127], o en problemas de máxima diversidad [84]. Los algoritmos PBIG son muy recientes en su utilización y se han aplicado a problemas tales como *minimum weight vertex cover problems* [25], *maximal covering location problems* [112], o problemas de nivelación de recursos [9]. Comentar que el término *Individuo* se seguirá utilizando para referirse a una solución perteneciente a una población, para así usar la misma nomenclatura que en los algoritmos genéticos.

5.3.1. Funcionamiento de un PBIG

Básicamente un algoritmo IG genera una secuencia de soluciones elegidas mediante heurísticas *greedy*, es decir, eligiendo la solución óptima en cada paso local con el objetivo de llegar a un óptimo global. El algoritmo IG clásico sólo trabaja con una solución pero sus variantes basadas en población extienden ese comportamiento utilizando un conjunto de soluciones que serán mejoradas de forma paralela [112], como es el caso del algoritmo de delimitación de núcleos rurales presentado en esta tesis.

En los algoritmos IG es común utilizar una operación conocida como *destruir y construir* que intenta mejorar los individuos. Este operador primero destruye parte de una solución y luego la reconstruye hasta volver a tener una solución completa [9]. En nuestro caso los algoritmos utilizan en su lugar la operación borrar y añadir (*remove-and-add operator*, *R&A Op.*), que tiene como objetivo borrar y añadir parcelas de los diferentes conjuntos de parcelas con la misma categoría, creando nuevas soluciones.

5.4. Características comunes en los algoritmos implementados

Al trabajar sobre un mismo ámbito, algunas características son comunes a los dos algoritmos que se han implementado. Concretamente: la representación de los individuos y el cálculo de las relaciones de vecindad entre parcelas. Estos dos aspectos juegan un papel muy importante en lo que se refiere al rendimiento y al tiempo de ejecución.

5.4.1. Codificación de los individuos

Cómo representar los individuos es una de las cuestiones más importantes a la hora de diseñar algoritmos potencialmente críticos en cuanto al consumo de recursos. En [89] los autores proponen dos métodos distintos para la representación de los individuos en problemas de ordenación territorial. La primera, llamada representación *Land Block* (LB), usa un array unidimensional en el que cada elemento del array, *land block*, representa una parcela existente. En la segunda representación, denominada *Percentage & Priority* (P&P), los genes tienen dos valores, el porcentaje y la prioridad de asignación de cada categoría. El problema de la primera representación es que normalmente necesita muchos bloques (tantos como parcelas), lo que hace que se eleven los recursos necesarios y el tiempo de ejecución. Por su parte, la segunda representación necesita definir más operaciones y realizar más iteraciones para llegar a un resultado aceptable. Existe un tercer método de representación basado en un array bi-dimensional, llamado grid [77], en el que la región para la que se hace el plan de ordenación se divide en celdas que representan pequeñas porciones del territorio. Este método fue rechazado para aplicarlo en los algoritmos desarrollados debido a la dificultad que existe a la hora de adaptar las soluciones proporcionadas por el algoritmo a las parcelas existentes en la zona de estudio, ya que, como se ha comentado, tanto en los procesos de delimitación de núcleos rurales como en los de ordenación territorial se recomienda no alterar las geometrías de las parcelas.

Finalmente la opción elegida fue la primera representación, LB, por los siguientes motivos: encaja perfectamente con la información geográfica de las parcelas disponible; las posibles exigencias de recursos computacionales debido a la elevada longitud de los individuos puede ser mitigada utilizando tipos de datos básicos y estructuras de datos eficientes (como la comentada por ejemplo en el apartado 5.4.2); y los posibles inconvenientes en cuanto al tiempo de ejecución pueden ser disminuidos con técnicas de computación de altas prestaciones.

5.4.2. Relaciones de vecindad entre parcelas

El valor de una parcela no sólo reside en las cualidades intrínsecas de la misma, si no que también influyen enormemente factores de su entorno. Es por esto que las

relaciones de vecindad entre parcelas, viarios, infraestructuras, etc., tienen mucha importancia en el tipo de problemas como los que se abordan en los siguientes capítulos.

Se considera que dos parcelas son vecinas si la longitud de la frontera compartida es mayor que cero. Esto implica que dos parcelas que se tocan en un sólo punto, no se consideran vecinas. El cálculo de las relaciones de vecindad es una operación espacial que es muy costosa computacionalmente, lo que implica tiempos de ejecución altos. Por tanto, en los algoritmos implementados estos cálculos se realizan en una etapa de preprocesado anterior a la ejecución de los algoritmos.

La información sobre la vecindad de las parcelas será consultada con frecuencia por los algoritmos, por lo que almacenar los datos en estructuras de acceso rápido es una cuestión fundamental. Teniendo en cuenta que el número de vecinos puede ser distinto entre las parcelas, la primera idea fue utilizar un array bidimensional de tamaño $L \times L$, siendo L el tamaño del individuo, por si una parcela tuviera a todas las demás como vecinas. No obstante, rápidamente se comprobó que no era muy eficiente en el acceso a los datos, debido sobre todo a que Java⁴ trata como arrays de arrays a los arrays bidimensionales.

Así pues se buscó otra estructura que permitiera un acceso más rápido a los datos, y se decidió utilizar dos arrays unidimensionales: uno para almacenar los vecinos de cada parcela y otro que servirá como índice. La idea es que la entrada i -ésima del array de índices indique la posición del array de vecinos en la que se guarda el primer vecino de la i -ésima parcela, donde $i=1, \dots, NP$, siendo NP el número de parcelas. Los vecinos de cada parcela son almacenados consecutivamente en el array de vecinos. En la Figura 1 se muestra un ejemplo: los vecinos de P2 son P1, P3, P6 y P7 y están almacenados a partir de la posición 4 del array de vecinos (4 es el valor de la segunda entrada en el array de índices).

Los detalles técnicos del cálculo de las vecindades entre parcelas son los siguientes. Se ha utilizado la librería JTS [118], concretamente su objeto *STRtree*. *STRtree* es la implementación utilizando el algoritmo STR [82] de lo que se conoce como árbol-R, una estructura de datos de tipo árbol similar al árbol-B (árbol binario de búsqueda en los que cada nodo puede poseer más de dos hijos, mantienen los da-

⁴Java es el lenguaje de programación utilizado para el desarrollo de los algoritmos. Se dan más detalles de los motivos de su elección en la Sección 5.5.

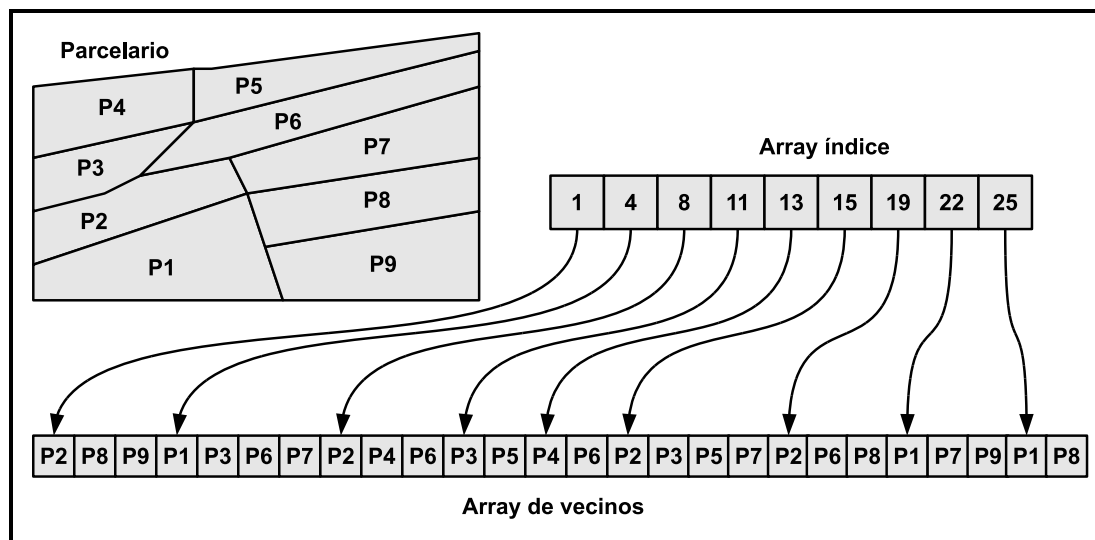


Figura 5.1: Arrays utilizados para almacenar la información sobre los vecinos.

tos ordenados y las inserciones y eliminaciones se realizan en tiempo logarítmico amortizado [33]) pero adaptado para almacenar información espacial [5].

Con el objeto STRtree de JTS podemos insertar pues geometrías en el árbol de manera que posteriormente se pueden hacer consultas sobre él. Se considera que esta estructura de datos es de sólo lectura en el sentido en que una vez construido el árbol, no se podrá añadir o quitar ningún nodo más.

Para insertar geometrías se utiliza el método *insert* al que, en nuestro caso, se le pasa la geometría de la parcela y su envolvente interna para que lo sitúe adecuadamente en el árbol. Para hacer consultas se utiliza el método *query*, el cual dada una envolvente, devuelve una lista de geometrías cuyos límites intersecan con ella. La primera vez que se llama a *query* es cuando se crea el árbol, impidiendo modificaciones en él.

Para saber qué parcelas son las que se corresponden con las geometrías devueltas por el método *query* utilizamos un objeto *Map* mediante el cual relacionamos cada geometría con la parcela a la que corresponde.

El método *query* se aplica a cada parcela, lo que le proporciona su lista de *posibles* vecinas. Si la longitud de la frontera de la intersección entre la *posible* vecina y la parcela que estamos examinando es mayor que 0, se podrá decir que existe una

relación de vecindad entre ambas. En otro caso, esta relación no existe.

5.5. Tecnologías utilizadas

El lenguaje de programación utilizado es Java (<http://java.com/>). Es importante destacar que su rendimiento con respecto a lenguajes nativos ha sido mejorado en los últimos años, gracias sobre todo al compilador Just-in-Time (JIT) de la máquina virtual de Java (JVM). Además, Java está siendo muy utilizado en sistemas de alto rendimiento debido a sus mejoras en los últimos años [125].

Este buen comportamiento del lenguaje de programación se produce sobre todo cuando se utilizan tipos de datos básicos, ya que trabajar con objetos es computacionalmente más costoso. Es por esta razón por la que, durante el diseño de los algoritmos implementados, se ha tratado de evitar la utilización de objetos cuando el consumo de recursos se prevé elevado. En cálculos más livianos computacionalmente, el uso de objetos o no, no tiene tanta importancia, dejando la elección en base a, sobre todo, la sencillez a la hora de programar.

La elección de Java como lenguaje de programación también se ha basado en el hecho de que existen muchas librerías y aplicaciones escritas en Java, o que tienen su versión para Java, lo que facilita el desarrollo y la integración de los algoritmos.

Para promover la utilización de los algoritmos, estos se han integrado en software SIG de escritorio. Concretamente se ha utilizado SEXTANTE sobre gvSIG, dos herramientas robustas, competentes y muy utilizadas dentro del software SIG libre. Para los cálculos espaciales y operaciones geométricas se han utilizado las librerías JTS y GeoTools, ambas fuertemente respaldadas y usadas por toda la comunidad SIG.

5.6. Resumen

En este capítulo se ha explicado por qué los algoritmos genéticos son una buena elección a la hora de diseñar e implementar soluciones que tengan que ver con asignación espacial, aunque para problemas de delimitación se ha optado por utilizar

otra clase de algoritmos como los *greedy* iterativos. El funcionamiento básico de estos dos tipos de algoritmos de búsqueda heurística también ha sido descrito.

Se ha dado también una visión general de los aspectos que comparten ambos algoritmos, como es la codificación de las soluciones y el cálculo de las relaciones de vecindad utilizando algunas estructuras de datos caracterizadas por su rápido acceso a la información.

Finalmente se comentaron algunos detalles de las tecnologías que se han utilizado durante todo el ciclo de desarrollo de los algoritmos.

Capítulo 6

Algoritmo PBIG para la delimitación de núcleos rurales

Los núcleos rurales son pequeños grupos de edificaciones situados entre ciudades en los que predominan los usos del suelo residenciales y las actividades relacionadas con la agricultura. Aunque esta última característica hoy en día ya no es tan significativa a la hora de definir los núcleos rurales [95], sí es cierto que ha condicionado la evolución y la distribución de muchas de las aldeas que existen hoy en día [69]. Este modelo puede ser representativo de los asentamientos de la Europa rural, presente también en el noroeste de la Península Ibérica [60], pero no exclusivo, ya que se puede encontrar por muchas regiones del mundo [58, 117, 96, 81, 121].

Los estudios más antiguos sobre los núcleos rurales se centran sobre todo en la identificación de la distribución espacial de los mismos. Por ejemplo, en [39] se distinguen dos tipos de núcleos rurales con características opuestas (granjas aisladas y aldeas), y numerosas graduaciones entre ellos (compactas irregulares, lineales, circulares, con crecimiento irregular, con crecimiento en los suburbios, etc.), mientras que en [117] se clasifican en asentamientos dispersos y asentamientos compactos. Estudios posteriores comenzaron a fijarse en las características funcionales y geométricas. En [96] por ejemplo se analiza la morfología de los asentamientos de una región de la India de acuerdo con el tipo de relaciones funcionales, la forma geométrica de sus calles y edificaciones, y los patrones existentes entre grupos de asentamientos. Por otro lado, en [69] se identifican núcleos rurales en Israel fijándose en su densi-

dad de edificación, su tamaño, su entorno, y en la presencia o ausencia de núcleos dispersos. Sin embargo estudios más recientes han intentado identificar las funcionalidades internas presentes en los núcleos rurales. Por ejemplo, en [121] se propone la separación entre las áreas residenciales y las áreas agrícolas para intentar realizar delimitaciones más claras y crear núcleos más compactos. En [58] se distinguen dos tipos de crecimientos de los asentamientos, concentrado y disperso, de acuerdo con el valor del *índice de forma*¹ junto a otras características. Más reciente es la diferenciación de tres tipos de núcleos rurales basándose en su funcionalidad turística, de recreo o agrícola [12]. No obstante, no se han encontrado estudios de métodos científicos o técnicas de planeamiento para la delimitación de núcleos rurales a los que haya que asignar categorías, más allá de las propias especificaciones y procedimientos establecidos por las leyes [81] y políticas [129] de las zonas afectadas o del método propuesto por [60] para la delimitación de zonas urbanas en regiones con poca densidad.

Este capítulo tiene como objetivo explicar en detalle las características de una nueva metodología para los procesos de delimitación de núcleos rurales y el desarrollo de un algoritmo de asignación de categorías del suelo para llevar a cabo las delimitaciones. Este algoritmo ha sido diseñado basándose en las actuales leyes de planeamiento gallegas y en criterios relacionados con distancias entre edificaciones, el número total de edificaciones, la densidad de edificación, la superficie ocupada, la superficie edificable y la compacidad del área delimitada. Para poder pues utilizar el algoritmo en otra región distinta a la gallega, lo único que hay que redefinir son las restricciones relacionadas con las leyes, ya que los demás criterios de la metodología son trasladables a cualquier lugar del mundo.

El algoritmo proporciona soluciones válidas y satisfactorias, esto es, soluciones que cumplan todas las restricciones y que tengan una calidad acorde a las necesidades del usuario. Las características de las parcelas (su pendiente, orientación, usos del suelo, etc.) y las relaciones entre los demás elementos del núcleo (edificaciones, carreteras, vías del tren, etc.) son clave para discernir entre la asignación de una u otra categoría. Cada una de estas variables han sido cuantificadas a través de un proceso analítico jerárquico (AHP por las siglas en inglés de Analytic Hierarchy

¹El *índice de forma* (*shape index* en inglés) es una medida de compacidad que relaciona el área y el perímetro de una superficie

Process) [24] en un sistema de apoyo a las decisiones multi-criterio utilizando para ello el software MPC 2.0² [113] y con la participación de 12 expertos en procesos de planificación.

Este capítulo se estructura como sigue: en la Sección 6.1 se describen cuales son los criterios que las demarcaciones han de cumplir, tanto las marcadas por las leyes vigentes como las definidas por la nueva metodología propuesta. La Sección 6.2 explica la etapa de preprocesado, donde las operaciones espaciales más costosas son ejecutadas. La Sección 6.3 introduce los aspectos generales del algoritmo mientras que sus fases son detalladas en la Sección 6.4. Finalmente, la Sección 6.5 muestra los resultados del algoritmo aplicándolo a varios asentamientos con características bien diferenciadas y la Sección 6.6 enumera las conclusiones obtenidas en este proyecto.

6.1. Criterios para la demarcación de núcleos rurales

Los procesos de delimitación son básicamente definidos por una serie de reglas impuestas por la legislación vigente. No obstante, los expertos en planificación, cuando se enfrentan a estos procesos aprovechan su experiencia para conseguir soluciones de más calidad de manera más eficiente. Acorde con estos criterios se define una nueva metodología para la asignación de núcleos rurales y asignación de categorías. Así, los expertos en planificación podrán comparar los resultados obtenidos ejecutando el algoritmo conforme a la ley y los obtenidos conforme a esta nueva metodología, con el objetivo de mejorar los actuales métodos de delimitación.

De aquí en adelante el término *edificación* se utilizará para cualquier construcción que esté contenida en una parcela. Especial atención tienen las *viviendas*, que son aquellas edificaciones destinadas a ser habitadas por personas. Las edificaciones pueden ser *tradicionales* o *nuevas*, dependiendo de su situación en el núcleo, materiales de construcción, altura y sobre todo antigüedad. Esta clasificación suele ser realizada por los expertos en planificación en trabajos de campo, *in situ*.

²Software diseñado para facilitar la aplicación de la metodología de toma de decisiones AHP basada en la comparación por pares.

Para una mejor comprensión del texto, la Tabla 6.1 agrupa todas las notaciones que se van a utilizar en las siguientes secciones de este capítulo.

6.1.1. Marco legal

La primera ley autonómica³ sobre el uso del suelo en Galicia data de 1985 y se denominó *Ley 11/1985 de Adaptación del Suelo de Galicia*, o LAGSA. Se publicó en el Diario Oficial de Galicia (DOGA) el 29 de agosto de 1985. En esta ley se distinguían tres tipos de áreas en los núcleos rurales: *núcleo rural existente de carácter tradicional*; *núcleo rural existente de reciente formación*; y *núcleo rural pre-existente*.

La segunda ley que apareció fue la *Ley 1/1997 del Suelo de Galicia*, publicada en el DOGA el 26 de marzo de 1997, que añade el concepto del *suelo de núcleo rural*, que si bien puede equivaler a la categoría de *suelo no urbanizable de núcleo rural tradicional* de la anterior ley, este nuevo concepto es más coherente con la realidad y además se establece en un nuevo régimen jurídico diferente al de las otras categorías.

La tercera ley en este ámbito es la *Ley 9/2002 de Ordenación Urbanística y Protección del Medio Rural de Galicia* publicada en el DOGA el 31/12/2002 [139]. En esta ley, y en sus posteriores modificaciones, se establecen las siguiente áreas en los núcleos rurales: área de núcleo rural y área de expansión del núcleo

Esta ley amplía el concepto de *consolidación* también al núcleo rural, ya que hasta el momento sólo estaba presente en los núcleos urbanos. La consolidación indica si un núcleo está suficientemente edificado como para poder considerarlo como tal. Esta ley establece que un núcleo rural debe tener un 50 % de consolidación como mínimo y el área de expansión un 33 %. Además, dentro del área del núcleo rural debe existir una separación máxima de 50 metros entre edificaciones tradicionales y la superficie mínima edificable debe ser de 300 m² para el área del núcleo rural y de 600 m² para el área de expansión.

Otra modificación de esta ley fue la denominada *Ley 2/2010 de Medidas Urgentes de Modificación de la ley 9/2002 de Ordenación Urbanística y Protección del Medio Rural de Galicia*, publicada en el DOGA del 31 de marzo de 2010 [142]. Cabe destacar

³Es la primera porque aunque existían algunas normas anteriores, estas tenían carácter provincial

Siglas	Significado
PBIG	Algoritmo population-based iterated greedy.
NR	Núcleo rural.
NRHT	Núcleo rural histórico tradicional.
NRC	Núcleo rural común.
Máx-NR	Delimitación máxima del NR.
Máx-NRHT	Delimitación máxima del NRHT.
Máx-NRC	Delimitación máxima del NRC.
GC	Grado de consolidación.
NE	Número de edificaciones.
SE	Superficie edificable.
SDN	Superficie de la delimitación del núcleo.
SPME	Superficie de parcela mínima edificable.
PEC	Parcelas con edificación.
PVSE	Parcelas vecinas directas.
PSM	Suma de las áreas de todas las parcelas de área menor que SPME.
DMC	Distancia media característica.
EN	Edificaciones nuevas.
VN	Viviendas nuevas.
ET	Edificaciones tradicionales.
VT	Viviendas tradicionales.
APT	Aptitud.
COMP	Compacidad.
NRHT_EDIF	Edificaciones en el NRHT.
NR_EDIF	Edificaciones en el NR.
NRHT_SUP	Superficie del NRHT.
NR_SUP	Superficie del NR.
R&A Op.	Operador remove-and-add (borrar-y-añadir) del PBIG.

Tabla 6.1: Resumen de las siglas utilizadas en el Capítulo 6.

que con esta ley se cambian los tipos de áreas del núcleo rural por: *área de núcleo rural histórico tradicional* o *NRHT*, *área de núcleo rural común* o *NRC* y *área de núcleo rural complejo*. Este último tipo no es más que el nombre que recibe un núcleo que tiene tanto NRHT como NRC.

Por lo tanto se puede decir, a modo de resumen, que la evolución de las leyes del suelo de Galicia en relación a las tipologías de áreas de los núcleos rurales, ha seguido el camino de intentar identificar lo que es tradicional y lo que es nuevo. Dada la desestructuración y pérdida de referencias morfológicas de los núcleos y de su valor patrimonial como conjunto, la utilización de un *grado de consolidación* como criterio también para la delimitación del área común, pretende incluir bastantes edificaciones recientes que se encontraban fuera de planeamiento con las leyes anteriores. Las dos últimas, la 9/2002 y la 2/2010, presentan reglas de delimitación basadas en distancias, pero como la realidad espacial es definida por la variabilidad morfológica de los núcleos, estos criterios pueden resultar incoherentes con el territorio y la gente que lo habita [13].

El 11 de mayo de 2011 se publicó en el DOGA la *Instrucción 4/2011, de 12 de abril, sobre la metodología de cálculo del grado de consolidación edificatoria en la delimitación del suelo de núcleo rural, al amparo de lo dispuesto en la Ley 2/2010, de 25 de marzo, de medidas urgentes de modificación de la Ley 9/2002, de 30 de diciembre, de ordenación urbanística y protección del medio rural de Galicia*. Según esta modificación, la ley anterior no era lo suficientemente flexible e impedía, en la práctica, delimitaciones que dieran respuesta a las necesidades de crecimiento y desarrollo de los asentamientos rurales, lo que hacía inviable gran parte de los planes generales.

Grado de consolidación

El grado de consolidación hace referencia al grado de densidad edificatoria del núcleo. Decimos que un área es consolidada cuando supera un determinado grado de consolidación. Pero no todas las edificaciones del núcleo son tenidas en cuenta para este cálculo, sólo lo harán aquellas que sean identificadas como tales en estudios previos.

Para el cálculo de la consolidación se han utilizado dos métodos: uno propuesto

por la Agencia de Protección de la Legalidad Urbanística (APLU) y otro basado en la relación entre parcelas. Con la Instrucción 4/2012 la Administración ha propuesto dos nuevos métodos, uno numérico y otro gráfico, con la intención de proporcionar más coherencia al cálculo de la consolidación. A continuación se explicará cada uno de ellos.

Método de la APLU: La Agencia de Protección de la Legalidad Urbanística, instrumento organizativo que colabora con la Consellería de Medio Ambiente, Territorio e Infraestructuras y con las administraciones municipales con el objetivo de mejorar la calidad en la ordenación urbanística, propone un método⁴ para el cálculo de la consolidación basado en el número de parcelas edificables y las edificadas.

Adaptando este método a nuestro algoritmo, el cálculo del grado de consolidación del núcleo (GC) se obtendría según la siguiente ecuación.

$$GC(\%) = \frac{NE}{0,8 * SE} \quad (6.1)$$

siendo NE el número de edificaciones y SE la superficie edificable, que a su vez es definida como:

$$SE = \frac{SDN}{SPME} \quad (6.2)$$

siendo SDN la superficie de la delimitación del núcleo y $SPME$ la superficie de parcela mínima edificable.

Método basado en la relación de parcelas: Sea PEC el número de parcelas con edificación tenidas en cuenta para la consolidación, $PVSE$ el número de parcelas vecinas directas (que toquen o que estén a menos de 10 metros) de un vial que no contenga edificaciones, y PSM la suma de las áreas de todas las parcelas cuyo área sea menor que $SPME$. Se define entonces el grado de consolidación mediante la siguiente ecuación:

$$GC(\%) = \frac{PEC}{PVSE + \left(\frac{PSM}{SPME}\right)} \quad (6.3)$$

⁴Accesible en http://www.axenciaurbanistica.es/index.php?option=com_content&id=100/&Itemid=97

Según la legislación vigente, el grado de consolidación mínimo que ha de tener un NRHT es del 50 %, y de un 33 % para el NRC.

Métodos propuestos por la Instrucción 4/2011: Consciente de la gran disparidad de criterios que se utilizaban para la definición de consolidación y más aún para su cálculo, la Secretaría General de Ordenación del Territorio y Urbanismo de la Consellería de Medio Ambiente, Territorio e Infraestructuras de la Xunta de Galicia, publica por primera vez en la Instrucción 4/2001 el método oficial que la administración va a utilizar para su cálculo. De esta manera clarifica y unifica el procedimiento para que los expertos en planificación puedan realizarlo tal cual va a ser examinado.

Esta Instrucción aclara pues ciertas definiciones que podrían no estar del todo claras en leyes anteriores. Estas son las más relevantes y las que afectan directamente al desarrollo del algoritmo propuesto:

- Grado de consolidación: relación entre el número de parcelas edificadas y el número de parcelas edificables, expresada en porcentaje, para cada ámbito concreto delimitado de núcleo rural.
- Superficie bruta del ámbito: superficie total delimitada de cada tipo de núcleo, incluyendo el viario y las dotaciones, medida sobre el terreno, o sobre la cartografía más detallada disponible.
- Parcela existente: la recogida en el parcelario de la cartografía oficial de Catastro.
- Parcela edificada: parcela en la que existe una edificación destinada al uso previsto en el planeamiento.
- Coeficiente corrector: parámetro cuyo valor, resultado del estudio de un número representativo de núcleos rurales, queda establecido en 1,75.

Los métodos propuestos para calcular el grado de consolidación de un núcleo rural son: el *método gráfico* y el *método numérico o simplificado*, ambos definidos basándose en el estudio de numerosos núcleos rurales pilotos.

El *método gráfico* es un mecanismo de simulación que consiste en proyectar sobre el parcelario existente las posibilidades de parcelamiento, obteniéndose así las posibles parcelas edificables según la ordenación prevista por el propio plan. Será el método aplicable con carácter general con el que se obtiene gráficamente el número de parcelas edificadas y el número de parcelas edificables.

$$GC(\%) = \frac{n^{\circ}_parcelas_edificadas}{n^{\circ}_parcelas_edificables} * 100 \quad (6.4)$$

Por su parte, el *método numérico* o de *cálculo simplificado*, permite determinar el número total de parcelas edificables aplicando el coeficiente corrector de 1,75 a las parcelas hipotéticas que resultan de dividir la superficie bruta del ámbito entre la parcela mínima edificable que el plan establezca. Tiene un carácter aproximativo, pero no se observa inconveniente en su empleo como instrumento ágil que es, aunque observando los resultados del estudio pormenorizado realizado en el que se basa la presente Instrucción, resulta conservador en relación con el método gráfico.

$$GC(\%) = 1,75 * \frac{n^{\circ}_parcelas_edificadas}{n^{\circ}_parcelas_edificables} * 100 \quad (6.5)$$

Para ambos métodos la manera de calcular el número de parcelas edificables es la siguiente:

$$n^{\circ}_parcelas_edificables = \frac{superficie_bruta_del_ambito}{SPME} \quad (6.6)$$

6.1.2. Nueva metodología propuesta

Aparte de las restricciones impuestas por las leyes se han definido una serie de criterios que deben seguirse para llegar a un resultado aceptable. Estos criterios se describen a continuación.

La Distancia Media Característica

La Distancia Media Característica (DMC) es una medida que, aun no formando parte de la ley, se ha incluido en el desarrollo de esta nueva metodología para com-

probar si se puede establecer un criterio más eficaz que el actual. Según las leyes vigentes, para que una parcela pueda formar parte del núcleo histórico-tradicional ha de distar 50 metros o menos de una edificación tradicional, no estableciendo ninguna distancia máxima para el núcleo común. Como alternativa a los 50 metros, la DMC intenta encontrar una distancia variable según la morfología del núcleo y que está directamente relacionada con la distancia entre edificaciones.

Se definen dos DMC: una aplicada al NRHT y otra al NRC. Para el cálculo de estos valores se siguen los pasos que se enumeran a continuación:

1. El usuario puede escoger el tipo de edificación que se utilizará para el cálculo de las distancias: *viviendas tradicionales* o cualquier *edificación tradicional* para el NRHT, o para el NRC *viviendas* o cualquier *edificación que cuente para los cálculos de consolidación*.
2. Se calcula el centroide de cada edificación del tipo seleccionado.
3. Para cada centroide, se calcula su distancia al resto de centroides.
4. Se ordenan las distancias obtenidas de menor a mayor.
5. Se halla la media de las x primeras distancias más pequeñas, siendo x un porcentaje que será determinado por el usuario como parámetro de entrada.

El resultado final obtenido es la Distancia Media Característica del NRHT o del NRC, según el tipo de edificación seleccionada en el paso 1.

Conjuntos de parcelas vecinas

Otro de los criterios definidos tiene que ver con las relaciones de vecindad entre parcelas. Se define como *polígono* al conjunto de parcelas vecinas entre sí pertenecientes a la misma categoría, NRHT o NRC. Un núcleo rural tiene que cumplir algunas restricciones relacionadas con los polígonos:

- Un NRHT ha de tener como mínimo un polígono con tres viviendas tradicionales. El resto de polígonos tienen que tener como mínimo dos.

- Todos los polígonos pertenecientes a un NRC han de tener al menos dos viviendas.

Las leyes no definen el número de viviendas mínimo que un núcleo debe tener, sino que sólo hacen mención al término *viviendas*, en plural. Por esta razón, es habitual utilizar la condición de *mínimo dos viviendas* en los procesos de delimitación. Sin embargo para el caso del NRHT es conveniente que al menos un polígono cuente con tres o más viviendas tradicionales.

Superficie máxima del núcleo

La superficie máxima de un núcleo debe ser limitada teniendo en cuenta su crecimiento estimado. Considerando la SPME definida por el plan de ordenación del propio municipio, el número de edificaciones existentes y el máximo número de nuevas edificaciones que deberían permitirse, la máxima superficie para la delimitación del núcleo puede ser calculada multiplicando la SPME por el total de edificaciones (actuales y futuros), con el objetivo de minimizar la modificación del entorno. Dependiendo del caso en el que nos encontremos, la superficie máxima del núcleo se calculará de diferente manera:

- Si el número de viviendas nuevas (VN) es o supera el 50 % del total de viviendas del núcleo, el objetivo es permitir la construcción, como máximo, de tantas futuras edificaciones como VN actualmente existan. Esto es, si hay 5 viviendas tradicionales (VT) y 10 VN, el núcleo debe tener espacio, como máximo, para 10 VN.
- Si el número de VN está entre el 25 % y el 50 % del total de viviendas, el objetivo será permitir, como máximo, la construcción de tantas edificaciones como VN haya (para el caso del NRC), o de la mitad de VN que existan (para el caso NRHT). Un ejemplo: si hay 8 VT y 12 VN en todo el núcleo, la nueva delimitación debe tener espacio para 8 VT más en NRC y para 4 VT más en el NRHT.
- Si el número de VN es menor del 25 % del total de viviendas del núcleo, el objetivo será permitir la construcción de, como máximo, un 50 % de las VN

existentes en el núcleo para el caso del NRC, un 25 % para el caso del NRHT, y un 50 % para el caso del NRHT cuando no existe NRC. Por ejemplo, si se tienen 10 VT y 2 VN, si existe NRHT y NRC, el núcleo debería tener espacio como máximo para 1 NV más para el caso del NRC (en el NRHT no se podría construir más). Por otro lado, si no existe NRC, se podría construir 1 NV en el NRHT.

Es importante que este criterio diferencie entre viviendas nuevas y viviendas tradicionales ya que la conservación de estas últimas es un objetivo prioritario de la política de desarrollo agrícola de la Unión Europea [63].

6.2. Etapa de preprocesado

Para formalizar el problema asumimos que la extensión máxima de un núcleo rural está previamente limitada por la selección de un conjunto de parcelas, de entre las que figuran en el Catastro, candidatas a formar parte del núcleo rural. Este conjunto de parcelas debería ser suficientemente grande como para incluir cualquier delimitación de tamaño razonable, aunque no muy grande como para que incremente demasiado el número de parcelas involucradas y por ende, el tiempo de computación. Entonces, dado un conjunto de parcelas catastrales, el problema consiste en asignar cada parcela a una de las dos categorías existentes (NRHT o NRC) o excluirla de la delimitación, maximizando una función de fitness definida y verificando las restricciones dadas. De esta manera, el número de posibles soluciones del problema es N^C , donde N es el número de parcelas y C el número de categorías (tres en este caso⁵).

La etapa de preprocesado se encarga pues de preparar la información que posteriormente utilizará el algoritmo durante su ejecución. Esto se hace en una etapa previa debido a que muchas de las operaciones realizadas son espaciales y consumen mucho tiempo y recursos computacionales. El beneficio es que con ejecutar una sola vez el preprocesado, el algoritmo podrá ser utilizado con diferentes parámetros de configuración sin necesidad de repetir la ejecución de las operaciones costosas.

⁵Se consideran tres categorías ya que la exclusión de una parcela de la delimitación es, a su vez, una categoría

El preprocesado también reduce el número de parcelas candidatas a formar parte del núcleo, es decir, aquellas que por su distancia a las edificaciones pueden pertenecer al núcleo. Las parcelas descartadas no serán utilizadas posteriormente por el algoritmo, reduciendo así el tamaño de los individuos y, en consecuencia, consumiendo menos recursos y reduciendo los tiempos de ejecución.

Al finalizar esta etapa se obtiene una posible solución al problema planteado, aunque será tarea del algoritmo principal validarla y utilizarla como base para la generación de los individuos de la población inicial.

Durante el preprocesado se seguirá la codificación de los individuos ya definida en la Sección 5.4.1.

6.2.1. Datos de entrada de la etapa de preprocesado

La Tabla 6.2 muestra los datos que obligatoriamente debe introducir el usuario para ejecutar la etapa de preprocesado. A parte de estos datos, también se puede proporcionar información, en formato *shapefile*⁶, sobre los usos del suelo de las parcelas, sus orientaciones, pendientes, carreteras de la zona, redes de abastecimiento y alcantarillado, elementos de iluminación, parques, jardines y áreas naturales. También, y meramente a nivel informativo y de clasificación, se puede indicar la condicionante espacial del núcleo (núcleo rural de montaña, de interior, de litoral o de proximidad a área urbana), su intensidad estructural de edificado (alto, medio, o bajo grado de edificación, o ninguno), y su estructura formal (círculo, cuadrado, rectángulo o irregular).

6.2.2. Datos de salida de la etapa de preprocesado

La Tabla 6.3 muestra los datos de salida que genera la etapa de preprocesado y que son necesarios para la posterior ejecución del algoritmo.

⁶Formato de archivo vectorial muy utilizado en el mundo SIG que, aunque es propiedad de la empresa ESRI, se ha convertido casi en un estándar.

Dato de entrada				Tipo	Descripción
Parcelario				Shapefile	Límites de las parcelas.
Edificaciones				Shapefile	Edificaciones y sus características.
Superficie NRHT	mínima	edificable		Doble	Superficie mínima que debe tener una parcela del NRHT para poder edificar en ella. El valor por defecto, y que marca la ley, es de 300m ² .
Superficie NRC	mínima	edificable		Doble	La superficie mínima que debe tener una parcela del NRC para poder edificar en ella. El valor por defecto, y que marca la ley, es de 600m ² .
Distancias DMC	incluidas	para la		Porcentaje	Porcentaje de distancias tenidas en cuenta para el cálculo de la DMC. Opciones: 25 %, 50 % y 75 %
Edificaciones tenidas en cuenta en el NRHT				Select	Tipo de edificación que se tendrán en cuenta para diferentes cálculos en la NRHT. Opciones: <i>viviendas tradicionales</i> y <i>edificaciones tradicionales</i> .
Edificaciones tenidas en cuenta en el NRC				Select	Tipo de edificación que se tendrán en cuenta para diferentes cálculos en la NRC. Opciones: <i>viviendas</i> y <i>edificaciones utilizadas para el cálculo de la consolidación</i> .
Superficie mínima de edificación en parcela				Doble	Superficie mínima que una parcela debe tener para considerar que una edificación pertenece a ella (ver Sección 6.2.3).

Tabla 6.2: Datos de entrada de la etapa de preprocesado para la delimitación de núcleos rurales.

Dato de salida		Tipo	Descripción
Datos de preprocesado		Serializable ⁷	Archivo con todos los datos necesarios para ejecutar el algoritmo.
Capa de preprocesado		Shapefile	Parcelario después de la ejecución de la etapa de preprocesado.
Carpeta para shapes		String	Ruta a la carpeta donde se guardarán los ficheros shapes generados y que son comentarios en las siguientes secciones.

Tabla 6.3: Datos de salida de la etapa de preprocesado para la delimitación de núcleos rurales.

6.2.3. Fases de la etapa de preprocesado

En los próximos apartados se describen las fases más importantes de la etapa de preprocesado. Desde ahora se empleará el término *edificaciones seleccionadas* para identificar el tipo de edificaciones escogidas en los parámetros de entrada *Edificaciones tenidas en cuenta en el NHRT* y *Edificaciones tenidas en cuenta en el NHC*, según el caso, de la Tabla 6.2.

Identificación de las parcelas y edificaciones

El algoritmo comienza leyendo todos los registros de los shapefiles indicados en los parámetros de entrada y los va insertando en estructuras de datos de acceso rápido como pueden ser las listas. Esto se hace porque como van a ser datos a los que habrá que acceder en más de una ocasión, es más rápido leerlos de una lista que directamente del shapefile.

La única información espacial que el algoritmo necesita es la que tiene que ver con las parcelas y las edificaciones. Como los datos disponibles normalmente no suelen ser muy precisos, puede ocurrir que una edificación esté en más de una parcela a la vez. Esto hace que el cálculo de, por ejemplo el grado de consolidación, proporcione resultados erróneos ya que contaría edificaciones que no deberían tenerse en cuenta. Para evitar esta situación se descartan aquellas edificaciones que no intersequen con ninguna parcela y, para las que lo hagan con más de una, se descartan aquellas cuya superficie de intersección sea menor a la especificada en el parámetro de entrada *Superficie mínima de edificación en parcela* de la Tabla 6.2. En caso de que ninguna intersección supere la superficie mínima, se considera que la parcela con más superficie de intersección será la parcela a la que pertenecerá la edificación.

El siguiente paso es la identificación de las parcelas que se corresponden con los viales. Estos vienen marcados en el shapefile de usos del suelo. Intersecando pues la capa de parcelas y la capa de usos se pueden identificar qué parcelas son en realidad viales. Como también en este caso puede haber imprecisiones entre ambas capas, se considera que una parcela es vial si la superficie de intersección entre ellas es mayor al 60 % de la superficie total de la parcela. En caso contrario, la parcela no se considerará como vial.

El motivo por el cual se identifican los viales es que estos suelen ser parcelas muy largas y muy estrechas, lo que hace que estén en contacto con otras muchas parcelas. Esta situación provoca que a la hora de calcular las relaciones de vecindad, dos parcelas muy alejadas entre sí puedan estar conectadas por los viales. Esta cercanía teórica no es real y produce resultados no satisfactorios. La solución propuesta es *cortar* los viales y convertirlos en segmentos más pequeños para que sólo puedan conectar parcelas muy próximas entre sí. Para realizar estos cortes se calcula un punto interior del vial y se trazan círculos concéntricos a él cada vez con mayor radio, de manera que entre los círculos vayan quedando trozos de vial más pequeños que serán los tenidos en cuenta.

Seguidamente a cada parcela se le asignan las características que le corresponden en base a los shapefiles adicionales que se indiquen en los parámetros de entrada. Esta información es la siguiente:

- **Orientación.** La orientación de una parcela puede tener los siguientes valores: *norte, sur, este, oeste o sin determinar*. Y como una parcela puede tener más de una, la orientación que mayor superficie tenga en la parcela será la que finalmente se le asigne.
- **Pendiente.** La pendiente se indica mediante rangos: *[0 %, 2 %), [2 %, 5 %), [5 %, 10 %), [10 %, 15 %), [15 %, 20 %), [20 %, 25 %), [25 %, 35 %), [35 %, 50 %), >50 %*. El rango predominante en una parcela será el asignado.
- **Usos.** En este caso se consideran todos los usos que tenga la parcela. Por muy poca superficie que estos usos ocupen, más adelante serán importantes a la hora de la evaluación de las soluciones.
- **Distancia a las avenidas, carreteras y plazas.**
- **Distancia a la red de saneamiento, de abastecimiento y a los puntos de iluminación.**
- **Distancia a parques, jardines y áreas naturales.**

Por otra parte, las edificaciones también juegan un papel fundamental a la hora de definir la delimitación de un núcleo. Las características analizadas son: si es

tradicional o nueva; si es vivienda o no; si se tiene en cuenta o no para el cálculo del grado de consolidación; su tipo habitacional; si es edificación histórica; si es edificación pública; si es edificación rural; y si tiene actividad rural.

Cálculo de las relaciones de vecindad

Para tener una visión espacial del núcleo se calculan todas las relaciones de vecindad entre las parcelas mediante la técnica explicada en la Sección 5.4.2.

Además de la vecindad de las parcelas entre sí, también se calcula la vecindad entre parcelas y viales. Se distinguirá pues entre parcelas vecinas directas a viales o vecinas indirectas (aquellas que son vecinas de alguna parcela vecina directa). Este cálculo sólo se realizará a parcelas cuya distancia a algún vial sea igual o menor a 10 metros.

Existencia del NRHT y su extensión máxima

Como se ha mencionada en la Sección 6.1.2, un núcleo rural no puede exceder de una cierta extensión máxima, que va a ser calculada conjuntamente con la existencia del NRHT y del NRC en esta etapa de preprocesado.

Para comprobar pues la existencia de un NRHT se crean buffers alrededor de las edificaciones seleccionadas. Cuando se utiliza la DMC, dichos buffers tienen de radio $1,5 * DMC$. A continuación todos esos buffers se unen creando así una nueva geometría compuesta por uno o más polígonos. Si alguno de los polígonos tiene más de dos viviendas tradicionales, se puede concluir que existe un NRHT. En caso contrario, el núcleo no tendría un área histórica-tradicional.

Si del paso anterior se concluye que existe un NRHT, se calcula cual es la superficie máxima que puede tener. Por la restricción de polígonos lo primero es eliminar los que tengan una o ninguna edificación seleccionada. Con los polígonos restantes se crea un nuevo buffer de $0,5 * DMC$ (o lo que es lo mismo, un buffer de $2 * DMC$ sobre las edificaciones seleccionadas). Lo siguiente es unir los polígonos resultantes que intersequen entre sí, y dejar con el tamaño anterior ($1,5 * DMC$) a los que no lo hagan.

En este punto la situación es la siguiente: algunas parcelas están fuera de los buffers, otras están totalmente dentro y otras sólo parcialmente dentro. Las que están sólo parcialmente incluidas han de tener al menos una vivienda dentro de los límites de los buffers para poder ser consideradas como parcela candidata, ya que si no la tienen serán excluidas.

Por otra parte, las parcelas que estén fuera de los buffers pero que tengan viviendas a menos de 10 metros de distancia de los mismos serán también incluidas en ellos. Para adherir dichas parcelas se crea un buffer de 15 metros alrededor de la vivienda en cuestión, se interseca con la parcela y el resultado se une al núcleo. Después de todas estas modificaciones el algoritmo comprueba otra vez que los polígonos cumplen todas las restricciones.

Si en vez de la DMC se utilizan los 50 metros que marca la ley, el proceso seguido es igual salvo que en vez de crear el primer buffer con radio $1,5 * DMC$ se creará con radio de 50 metros y sin crear un segundo buffer.

La Figura 6.1 muestra todos estos pasos aplicados al núcleo rural de Corredoiros perteneciente al municipio gallego de Guitiriz. En la primera imagen se puede observar el buffer de $1,5 * DMC$ alrededor de las edificaciones seleccionadas. Como existen algunos polígonos con dos o más edificaciones seleccionadas en su interior, el núcleo podría tener una parte histórica-tradicional. El polígono con una sola edificación seleccionada es rechazado y ya no aparece su buffer en la imagen 2. En esta imagen se puede ver cómo a los dos polígonos que se conservan se les añade un buffer de $0,5 * DMC$ y se juntan en uno solo que será el utilizado en la delimitación final del núcleo, como muestra la imagen 3.

Existencia del NRC y su extensión máxima

Para comprobar la existencia del NRC se contemplan dos escenarios: uno en el que exista un NRHT y otro en el que no. En el primer escenario, la existencia del NRC dependerá de las edificaciones que estén dentro de las parcelas asignadas a la categoría NRHT, por tanto, será el algoritmo quien se encargue de esta comprobación, ya que se necesita conocer el NRHT real en cada momento. En el segundo escenario se crea un buffer sobre cada edificación seleccionada de $2 * DMC$ y, si hay algún polígono con al menos dos viviendas dentro del buffer, concluimos que existe

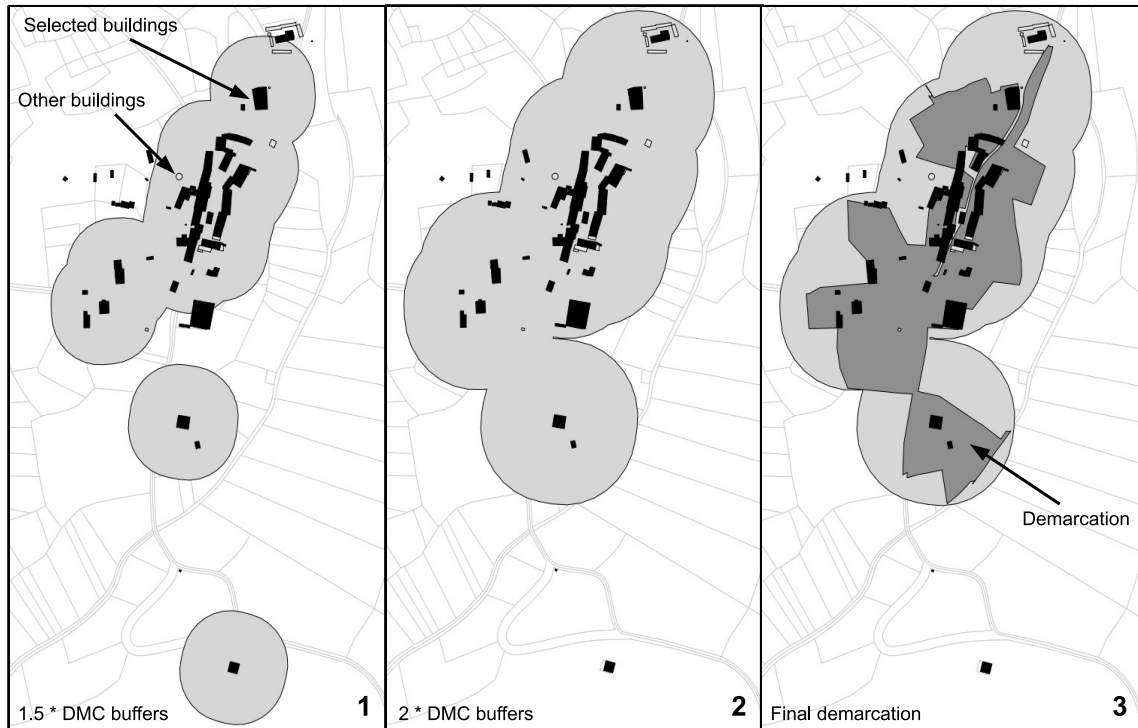


Figura 6.1: Buffers alrededor de las edificaciones seleccionadas y la delimitación final.

el NRC.

Si de lo anterior se puede concluir que existe NRC, su superficie máxima será calculada como se hizo con el NRHT pero utilizando la DMC del NRC en vez la del NRHT.

Resumiendo, en los dos últimos apartados se han calculado tres nuevas geometrías: la delimitación máxima del NRHT utilizando la DMC, la delimitación máxima del NRHT utilizando la ley, y la delimitación máxima del NRC. Estas geometrías serán guardadas en la carpeta de shapes (ver Tabla 6.3).

De ahora en adelante la delimitación máxima del NRHT se denotará como *Máx-NRHT* y la delimitación máxima del NRC como *Máx-NRC*. La unión de las parcelas de Máx-NRHT con las parcelas del Máx-NRC que no estén en Máx-NRHT, será la delimitación máxima del núcleo y será llamada *Máx-NR*.

Identificación de las parcelas candidatas

Una vez definida la superficie máxima total de las parcelas asignadas a las dos categorías rurales se puede saber qué parcelas son candidatas a formar parte del NRHT y del NRC. El resto de parcelas ya no serán tenidas en cuenta en el algoritmo.

Para calcular pues las parcelas candidatas se interseca cada parcela con las delimitaciones máximas anteriormente calculadas. Las que tengan una superficie de intersección mayor que 0, serán las candidatas a formar parte del núcleo.

Es importante destacar que para los cálculos que se harán posteriormente en el algoritmo, sobre todo para la comprobación de la restricción de la superficie máxima del núcleo, es necesario conocer qué superficie de las parcelas forma parte del núcleo, ya que, como se ha comentado anteriormente, algunas parcelas pueden estar incluidas sólo parcialmente.

Creación de los individuos con máxima extensión

La fase final del preprocesado es crear dos soluciones con las parcelas candidatas calculadas en el punto anterior. Una solución corresponderá a la aplicación de la DMC y la otra a la aplicación de los 50 metros marcados por la ley. Estas soluciones serán la base para la construcción de la población inicial en el algoritmo, y además, también tendrán una importancia por sí mismas debido a cómo han sido obtenidas, aspecto que se explicará en las siguientes secciones.

6.3. Algoritmo PBIG para la delimitación de núcleos rurales

Como ya se ha comentado en la Sección 5.3, la primera intención era utilizar un algoritmo genético para resolver este problema. Estudiando las características del mismo y al comenzar el diseño del algoritmo se observó que el operador de cruce tradicional utilizado en la mayoría de los algoritmos genéticos, en este problema en particular no tenía sentido. El motivo es que este operador no tiene en cuenta las relaciones de vecindad entre parcelas y, al cambiar las parcelas de categoría sin

importar su localización, se generan muchas soluciones no válidas, lo que retrasa la convergencia del algoritmo. Suprimiendo el operador de cruce, el de selección ya no es necesario y sólo quedaría el operador de mutación, que sería aplicado a todos los individuos de la población. Vista la situación se optó por aplicar una estrategia basada en algoritmos *greedy* iterativos basados en población (PBIG) que, como también se comenta en la Sección 5.3, se ajustan perfectamente a las necesidades de nuestro problema.

Por tanto, el algoritmo PGIB es el encargado de encontrar una solución satisfactoria al problema de delimitación de núcleos rurales. Esto consiste en asignar a cada parcela un valor que indique si pertenece al NRHT, al NRC o a ninguno de los dos. En las siguientes subsecciones se describen los datos de entrada y salida del algoritmo, las clases utilizadas y sus relaciones, el cálculo de las aptitudes de las parcelas, y la validación y evaluación de los individuos. Posteriormente, la Sección 6.4 está dedicada a la descripción de las fases por las que pasa el algoritmo PBIG.

6.3.1. Datos de entrada del algoritmo PBIG

La Tabla 6.4 muestra qué datos se necesitan para la ejecución del algoritmo PBIG. Son todos obligatorios.

Dato de entrada	Tipo	Descripción
Tamaño población	Entero	Número de individuos de la población.
Número máximo de iteraciones	Entero	Número de iteraciones que el algoritmo ejecutará.
Distancia a utilizar	Select	Distancia utilizada para los cálculos. Opciones: <i>DMC</i> o <i>50 metros de la ley 2/2010 del suelo</i> .
Densidad de edificación del NRHT	Doble	Mínima densidad de edificación para el NRHT. Rango $[0, 1]$. Valor por defecto: <i>0,5</i> . Restricciones: sólo disponible cuando se seleccione la <i>DMC</i> como la <i>distancia a utilizar</i> .

Dato de entrada	Tipo	Descripción
Densidad de edificación del NRC	Doble	Mínima densidad de edificación para el NRHT. Rango $[0,1]$. Valor por defecto: $0,33$. Restricciones: sólo disponible cuando se seleccione la DMC como la <i>distancia a utilizar</i> .
Método para el cálculo de la densidad de edificación	String	Método utilizado para el cálculo de la densidad de edificación. Opciones: <i>método de la APLU</i> o <i>método basado en la relación entre parcelas</i> (sólo disponible cuando se seleccione la DMC como la <i>distancia a utilizar</i>).
Aptitud	Doble	Peso para la aptitud en la función objetivo. Rango $[0,1]$.
Ratio de edificaciones en el área tradicional	Doble	Peso para la relación entre el número de edificaciones del área tradicional de la solución y el número de edificaciones del Máx-NRHT. Rango $[0,1]$.
Ratio de todas las edificaciones	Doble	Peso para la relación entre el número de edificaciones de la solución y el número de edificaciones del Máx-NR. Rango $[0,1]$.
Ratio de superficies del área tradicional	Doble	Peso para la relación entre la superficie del área tradicional de la solución y la superficie del Máx-NRHT. Rango $[0,1]$.
Ratio de toda la superficie	Doble	Peso para la relación entre la superficie de la solución y la superficie del Máx-NR. Rango $[0,1]$.
Compacidad	Doble	Peso para la compacidad en la función objetivo. Rango $[0,1]$.
Recortar solución	Boolean	Si el usuario escoge <i>Sí</i> , la solución obtenida se recortará como se indica en 6.4.4. Si el usuario escoge <i>No</i> , la solución no se recortará y la delimitación tendrá los límites de las parcelas.

Dato de entrada	Tipo	Descripción
Pesos de criterios para NRTH	String	Ruta al archivo con los pesos para las características de las parcelas del NRHT.
Pesos de criterios para NRC	String	Ruta al archivo con los pesos para las características de las parcelas del NRC.
Datos del preprocesado	Serializable	Archivo serializado con los datos obtenidos de la etapa del preprocesado.

Tabla 6.4: Datos de entrada del algoritmo PBIG de delimitación de núcleos rurales.

6.3.2. Datos de salida del algoritmo PBIG

La Tabla 6.5 muestra los datos de salida que produce el algoritmo PBIG.

Dato de salida	Tipo	Descripción
Aptitudes	Shapefile	Aptitudes edificatorias de cada parcela, calculadas como se describe en 6.3.4
Archivo con la solución	Serializable	Archivo serializable para poder recuperar la solución en cualquier momento.
Solución	Shapefile	Delimitación del núcleo rural, diferenciando el NRHT y el NRC.

Tabla 6.5: Datos de salida del algoritmo PBIG para la delimitación de núcleos rurales.

6.3.3. Descripción de las clases

Como se ha comentado anteriormente en la Sección 5.5, hoy en día Java es comparable a lenguajes como C o Fortran en cuanto a velocidad de ejecución, sobre

todo cuando se utilizan tipos de datos básicos. Es por esto que en el preprocesado, una fase bastante costosa computacionalmente, se utilizan esos tipos de datos con el objetivo de ganar en eficiencia y ahorrar en uso de memoria y en tiempo de ejecución. La idea inicial cuando se encaró el diseño del algoritmo PBIG fue seguir esta estrategia de codificación basada en arrays, pero como se observó que no iba a ser crítico en cuanto a tiempo (las ejecuciones hasta llegar a una solución aceptable no son largas) ni tampoco en cuanto al uso de memoria (al ser núcleos rurales con un número de parcelas relativamente pequeño, no se necesita mucha cantidad de memoria), y unido esto a la dificultad propia del problema (ya que hay que trabajar con muchos grupos de parcelas), se decidió utilizar las clases Java representadas en la Figura 6.2. Justo lo contrario ocurre con el problema de ordenación del territorio resuelto por algoritmo genético explicado en el Capítulo 7: al ser municipal su ámbito de aplicación, trabaja con tantas parcelas que es preciso utilizar óptimamente los recursos, haciendo inviable la utilización de objetos.

La Figura 6.2 muestra que un individuo (clase *Individual*) contiene todas las parcelas que se han marcado como candidatas para formar parte del núcleo rural en la etapa de preprocesado. Cada parcela perteneciente al individuo (*CategoryPlot*) guarda información de los datos de la parcela (*Plot*) y de la categoría asignada en ese instante a esa parcela. También tiene referencias a un núcleo tradicional (*TraditionalArea*) y un núcleo común (*CommonArea*), ambos heredados de la clase *Area*, que tiene a su vez referencias a parcelas, edificaciones (cada edificación es una clase *Building*) y polígonos.

Un polígono (*Polygon*) también guarda referencias al área donde está y a las parcelas y edificaciones que contiene. También mantiene información de las parcelas que forman la frontera (*Border*) del polígono, que puede ser de dos tipos: interna o externa. Por ejemplo, la frontera interna de un polígono que forme parte del NRHT contiene las parcelas que pertenezcan al NRHT pero que también tengan alguna vecina que no forme parte del NRHT. Por otro lado, una parcela formará parte de la frontera externa de un NRHT cuando dicha parcela no pertenezca a este pero tenga alguna vecina que sí pertenezca. Por su parte, la frontera interna de un polígono del NRC está formada por parcelas que formen parte del NRC pero que tengan alguna vecina que no esté asignada a ninguna categoría. Y por último, formarán parte de la frontera externa del NRC aquellas parcelas no asignadas a ninguna categoría pero

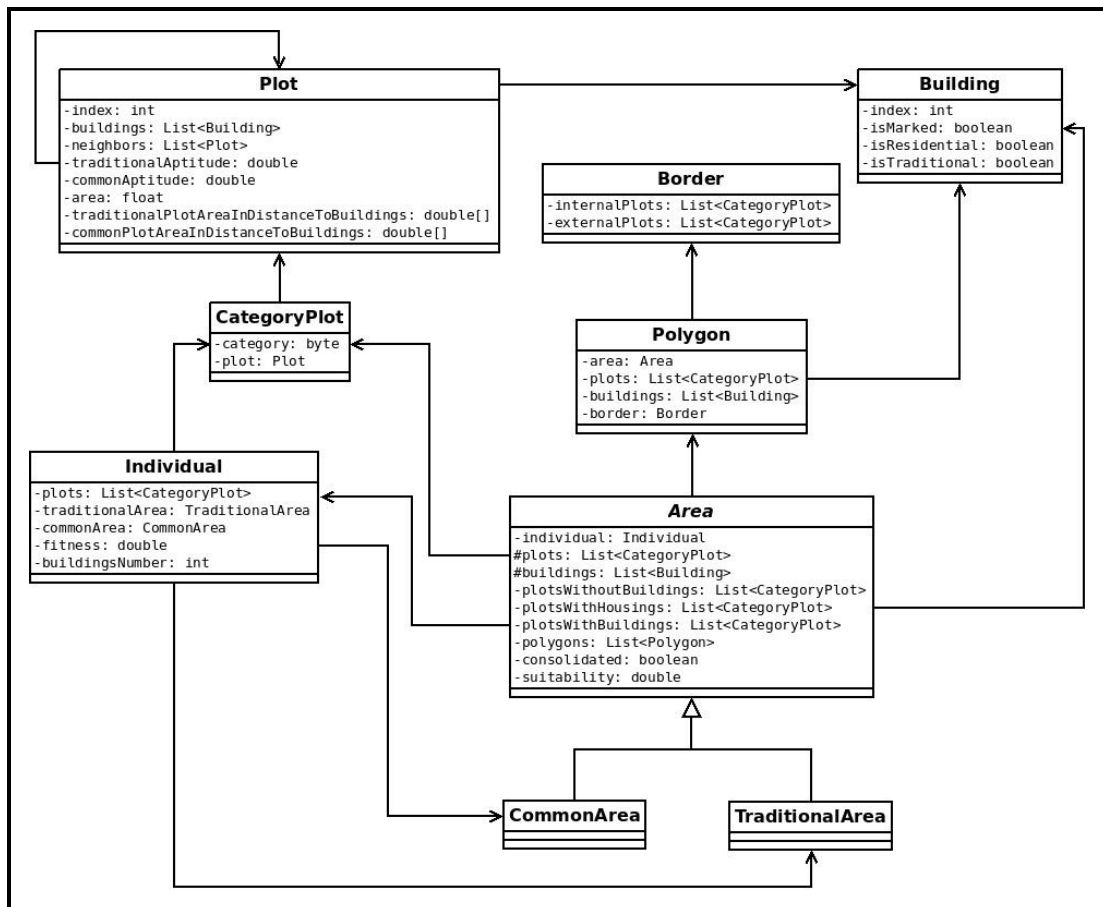


Figura 6.2: Diagrama de clases utilizadas en el algoritmo PBIG de la delimitación de núcleos rurales.

que tengan alguna vecina asignada al NRC.

6.3.4. Cálculo de las aptitudes de las parcelas

Dependiendo de las particularidades de las parcelas y de las características de sus edificaciones, algunas parcelas pueden ser más aptas para formar parte del NRHT, del NRC o incluso de ninguno en particular. Las propiedades escogidas, previamente agrupadas por unidades territoriales y colectivas [109], para el cálculo de las aptitudes de las parcelas son: la habitabilidad de sus edificaciones, si es pública o si es rural; sus usos; las infraestructuras cercanas, calles, plazas, saneamiento, abastecimiento, iluminación, parques y áreas de recreo; la pendiente del terreno e incluso su orientación. Cada una de estas características tiene asignada dos pesos, uno para el NRHT y otro para el NRC, acordes a su importancia y que han sido proporcionados por expertos en planificación [14]. Estos pesos se guardan en sendos archivos que se leerán desde el algoritmo en la ruta especificada en los parámetros de entrada.

6.3.5. Validación de individuos

Cada individuo representa una solución válida al problema planteado, y se entiende por solución válida a la solución que cumpla todas las restricciones.

El algoritmo PBIG adapta las restricciones explicadas en las secciones 6.1.1 y 6.1.2 de la siguiente manera:

- Restricción de área: el área de la solución ha de ser menor o igual que el área máxima del núcleo.
- Restricción de polígonos: cada polígono debe tener un número mínimo de viviendas.
- Densidad de edificación: el núcleo rural tiene que ser consolidado.

Otras restricciones como la DMC o los 50 metros definidos por ley son tenidas en cuenta implícitamente durante todo el proceso de delimitación, incluido el preprocesado.

6.3.6. Evaluación de individuos

La evaluación de un individuo se basa en la aplicación de una función objetivo que proporciona un valor relacionado con sus características. Cuanto mayor sea ese valor, mejor será el individuo. La función objetivo consta de seis variables cada una con un peso especificado por el usuario como parámetro de entrada. Estas variables son:

- Aptitud (APT). Como se ha comentado en 6.3.4, cada parcela posee un valor de aptitud que viene determinado como la media de las aptitudes de sus parcelas.
- Compacidad (COMP). La compacidad es una medida que indica la regularidad de la forma geométrica del núcleo rural. Es preferible por ejemplo un núcleo con una forma similar a la circular o a la cuadrada que rectangular con tamaños de lados muy desiguales. Para el cálculo de la compacidad se aplica la siguiente ecuación, que está basada en la circularidad [94], esto es, cuanto más se parezca la forma del núcleo a un círculo, mayor será su valor:

$$COMP = 4\Pi \frac{\sum_{i=0}^{NP_{NRHT}} \frac{area_i}{perimetro_i^2}}{NP_{NRHT}} + 4\Pi \frac{\sum_{i=0}^{NP_{NRC}} \frac{area_i}{perimetro_i^2}}{NP_{NRC}} \quad (6.7)$$

donde $area_i$ y $perimetro_i$ son la superficie y el perímetro del i -ésimo polígono asignado a una de las categoría respectivamente, y NP_{NRHT} y NP_{NRC} son el número de polígonos del NRHT y del NRC respectivamente.

- Edificaciones en el NRHT (NRHT_EDIF). Relación entre el número de edificaciones en el NRHT del individuo y el número de edificaciones en el NRHT del Máx-NRHT.
- Edificaciones en el NR (NR_EDIF). Relación entre el número de edificaciones en todo el NR del individuo y el número de edificaciones en el NR del Máx-NR.
- Superficie del NRHT (NRHT_SUP). Relación entre la superficie del NRHT en el individuo y la superficie del NRHT del Máx-NRHT.
- Superficie total del NR (NR_SUP). Relación entre la superficie total del NR en el individuo y la superficie del NR del Máx-NR.

Por tanto, la función objetivo (FO) queda definida como:

$$FO = w_1 * APT + w_2 * COMP + w_3 * NRHT_EDIF + w_4 * NR_EDIF + w_5 * NRHT_SUP + w_6 * NR_SUP \quad (6.8)$$

siendo w_i , con $\sum_{i=1}^6 w_i = 1$, los pesos de cada una de las 6 variables.

6.4. Fases del algoritmo PBIG para la delimitación de núcleos rurales

El algoritmo PBIG consta de una serie de fases que harán que, a partir de los datos obtenidos en el preprocesado y teniendo siempre en cuenta las restricciones del problema, los individuos de la población evolucionen creando a su vez nuevos individuos. De estos sólo los que mejor valoración obtengan en el cálculo de la función de evaluación formarán parte de las nuevas poblaciones y, por tanto, seguirán evolucionando. En las siguientes secciones se detallan estas fases, pero antes la Tabla 6.6 explica el significado de las variables y funciones utilizadas en el pseudocódigo del algoritmo PBIG explicado en la Figura 6.3.

6.4.1. Validación del Máx-NR

Si el Máx-NR (individuo que representa la delimitación máxima del NR) calculado en la etapa de preprocesado cumple todas las restricciones (línea 1), este se guarda en un archivo ya que puede resultar de utilidad en estudios posteriores a la ejecución del algoritmo. Hay que tener en cuenta que esta solución es teóricamente bastante compacta, ya que está basada en los buffers creados alrededor de las edificaciones. No obstante, si el Máx-NR no cumple alguna de las restricciones, se actuará de diferente manera dependiendo de cuál no cumpla. Si no cumple la restricción de polígono, se procede a sacar del núcleo los polígonos conflictivos (líneas 2-6). Si con estos cambios resulta que el Máx-NR es válido, la solución devuelta por el algoritmo será el Máx-NR sin los polígonos borrados (línea 7). En caso de que sean otras las

Nombre	Significado
P	Población.
M	Tamaño de la población (número de individuos).
P_i	i -ésimo individuo de la población.
$A-P_i$	Área del P_i asignado a la categoría A , donde A es <i>NRHT</i> o <i>NRC</i> .
$A_{int}-P_i$	Borde interno de la categoría A de P_i , donde A es <i>NRHT</i> o <i>NRC</i> .
$A_{ext}-P_i$	Borde externo de la categoría A de P_i , donde A es <i>NRHT</i> o <i>NRC</i> .
$pol(X)$	Función con la que se obtienen los polígonos de un conjunto de parcelas X .
$f(P_i)$	Función de fitness.
$add(A-P_i)$	Función que añade una parcela aleatoria del borde externo de $A-P_i$ al borde interno de $A-P_i$.
$remove(A-P_i)$	Función que borra una parcela aleatoria del borde interno de $A-P_i$.
$flip()$	Devuelve <i>verdadero</i> o <i>falso</i> con probabilidad de 50 % cada una.

Tabla 6.6: Significados de las variables y funciones del pseudocódigo del algoritmo PBIG para la delimitación de núcleos rurales.

restricciones que no se cumplen, se pasará a las siguiente fase.

6.4.2. Inicialización de la población

El algoritmo utiliza el Máx-NR como base para crear los individuos de la población inicial (línea 8). Utilizar una solución como base para la creación de nuevos individuos tiene dos ventajas: la población inicial se genera de manera más rápida, y los nuevos individuos pueden aprovecharse de que la solución base ya es relativamente buena y, al menos teóricamente, los individuos obtenidos a partir de ella serán mejores que los que se puedan generar de manera aleatoria [25].

El proceso de creación de estos individuos es el siguiente:

- Si el Máx-NR tiene NRHT pero no es consolidado, el algoritmo borra una parcela aleatoria de los polígonos no consolidados pero siguiendo este orden (mientras existan parcelas): parcelas sin edificación, parcelas con edificaciones que no sean viviendas, parcelas con viviendas. Esto se hace hasta que todos los polígonos del NRHT queden consolidados, o hasta que ya no quede ningún

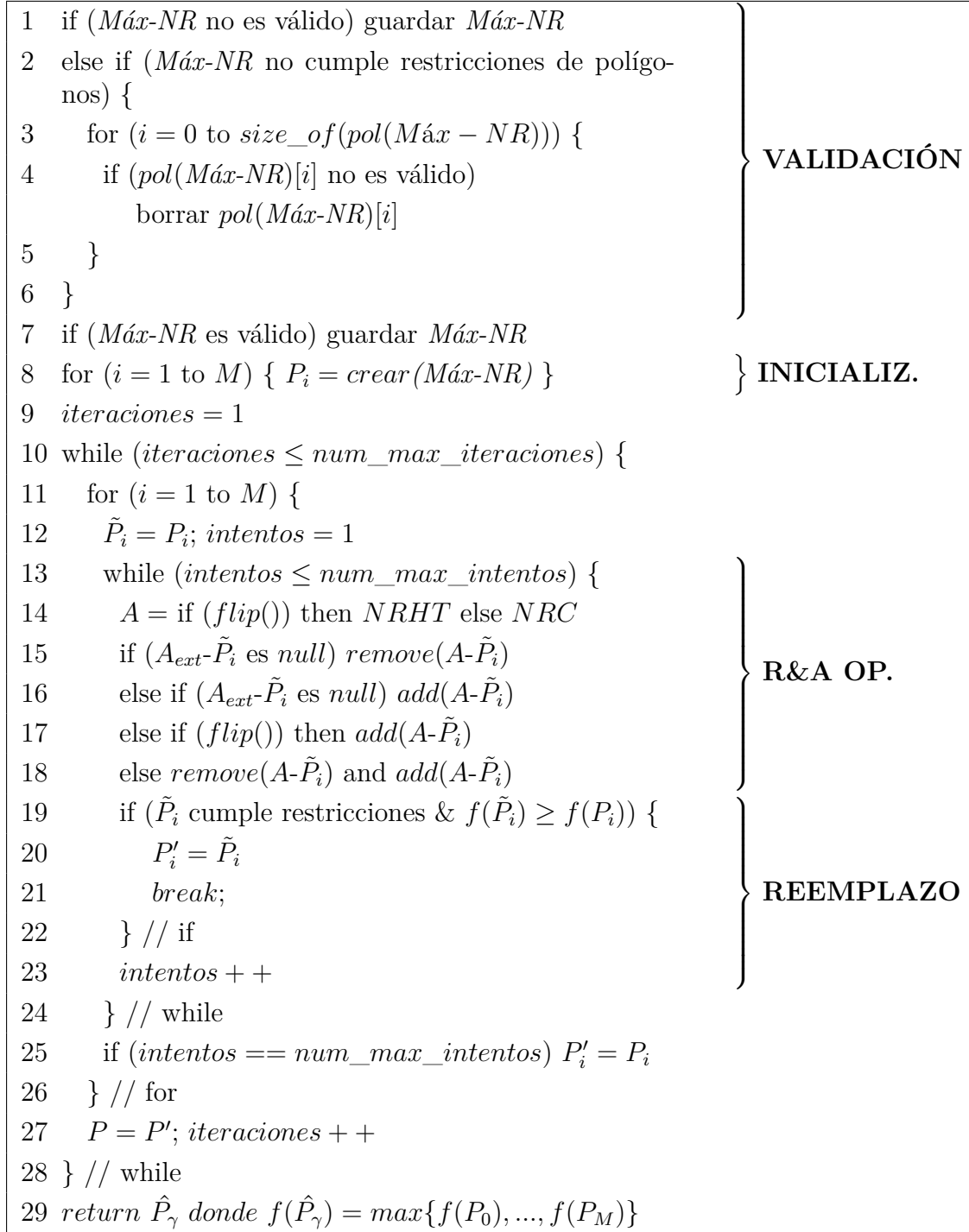


Figura 6.3: Pseudocódigo del algoritmo PBIG para la delimitación de núcleos rurales.

polígono en el NRHT.

- Si el Máx-NR, con las anteriores modificaciones, tiene NRC pero no es consolidado, el algoritmo hace lo mismo que en el supuesto anterior pero borrando parcelas de los polígonos no consolidados del NRC.
- Si el Máx-NR, con las anteriores modificaciones, excede la superficie máxima, el algoritmo borra una parcela aleatoriamente pero siguiendo el siguiente orden (mientras existan parcelas): parcelas sin edificaciones en NRC, parcelas sin edificaciones en NRHT, parcelas con edificaciones que no sean viviendas en NRC, parcelas con edificaciones que no sean viviendas en NRHT, parcelas con viviendas en NRC, parcelas con viviendas en NRHT. Se parará cuando la restricción de superficie sea satisfecha.
- Si el resultado de todas estas modificaciones no cumple el mínimo número de edificaciones por polígono, el algoritmo borra los polígonos conflictivos.
- En este paso todas las restricciones son comprobadas otra vez. Si alguna de ellas falla, el individuo es rechazado y se vuelven a ejecutar los pasos para crear otro. Si todas las restricciones son satisfechas, el individuo pasa a formar parte de la población inicial.

En el hipotético caso de que esta fase no sea capaz de crear suficientes individuos para completar la población inicial, el algoritmo se detendrá después de un número determinado de intentos e informará al usuario de tal circunstancia.

6.4.3. Operador *borrar-y-añadir* (*remove-and-add*, *R&A*)

El operador R&A es aplicado a cada uno de los individuos de la población y consiste en primero, elegir una de las categorías (NRHT o NRC) aleatoriamente y con una probabilidad del 50 % cada una (línea 14 en Figura 6.3). Si casualmente el individuo no dispone de la categoría elegida el individuo no sufre ninguna modificación. Si la categoría seleccionada no tiene parcelas en su borde externo, pero sí en el interno, el algoritmo borra parcelas de dicho borde (línea 15). Si la categoría seleccionada no tiene parcelas en su borde interno, se añadirán del externo (línea

16). Por otra parte, si el núcleo tiene ambos bordes, el algoritmo escoge aleatoriamente y con probabilidad del 50 % entre: añadir una parcela del borde externo al interno (línea 17), o borrar una parcela del borde interno y luego añadir otra del borde externo al borde interno (línea 18). Este reemplazo se realiza con el objetivo de maximizar la superficie de la delimitación y el número de edificaciones que hay en ella. Hay que recordar también que la parcela añadida tiene que tener una edificación lo suficientemente cerca como para satisfacer la restricción de distancia comentada a lo largo de la Sección 6.1.

La Figura 6.4 muestra dos operaciones R&A ejercidas sobre un individuo en 5 imágenes nombradas de la A a la E. El individuo original es representado en la imagen A. La primera operación ocurre en B: la categoría seleccionada es la NRHT y la acción consiste en añadir una parcela del borde externo al interno de la NRHT. En la imagen C la parcela seleccionada ya está asignada a la categoría NRHT. La segunda operación consiste en borrar y añadir: la parcela pintada en negro de la imagen C es borrada del NRHT y la parcela pintada en negro de la imagen D es asignada al NRHT. Nótese que esta última parcela es extraída del NRC. Sin embargo al NRC no se le permite extraer parcelas del NRHT, sólo podrá incorporar parcelas vecinas que no estén asignadas a ninguna categoría, es decir, de su borde exterior.

Una vez las operaciones son realizadas, el algoritmo comprueba si el nuevo individuo satisface todas las restricciones y si el valor de su función de fitness es mayor que el valor del individuo original. Si ambas condiciones son ciertas, el nuevo individuo es incluido en la nueva generación (etapa de reemplazo, líneas 19 a 22), si no, el individuo es modificado otra vez. Si varias iteraciones después el individuo aún no satisface las condiciones, el individuo original será el que forme parte de la nueva generación (línea 27).

6.4.4. Obtención de la solución final

La solución final será el individuo de la última generación con el valor de fitness más alto. La solución, representada por las clases definidas en la Sección 6.3.3, es serializada y guardada en un archivo. También se crea un archivo shapefile con las categorías asignadas a las parcelas.

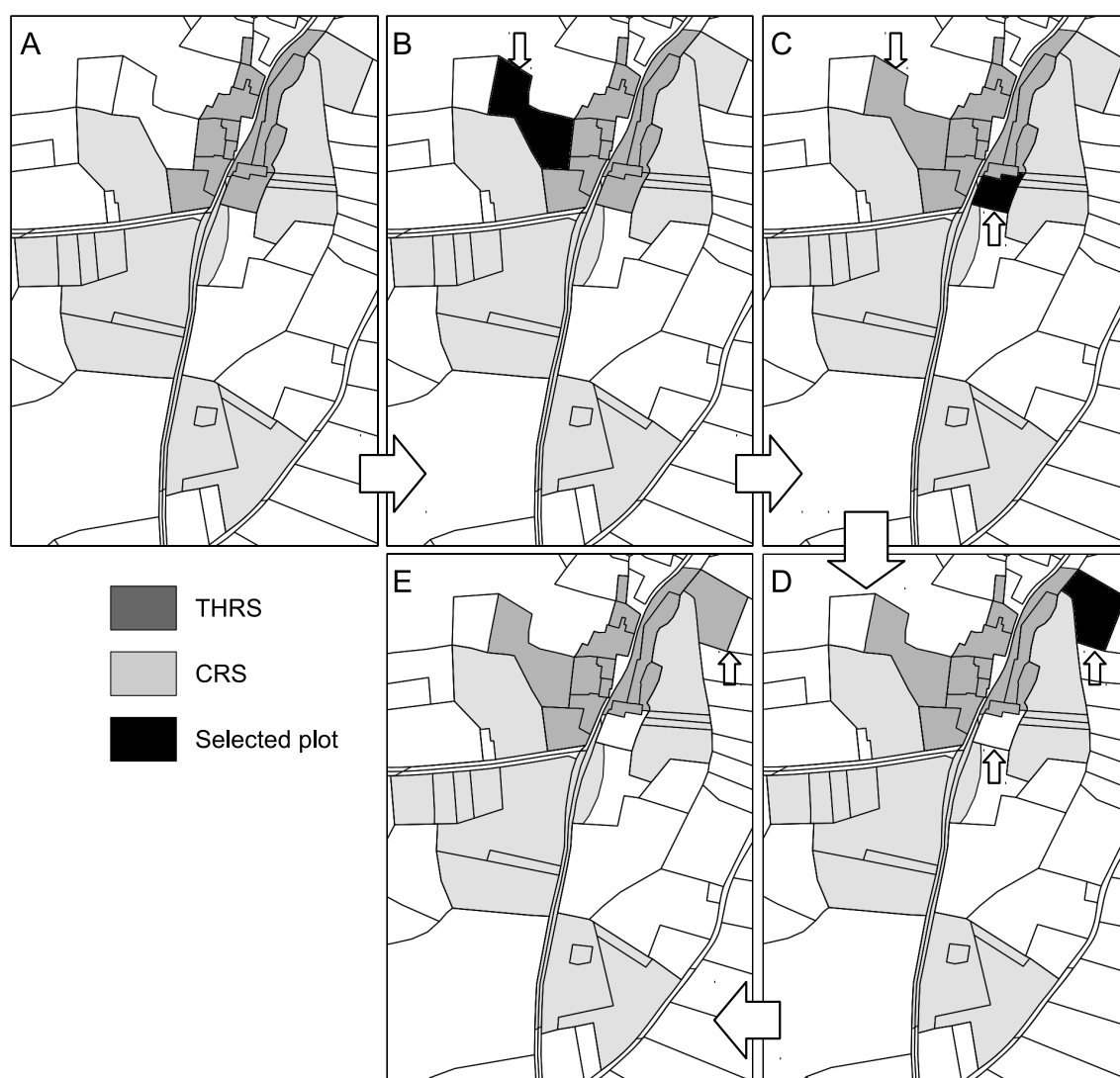


Figura 6.4: Operador R&A aplicado a un individuo (dos ejecuciones consecutivas).

Para transformar la solución de objetos a geometrías, existen dos opciones dependiendo de lo indicado en el parámetro *Recortar solución* de la Tabla 6.4. Si se especificó *No*, las geometrías de las categorías serán el resultado de la unión de las parcelas que formen parte de dichas categorías. En cambio, si se especificó *Sí*, habrá que recortar las geometrías por los buffers calculados en el preprocesado, siguiendo los siguientes pasos:

1. Para cada parcela se identifica el buffer con el que fue seleccionada como candidata (cálculo realizado en el preprocesado) y se interseca con las edificaciones que están contenidas en la misma categoría que la parcela.
2. Se realiza la unión de las intersecciones.
3. Si la parcela en cuestión está totalmente contenida en la unión, se mantiene la parcela entera.
4. Si la unión corta a la parcela por algún punto, sólo se mantendrá la parte de la parcela que queda solapada con la unión.
5. Si resultase que después de hacer estos cálculos el trozo de parcela solapado por la unión tuviese varias geometrías, se eliminarían aquellas que no tuviesen ninguna edificación.
6. Finalmente se unen todas las parcelas y los trozos de ellas pertenecientes a las categorías.

6.5. Casos de estudio

Galicia tiene casi la mitad de los asentamientos rurales localizados en España, en total 30.091. El 89 % de los asentamientos gallegos tienen una población menor de 100 habitantes y el 16 % de la población gallega viven en asentamientos con menos de 10 habitantes [51]. Estos datos explican la importancia que tiene en esta región la buena definición de planes de delimitación de núcleos urbanos y la multitud de aspectos problemáticos asociados a estos procesos.

Para la evaluación del algoritmo se han escogido tres asentamientos rurales del municipio gallego de Guitiriz: Viladonga, Ferreira y Saa, ya que cuentan con diferentes características morfológicas y patrones espaciales. El algoritmo fue ejecutado siete veces para cada asentamiento: en una ejecución se utilizaron los criterios marcados por las actuales leyes de planeamiento, y en el resto de las ejecuciones se utilizaron diferentes criterios alternativos especificados por la nueva metodología propuesta. La configuración de cada ejecución es resumida en la Tabla 6.7. Las soluciones fueron evaluadas acorde con los valores obtenidos por la función de fitness definida en la Sección 6.3.6. Los pesos utilizados para las seis variables de esta función fueron: 0,3 para APT, 0,2 para COMP, 0,2 para NR_SUP, 0,1 para NRHT_SUP, 0,1 para NRHT_EDIF, 0,1 para NR_EDIF, todas ellas basadas en estudios previos [14].

6.5.1. Núcleo rural de Viladonga

El asentamiento rural de Viladonga cuenta con 536 parcelas y está caracterizado por tener un patrón espacial mononuclear con edificaciones agrupadas y por estar organizado a lo largo de una red de carreteras secundarias cercanas a la red principal. Los resultados obtenidos para este asentamiento, plasmados en la Figura 6.5, muestran que la solución de demarcación siguiendo los criterios que marca la ley alcanza un valor de 0,56 y la mejor solución obtenida con el métodos alternativo (B1) alcanza un valor de 0,59. Ambas soluciones asignan sólo la categoría NRHT. En la misma figura también se pueden comparar los valores que adquieren los distintos componentes de la función objetivo (ver desglose de cada una de las barras de la Figura 6.5). De esta manera se puede comprobar cómo las principales diferencias entre ambas soluciones radican en los valores de las variables NR_EDIF y NR_SUP⁸. Como se puede observar en la Figura 6.6, la solución acorde a la ley excluye una vivienda tradicional más que la solución B1, y asigna menos superficie para el futuro crecimiento del asentamiento. Por su parte, la solución B1 muestra una mejor adaptación a las características morfológicas del asentamiento ya que la demarcación que obtiene incluye un mayor número de viviendas tradicionales y parcelas vacías. Este hecho es debido a que sólo se utiliza el 25 % de las distancias más cortas entre edificios para el cálculo de la DMC. El valor de la DMC para B1 es

⁸En la gráfica se corresponden con BUILD_RS y AREA_RS respectivamente, por sus nombres en inglés

Solución	Distancias	Tipo de edificaciones		Cálculo de la DMC
		NRHT	NRC	
Por ley	50 metros	ET	V	Método APLU
A1	DMC (25 %)	VT	V	Método alternativo
A2	DMC (50 %)	VT	V	Método alternativo
A3	DMC (75 %)	VT	V	Método alternativo
B1	DMC (25 %)	ET	V	Método alternativo
B2	DMC (50 %)	ET	V	Método alternativo
B3	DMC (75 %)	ET	V	Método alternativo

Tabla 6.7: Configuración de las ejecuciones en los casos de prueba del PBIG para la demarcación de núcleos rurales.

31.53 metros, que es más pequeña que los 50 metros que marca la ley, lo que permite una demarcación más compacta alrededor de las edificaciones del asentamiento. A medida que el porcentaje de distancias utilizadas crece, el valor de la DMC también lo hace (31,53 metros para B1, 48,69 metros para B2 y 64,94 metros para B3), lo que hace disminuir el valor total de la función de fitness. De estos datos podemos concluir que la posibilidad de modificar los valores de los parámetros utilizados en la metodología propuesta, permite representar mejor los patrones espaciales de los asentamientos en la solución final y, por tanto, conseguir demarcaciones adaptadas a las características de los asentamientos.

6.5.2. Núcleo rural de Ferreira

El asentamiento rural de Ferreira tiene 979 parcelas con un patrón espacial lineal con edificaciones distribuidas a lo largo de una carretera. Las viviendas nuevas están localizadas en los límites del asentamiento, principalmente en el norte, mientras que el resto de la superficie está ocupada por viviendas tradicionales. Los datos que aporta la Figura 6.7 muestran que la solución obtenida con los criterios legales tiene un valor de fitness total de 0,48, mientras que casi todas las soluciones obtenidas mediante la metodología propuesta tienen un valor más alto (la B3 es la única más baja con un 0,47), siendo la solución A2 la mejor con un valor de 0,54. Comparando

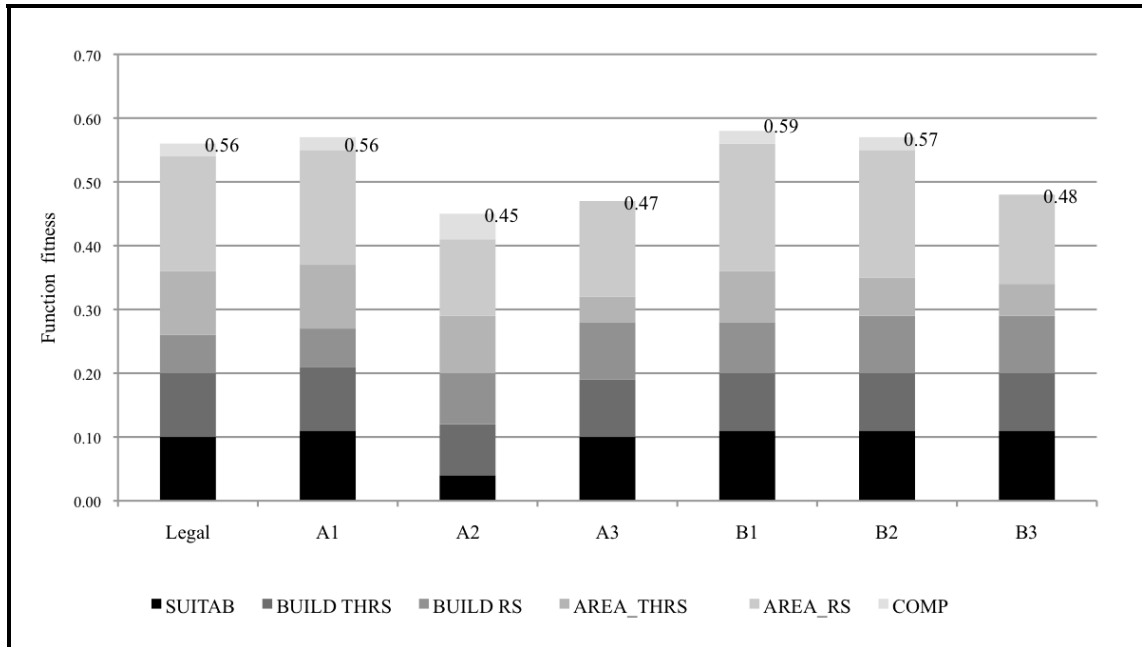


Figura 6.5: Valores de fitness de las ejecuciones del algoritmo PBIG sobre el asentamiento rural de Viladonega.

la solución acorde a la ley y la A2, las principales diferencias se dan en las variables de compacidad, aptitud y área total del núcleo⁹. La Figura 6.8 muestra que la solución acorde con la ley asigna el asentamiento entero a la categoría NRHT, mientras que la A2 lo divide en NRHT y NRC, lo que hace que la solución acorde con la ley presente una fractura en la continuidad de la demarcación, ausente en la solución A2. Por otra parte, se puede observar también que la solución A2 es más regular en sus bordes, lo que hace que los valores de las variables COMP y NR_SUP sean más altos. Una vez más, la utilización de la DMC hace que se obtengan resultados más acordes con las características del asentamiento que la utilización del valor fijo de 50 metros marcados por la ley.

6.5.3. Núcleo rural de Saa

El asentamiento rural de Saa cuenta con 987 parcelas y presenta un patrón espacial binuclear con edificaciones distribuidas a lo largo de una carretera principal y

⁹Estas variables equivalen a APT, COMP y NR_SUP, que en la gráfica se denotan por su nombre en inglés SUITAB, COMP y AREA_RS

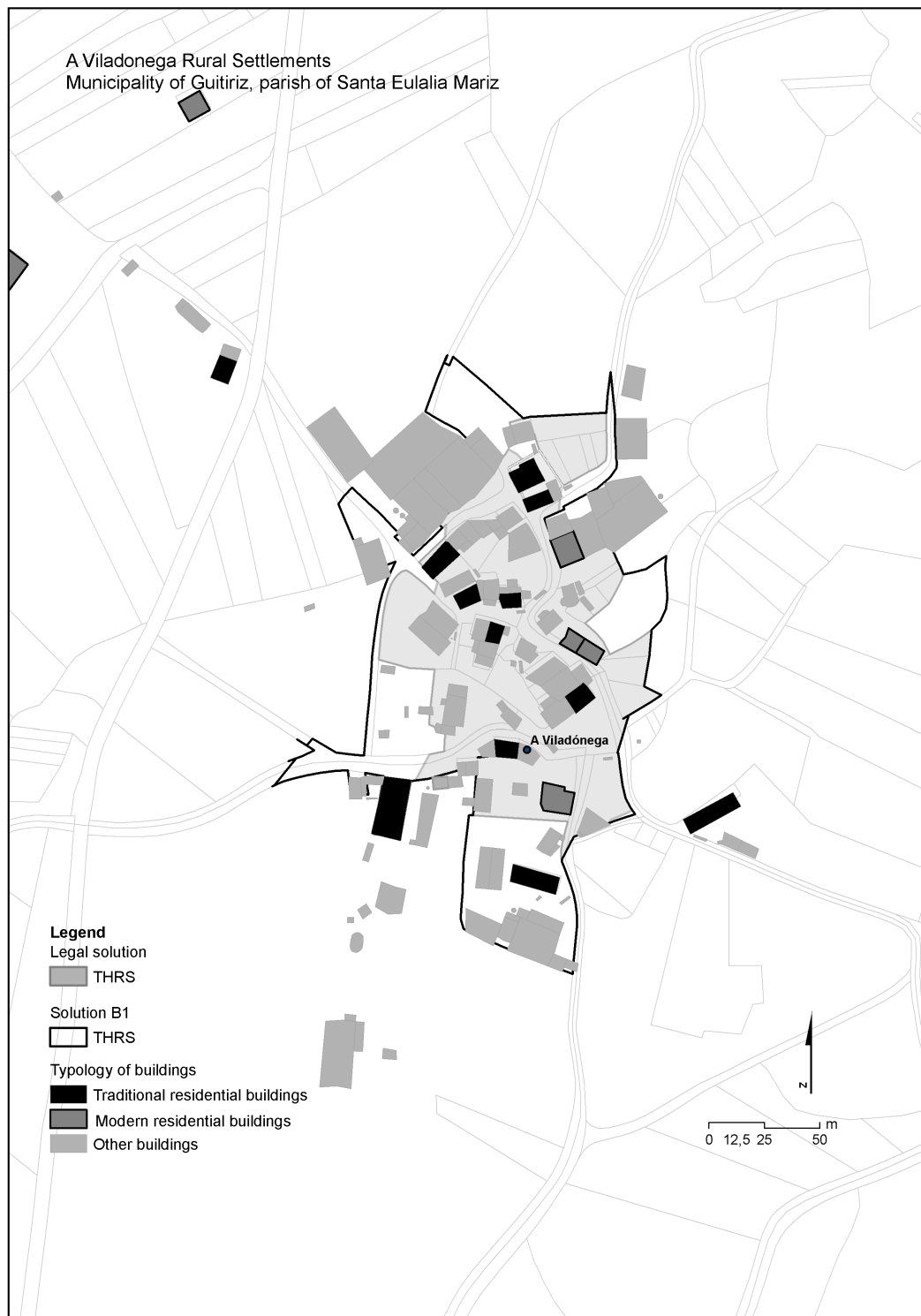


Figura 6.6: Demarcación del asentamiento rural de Viladonga obtenido por la solución acorde a la ley y por B1.

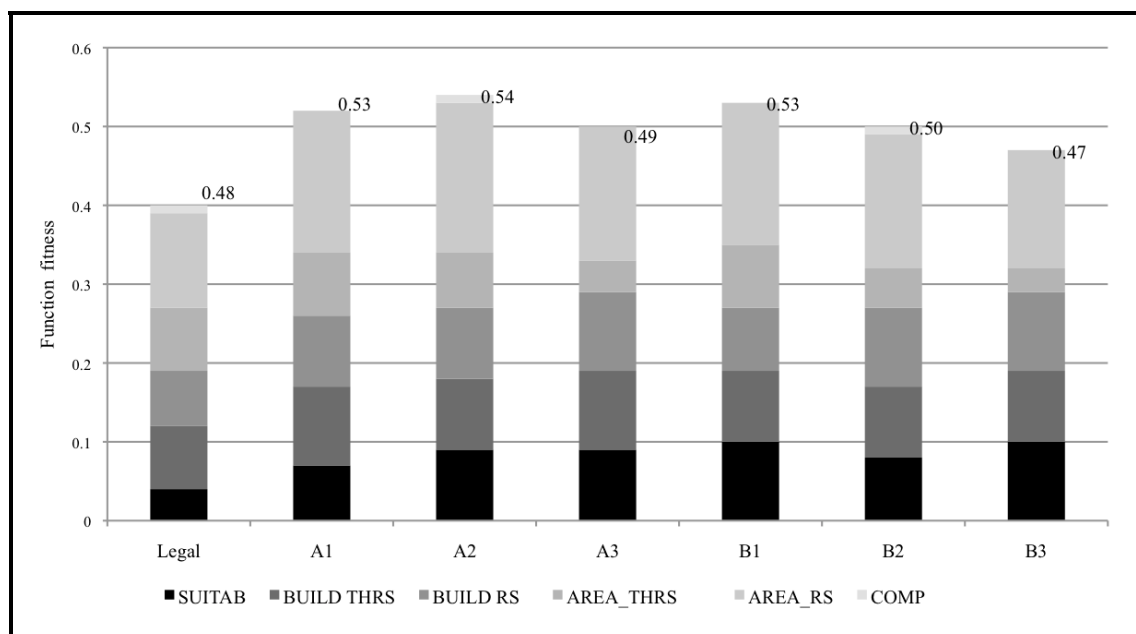


Figura 6.7: Valores de fitness de las ejecuciones del algoritmo PBIG sobre el asentamiento rural de Ferreira.

sus ramificaciones, formando dos áreas diferenciadas. En ambas áreas la dispersión de las viviendas tradicionales es menor que en los asentamientos anteriormente analizadas, mientras que las nuevas viviendas están localizadas en los límites norte y sur del asentamiento pero menos dispersas. Como se puede observar en la Figura 6.9, la demarcación obtenida con los criterios legales tiene un valor de fitness de 0,43, mientras que el resto de soluciones obtenidas con el método alternativo se obtienen iguales o mejores resultados, a excepción de la solución B3. La mejor solución de todas es la A1, con un valor de fitness del 0,52. La mayor diferencia entre ambas soluciones es la variable SUP_NR. La Figura 6.10 muestra que las dos soluciones asignan parcelas a las categorías NRHT y NRC pero la solución acorde con la ley presenta una división del NRHT en dos polígonos separados por otro polígono perteneciente al NRC, creando geometrías muy irregulares. Esto hace que el valor de la compacidad de esa demarcación sea nulo. Por su parte, la solución A1 tiene sólo un polígono para la categoría NRHT, permitiendo una unión espacial con una forma regular entre ambas áreas de viviendas tradicionales, incluyendo en el medio parcelas vacías para la futura construcción de viviendas nuevas. A su vez, una superficie pequeña al sur del asentamiento es asignada a la categoría NRC. También se puede

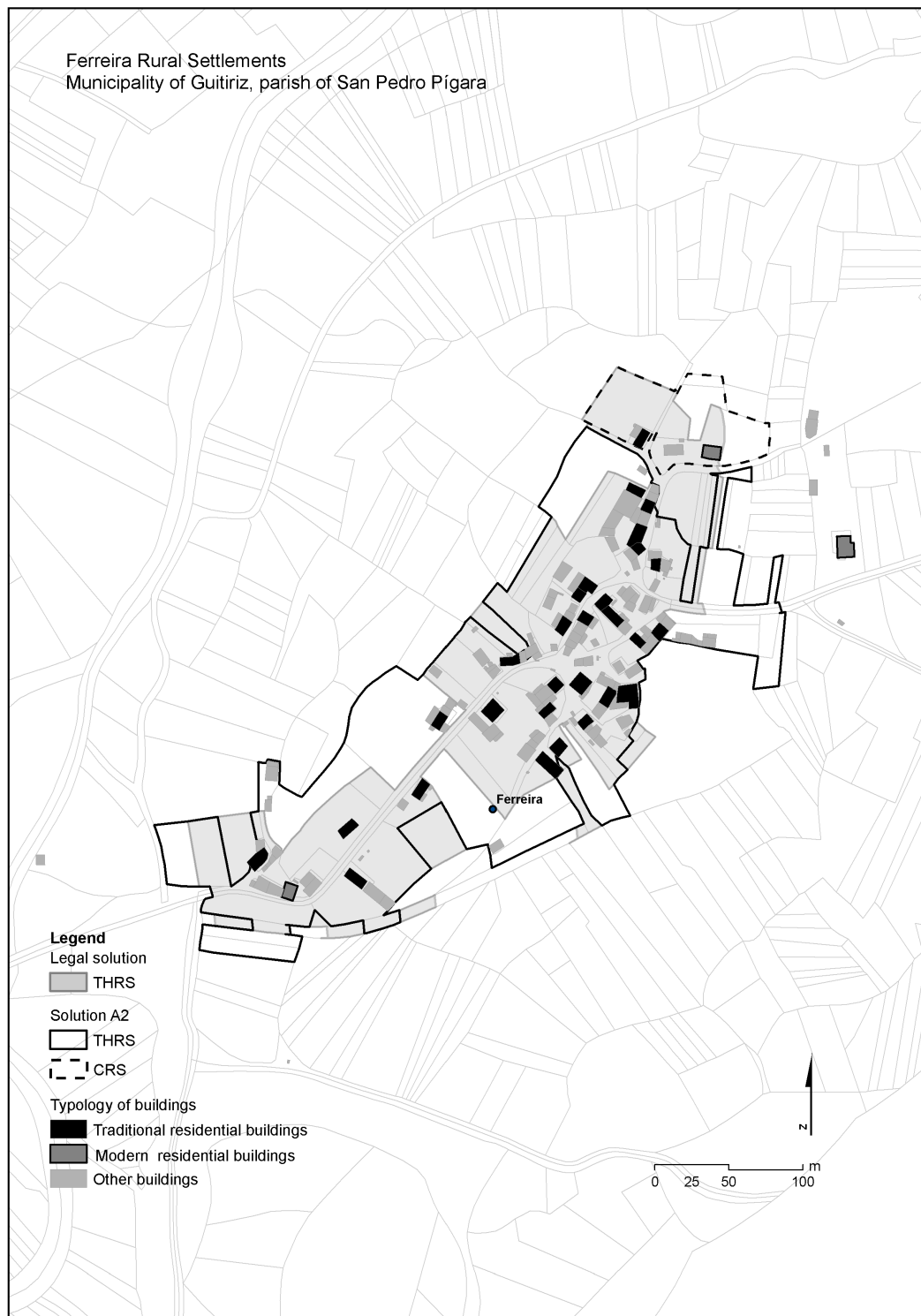


Figura 6.8: Demarcación del asentamiento rural de Ferreira obtenido por la solución acorde a la ley y por A2.

observar que a medida que el porcentaje de distancias utilizadas para los cálculos incrementa, el valor de la DMC decrece, obteniéndose peores resultados.

6.5.4. Análisis del comportamiento del algoritmo PBIG en casos reales

Con respecto al comportamiento del algoritmo durante su ejecución, la Figura 6.11 muestra ejemplos de la evolución del valor de la función de fitness a lo largo de, en este caso, 600 iteraciones. La gráfica se ha obtenido realizando tres ejecuciones sobre el núcleo rural de Ferreira con la configuración A2 como parámetros del algoritmo, y otros tres sobre el de Viladonga con la configuración B1. Como este último núcleo tiene bastante menos parcelas que el resto de los analizados, el número de parcelas con las que puede trabajar el operador R&A es muy reducido. Este hecho explica por qué el fitness alcanza un máximo muy pronto y todas las ejecuciones consiguen ese mismo máximo valor.

Con estos ejemplos de núcleos bien distintos morfológicamente y siguiendo patrones espaciales diferenciados, se puede concluir que la metodología alternativa propuesta siempre ofrece mejores resultados, en cuanto al valor del fitness se refiere, que las soluciones obtenidas aplicando los criterios legales actuales. Esto es posible principalmente debido a la posibilidad de adaptar los valores de los parámetros a las características morfológicas y espaciales de los asentamientos. En la idoneidad de una solución también influye de manera importante qué distancias se utilicen para los cálculos.

6.5.5. Cálculo del entorno inmediato de un núcleo rural

Conocer lo que rodea a un núcleo puede resultar de tanta importancia como una buena delimitación, ya que a partir del entorno de un núcleo rural se pueden tomar decisiones sobre otras actividades, como por ejemplo en procesos de concentración parcelaria [20].

El entorno inmediato de un núcleo debe satisfacer dos cuestiones básicas: debe contener al núcleo rural; y su geometría debe constar de un sólo polígono. Los pasos

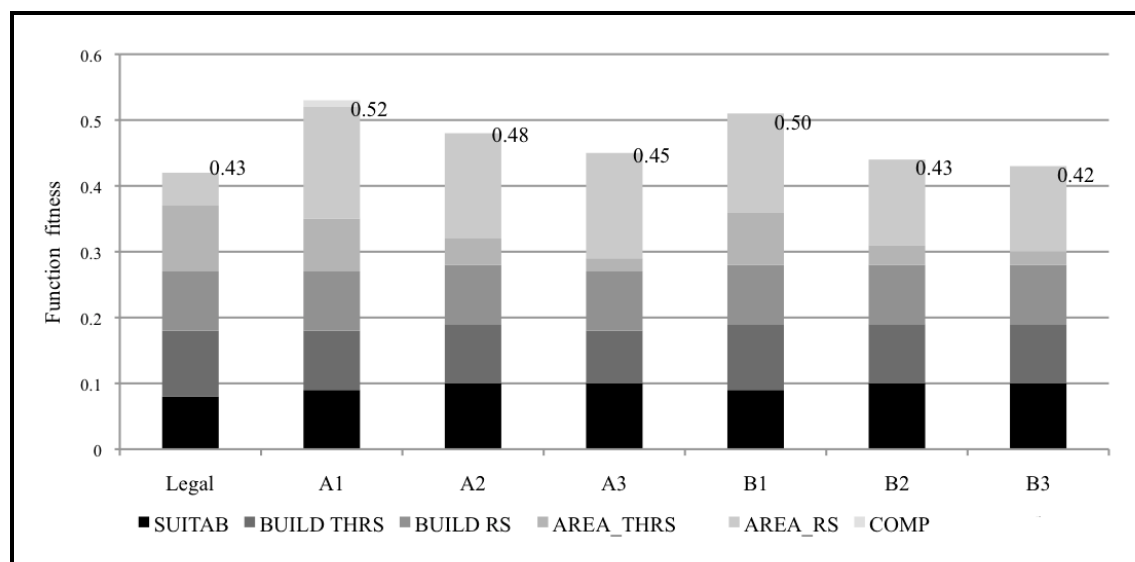


Figura 6.9: Valores de fitness de las ejecuciones del algoritmo PBIG sobre el asentamiento rural de Saa.

para su obtención son los que siguen.

Basándose en la delimitación del núcleo rural proporcionada por el algoritmo PBIG, se obtiene una envolvente base calculada mediante la unión de todos los polígonos de la delimitación. Como en esta nueva geometría pueden quedar huecos, lo primero que se debe hacer es cubrirlos. Para ello se calculan los puntos más próximos entre todos los polígonos y se trazan líneas rectas entre ellos para que todas las parcelas que las toquen, unirlas a la envolvente base. Si aún así quedan huecos, estos se obtienen con las funcionalidades que nos proporciona la librería JTS, y se unen las parcelas que queden en dichos huecos a la envolvente base.

A partir de esta envolvente base se crea una geometría que define el entorno inmediato. Esta nueva geometría se obtiene calculando un buffer de radio r , siendo r el resultado de multiplicar la DMC por un valor establecido por el usuario como parámetro de entrada. De este buffer se extrae la envolvente base creando una geometría con un hueco en el centro y sobre él se calculan las parcelas intersecadas teniendo en cuenta algunos aspectos. El primero es que las parcelas que formen parte del entorno inmediato al núcleo han que tener una aptitud mínima¹⁰. El segundo aspecto a tener

¹⁰La aptitud a cada parcela se establece en base a datos introducidos por el usuario y que le permite calcular distintos tipos de entornos (forestales, de recreo, de cultivo, etc.)

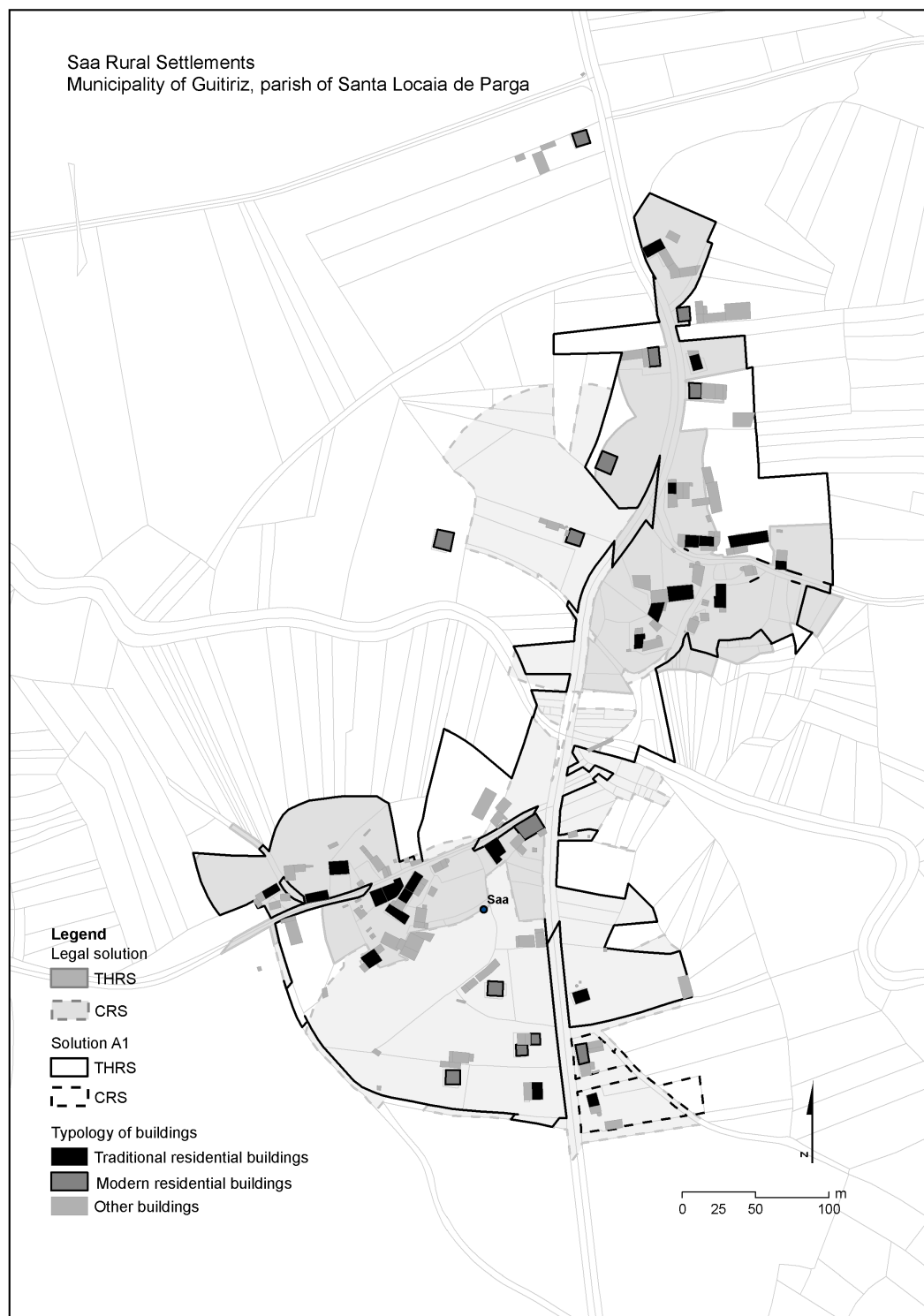


Figura 6.10: Demarcación del asentamiento rural de Saa obtenido por la solución acorde a la ley y por A1.

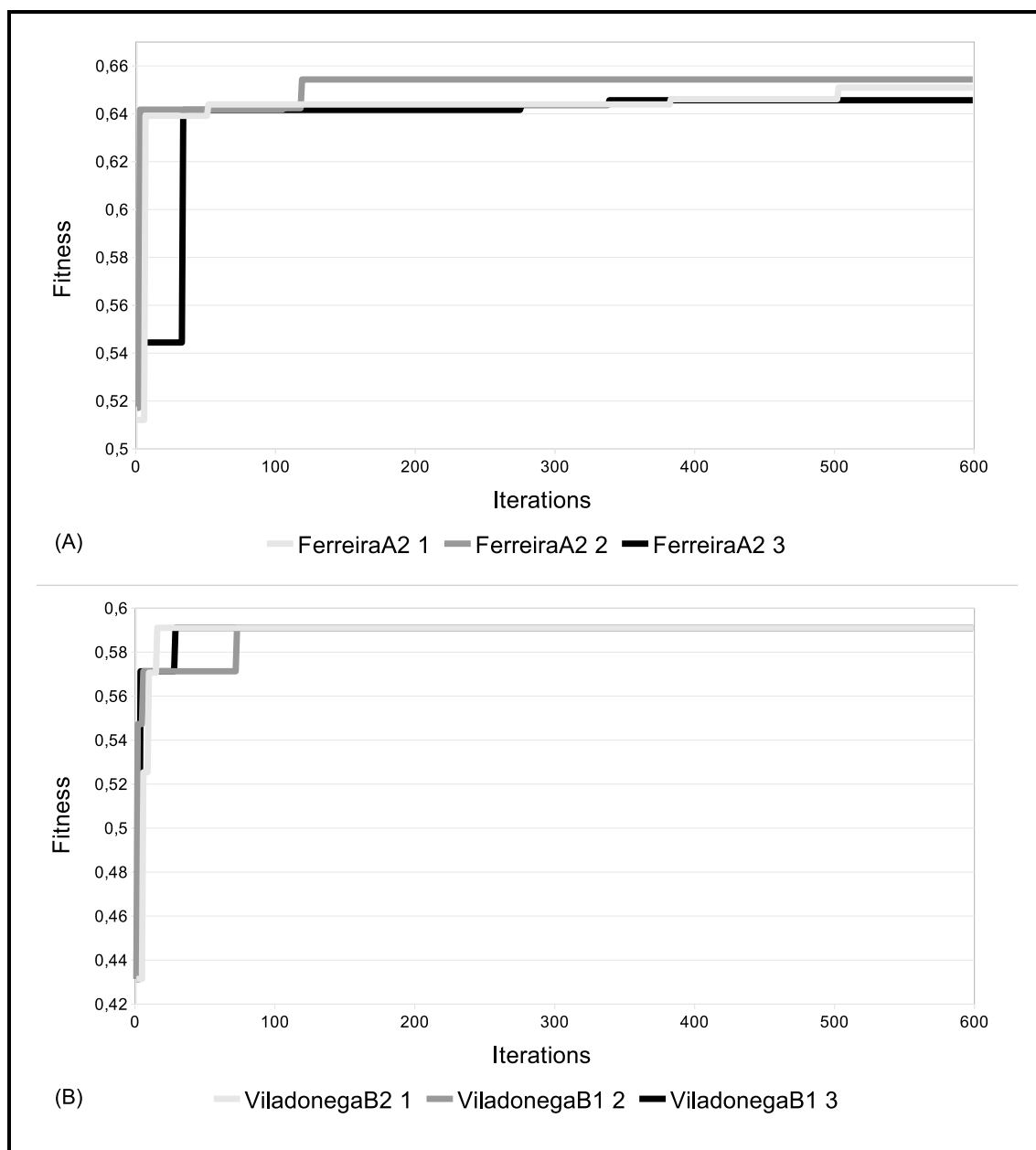


Figura 6.11: Evolución del fitness a lo largo de las iteraciones.

en cuenta es la superficie de la parcela: si esta es menor a 600 metros cuadrados toda la parcela formará parte del entorno, y si es mayor, sólo formará parte del entorno el resultado de intersecar la parcela con el buffer calculado previamente.

La Figura 6.12 muestra la delimitación del núcleo rural de Donis, en Lugo, donde la zona marrón corresponde al NRHT y la gris al NRC. También se muestran los puntos más cercanos entre ambos polígonos y su unión en línea recta, utilizada posteriormente servirá para la creación de un único polígono. Y en la Figura 6.13 se muestra el entorno inmediato del mismo núcleo calculado en base a un buffer de 0,5 la DMC y, en este caso, utilizando todos los usos del núcleo.

6.6. Resumen

En este capítulo se ha detallado el desarrollo de una metodología para los procesos de delimitación de núcleos rurales y un algoritmo *greedy* iterativo basado en población para aplicar a dichos procesos. Este algoritmo, que puede ser integrado en varios software SIG como por ejemplo gvSIG, proporciona una herramienta de gran potencial para la generación y comparación de diferentes alternativas de demarcaciones, siguiendo para ello criterios establecidos por las leyes actuales de ordenación del suelo y por la metodología propuesta.

La aplicación del algoritmo a varios asentamientos rurales reales con diferentes patrones espaciales, ha demostrado la habilidad del algoritmo para proporcionar soluciones de demarcación que se ajustan a las características morfológicas de los núcleos. Los resultados obtenidos pueden ser también un buen punto de partida para guiar a los expertos a la hora de configurar los parámetros del algoritmo en base a las características de las zonas donde quieran aplicarlo. El algoritmo también puede servir como herramienta de comparación entre delimitaciones obtenidas por diferentes medios.

En todas los núcleos analizados la nueva metodología propuesta siempre ha conseguido una solución mejor que la obtenida por criterios legales. Esto es debido a la posibilidad de poder variar los valores de los parámetros utilizados de forma adaptativa según la morfología del núcleo.

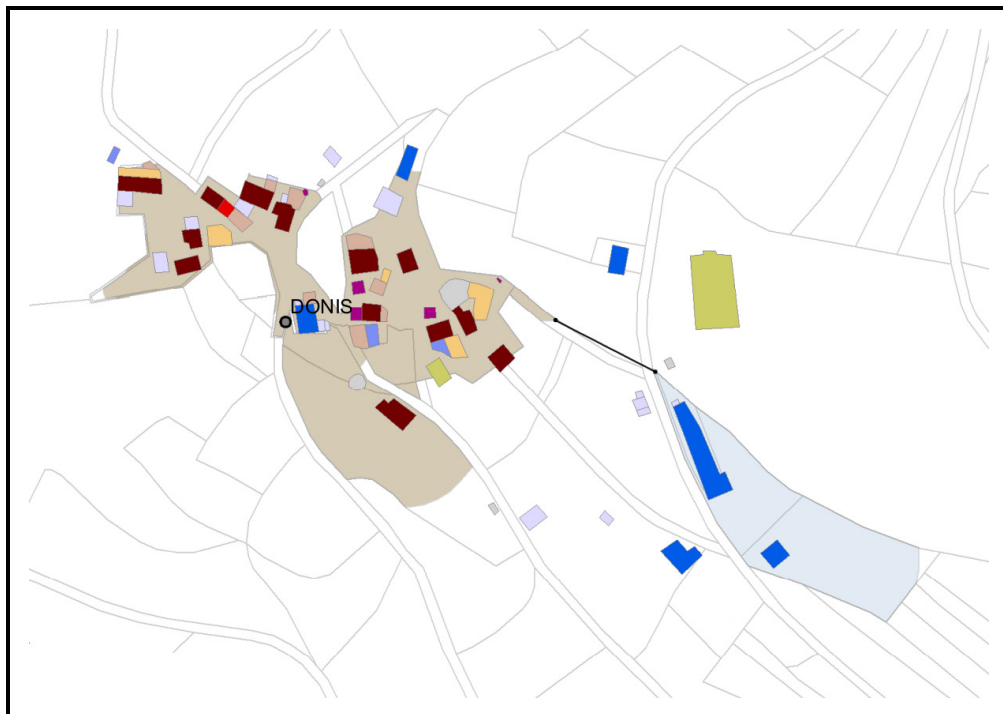


Figura 6.12: Delimitación del núcleo rural de Donis (Lugo).

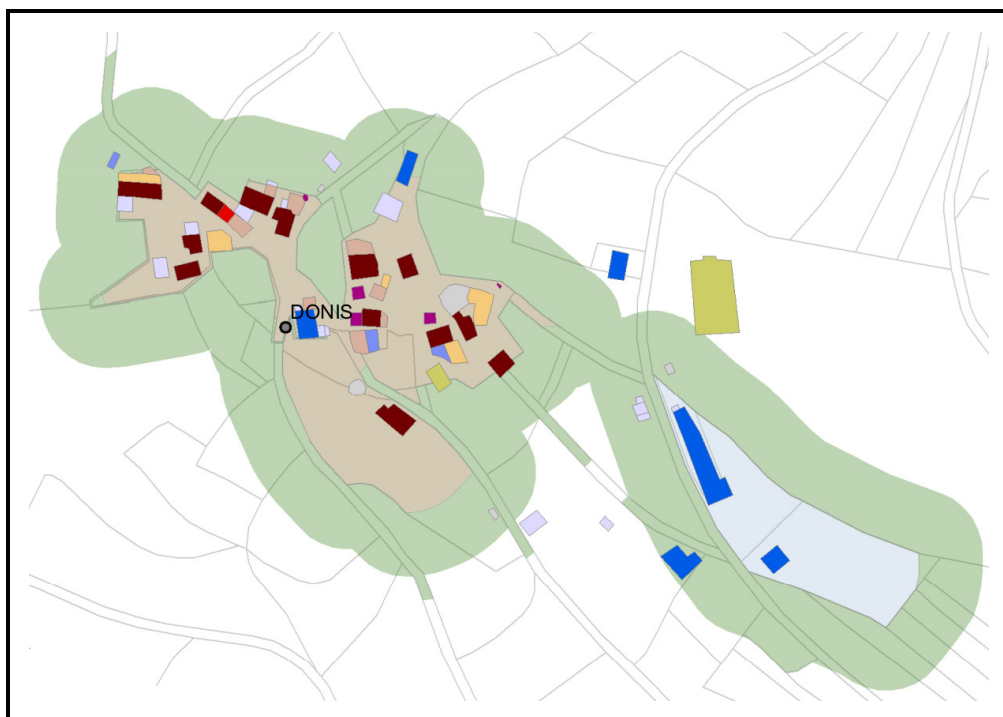


Figura 6.13: Delimitación y entorno inmediato del núcleo rural de Donis (Lugo).

También se ha automatizado el cálculo del entorno más próximo de un núcleo rural en base a su delimitación. Esto ayudará a los expertos en planificación a obtener datos útiles de manera rápida y sencilla que podrán ser utilizados para la realización de otras actividades, como puede ser, la concentración parcelaria.

En cuanto a publicaciones y participaciones en congresos, este proyecto se ha estructurado en dos partes, una inicial de estudio de los aspectos legales y los criterios de delimitación (no abordados en esta tesis) y una parte final de diseño e implementación del algoritmo en base a esos aspectos y criterios (que es lo que abarca este capítulo de la tesis). Así pues, de la parte inicial se han publicado los trabajos referenciados en [15, 16, 19] estando aún otros en periodo de revisión. Y en lo que respecta a la segunda parte, se han publicando los trabajos referenciados en [104, 103, 18], estando [17] en fase de revisión.

Capítulo 7

Algoritmo genético paralelo para la ordenación del territorio

Con los núcleos rurales ya delimitados por el algoritmo PBIG explicado en el capítulo anterior, es el momento de asignar las distintas categorías del suelo al resto de parcelas de un municipio. Por lo visto en la sección 5.2, los algoritmos genéticos parecen ser los mejores candidatos para buscar soluciones a este tipo de problemas.

Numerosos estudios han reconocido la naturaleza multi-objetivo de los problemas de planificación [74]. Estos objetivos a menudo incluyen criterios numéricos, como la idoneidad (aptitud) de asignación de categorías del suelo a parcelas [4, 35, 46, 137], y criterios espaciales, sobre todo la compacidad de las manchas¹ [1, 2, 45, 74, 77, 120]. La manera más simple para evaluar la compacidad está basada en cálculos sobre cada categoría, como puede ser el número de manchas [2, 74], la mancha más grande de cada categoría [2, 74], o el número de celdas vecinas con la misma categoría [1, 77]. Otros métodos más complejos para el cálculo de la compacidad están basados en la relación entre el área y el perímetro de cada mancha, como la media de las tasas del número de celdas en el perímetro de una mancha y el número total de celdas de la propia mancha [74], o la tasa media del perímetro dividido entre la raíz cuadrada del área de cada mancha [2]. En [26] se desarrolló un método más sofisticado para

¹La compacidad es una medida que indica la regularidad de una forma geométrica, que ya ha sido explicada en el Capítulo 6 pero en relación a un núcleo rural. En este capítulo la compacidad se aplica bien a un conjunto de parcelas con una misma categoría, o bien a las manchas, que se definen como *el conjunto de parcelas vecinas con una misma categoría*.

el diseño de las manchas basado en un algoritmo genético pero sólo aplicable a un entorno raster. En nuestro caso se ha optado por utilizar la aptitud y la compacidad como principales características a la hora de evaluar las soluciones proporcionadas por el algoritmo. Ambas serán detalladamente explicadas a lo largo de este capítulo.

La gran cantidad de parcelas que se ven involucradas en los planes de ordenación municipal hace que el coste computacional de los algoritmos al ejecutar un número de iteraciones suficientes como para explorar de manera amplia el espacio de búsqueda, sea muy alto. Por esta razón la utilización de tecnologías de computación de altas prestaciones, y en concreto la computación paralela, se hace fundamental a la hora de conseguir algoritmos que sean útiles y utilizables por los expertos. Hace aproximadamente 10 años, [30] y [90] apuntaron que el impacto de la computación paralela en el ámbito de la información geográfica era escasa y que se hacía necesario la implementación de más herramientas capaces de aprovechar las características paralelas de los procesadores que, en aquella época, empezaban a aparecer en el mercado. Hoy en día la variedad y proliferación de procesadores multi-núcleo para ordenadores personales, servidores e incluso para teléfonos móviles y tabletas, hace que cada vez aparezcan más aplicaciones paralelas en todos los ámbitos de la computación. No obstante, aún son muchas las que no sacan partido de todos los beneficios de la paralelización y se echa en falta más aprovechamiento de los recursos disponibles.

Volviendo al problema que nos ocupa de ordenación del territorio, en la literatura consultada no hay ejemplos de cómo manejar situaciones en las que el número de parcelas sea muy elevado. Es por esto que una de las principales tareas de este trabajo también es diseñar e implementar diferentes estrategias de paralelismo aplicadas al algoritmo para reducir sus tiempos de ejecución y para mejorar la calidad de sus resultados.

La estructura de este capítulo es la siguiente: la Sección 7.1 introduce el problema de la ordenación territorial y la complejidad del proceso de asignación de categorías del suelo a parcelas optimizando ciertos criterios. La Sección 7.2 describe la etapa del preprocesado, mientras que la Sección 7.3 profundiza en el algoritmo genético, sus fases y cómo se calcula el valor de la función de fitness. La Sección 7.4 explica cómo se ha paralelizado el algoritmo para que pueda ser ejecutado en tres entornos distintos: en sistemas multi-núcleo, sistemas clústers y sistemas clústers con nodos

multi-núcleo. La Sección 7.5 estudia la aplicación del algoritmo paralelo en un caso real en el municipio de Guitiriz (Lugo). Y por último las secciones 7.6 y 7.7 exponen funcionalidades adicionales del algoritmo y algunas conclusiones obtenidas, respectivamente.

7.1. Definición del problema de ordenación territorial

El problema de ordenación del territorio puede ser formulado como un problema de optimización en la cual cada trozo de tierra es asignado a la mejor categoría basándose en ciertos criterios y restricciones.

Las leyes sobre ordenación urbanística de Galicia [139, 142] comentadas en la Sección 6.1.1 del capítulo anterior, definen un conjunto de 15 categorías del suelo. Estas leyes también describen las restricciones que cada categoría ha de cumplir (e.g. el área mínima y máxima de cada una) y establecen las condiciones por las cuales un trozo de tierra puede, no puede, o debería ser asignado, a una categoría. La asignación de algunas de estas categorías está determinada de manera inequívoca por restricciones legales (e.g. la cercanía de una parcela a un río). Otras sin embargo, tienen mayor libertad a la hora de asignarse pero siempre respetando las restricciones impuestas. Por estas razones distinguiremos entre *categorías fijas* (las primeras) y *categorías no fijas* (las segundas).

Las leyes y los expertos en ordenación territorial aconsejan que el proceso de asignación debe considerar las fronteras reales de las parcelas del municipio, esto es, una parcela no debería ser dividida en partes con diferentes categoría. Es por esto que la unidad mínima utilizada en los algoritmos implementados sea la propia parcela, cumpliendo así con las condiciones para poder utilizar la codificación comentada en la Sección 5.4.1. Algunas de las parcelas tienen una categoría asignada directamente por ley (e.g. espacios naturales) y son llamadas *parcelas fijas*; al resto se les llama *parcelas no fijas*.

El problema de ordenación territorial puede ser resumido como *la asignación de una categoría no fija a cada parcela no fija*. Sea C el número de categorías no fijas,

y N el número de parcelas no fijas, el número de posibles soluciones es C^N . No obstante, y debido a las restricciones que se comentarán más adelante, el número de posibles soluciones será menor.

Por lo general las regulaciones sobre los planes de ordenación territorial no establecen explícitamente métodos para evaluar cuantitativamente y comparar planes entre sí. Hay muchas variables de diferente naturaleza que pueden considerarse para evaluar cualquier solución al problema. Este trabajo considera dos aspectos principales: la compacidad de la solución y la aptitud que cada parcela tiene respecto a la categoría asignada.

El problema de ordenación territorial es formalmente definido como *la optimización de la aptitud y la compacidad en un espacio de búsqueda definido por todas las posibles combinaciones de parcelas no fijas con categorías no fijas y sujeto a las restricciones dadas*. O lo que es lo mismo, se trata de *maximizar el valor de la una función de fitness (o de evaluación) que representa un compromiso entre la aptitud y la compacidad*.

Al adaptar esta definición al ámbito de los algoritmos genéticos, se tiene que cada gen de los individuos representa una parcela, con lo cual, y debido a la gran cantidad de parcelas con las que se trabaja, el tamaño de los individuos será muy grande, con las implicaciones negativas que esto tiene en cuanto a rendimiento y consumo de recursos. A su vez, los alelos (valores de los genes) representan una etiqueta que identifica de forma inequívoca la categoría no fija asignada a la correspondiente parcela no fija.

Cabe destacar que paralelamente al desarrollo de este algoritmo genético y dentro del contexto del proyecto de investigación en el que se enmarcan los trabajos realizados en el capítulo 6 y en este mismo, también se diseñó e implementó otro algoritmo con el que se intenta resolver el mismo problema de asignación de categorías del suelo pero mediante el método de *simulated annealing* [124].

7.2. Etapa de preprocesado

Prevía a la ejecución del algoritmo genético se hace necesaria una etapa de preprocesado similar a la que ya se ha descrito en la Sección 6.2 correspondiente al algoritmo PBIG para la delimitación de núcleos rurales. En este caso se deben identificar tanto las parcelas con categoría fijas como las parcelas fijas con categorías no fijas. Las primeras serán descartadas ya que su categoría es, por ley, invariable y no serán tenidas en cuenta por el algoritmo genético. Pero las segundas, aún a pesar de que no se pueden alterar sus categorías, sí que han de ser tenidas en cuenta ya que influyen directamente en algunos cálculos como, por ejemplo, la compacidad.

Para este cálculo, el de la compacidad, también es necesario conocer la longitud de las fronteras entre las parcelas, información proporcionada por las relaciones de vecindad obtenidas por el método explicado en la Sección 5.4.2.

7.2.1. Datos de entrada de la etapa de preprocesado

La etapa de preprocesado necesita, en primer lugar, un archivo shapefile con las geometrías de las parcelas de la zona donde se va a realizar el plan de ordenación, que normalmente será un municipio; y en segundo lugar, algunos datos sobre las categorías no fijas que pueden ser asignadas, concretamente se necesita saber la superficie máxima y mínima que debe ser asignada y los pesos de aptitud y compacidad aplicados a dicha categoría.

También pueden ser aportados (no es obligatorio) otros archivos shapes para dotar al algoritmo de mayor información sobre la zona. Estos datos son los relacionados con núcleos rurales histórico-tradicionales, núcleos rurales comunes, zonas de protección patrimonial, líneas de costa, núcleos de reservas de la biosfera, humedales, ZEPA², LIC³, ríos, aguas superficiales, vías, líneas de ferrocarril, líneas eléctricas, antenas telefónicas, gaseoductos, parques eólicos, mancomunidades, concentraciones parcelarias, terrenos incendiados, canteras y zonas improductivas.

²Zonas de Especial Protección para las Aves

³Lugares de Importancia Comunitaria

7.2.2. Datos de salida del preprocesado

La etapa de preprocesado genera un archivo Java *serializado* que contiene información que el algoritmo genético va a necesitar en la siguiente etapa. También se genera un shapefile con datos calculados sobre las parcelas, como por ejemplo las aptitudes a cada categoría.

7.3. Algoritmo genético para la asignación de usos del suelo

El algoritmo genético es el encargado de proporcionar una solución válida y suficientemente buena al problema de ordenación territorial que se plantea. Además la solución debe ser obtenida en un tiempo razonable para que el algoritmo sea utilizable.

Desde un punto de vista formal podemos definir el problema de la siguiente manera. Sea N el número de parcelas, cada parcela j está geométricamente caracterizada por su área a_j y su perímetro p_j . Sea C el número de categorías, cada parcela tiene asignada una única categoría. Sea N_i el conjunto de parcelas asignadas a cada categoría i -ésima y a_{ij} y p_{ij} el área y el perímetro de la j -ésima parcela en el i -ésimo conjunto, respectivamente. Además ap_{ij} denota la aptitud de la parcela j para la i -ésima categoría, valor que será especificado como dato de entrada.

A continuación se describen los datos de entrada y salida del algoritmo, las fases del algoritmo genético, y la función de fitness utilizada.

7.3.1. Datos de entrada del algoritmo genético

Parte de la información que el algoritmo necesita ya fue obtenida en la etapa de preprocesado. El resto de datos, relacionados sobre todo con el propio algoritmo genético y con su paralelización, han de ser especificados por el usuario como parámetros de entrada.

La Tabla 7.1 muestra todos los datos que el algoritmo genético necesita para

poder ejecutarse. Todos ellos son obligatorios.

7.3.2. Datos de salida del algoritmo genético

El algoritmo genético genera dos salidas: un shapefile con la solución al problema con todas las parcelas asignadas a una de las categorías, y un archivo serializado con la solución codificada de tal forma que sería posible volver a generar el shape de la solución en un futuro. Esto tiene la ventaja de que se podrían almacenar solamente los archivos serializados (ya que son de menor tamaño que los shapes) y cuando surgiese la necesidad de utilizar el shape con la asignación de categorías, se podría volver a generar con un programa que también se ha implementado.

7.3.3. Fases del algoritmo genético

La Figura 7.1 muestra el pseudo-código del algoritmo genético secuencial. En los siguientes párrafos serán descritos todos los pasos efectuados utilizando como referencia el número de línea del pseudo-código.

Sea P una población con M individuos, P_i un individuo de tamaño N (número de parcelas), $F(P_i)$ el valor de la función objetivo de P_i , $P_{i,j}$ la categoría del suelo del j -ésimo gen (parcela) de P_i , y Q el número de genes que han de ser mutados.

Los individuos iniciales con los que empezará a funcionar el algoritmo son creados de manera aleatoria (línea 1), pero esta población debe ser válida en el sentido de que deben satisfacer todas las restricciones impuestas por el problema. Si las restricciones son muy severas puede ocurrir que la población no se complete por la dificultad de crear individuos válidos. En este caso, y después de varios intentos, el algoritmo parará su ejecución y avisará al usuario de este contratiempo para que rebaje las restricciones. Una alternativa a la creación aleatoria de individuos puede ser empezar con individuos ya válidos si el usuario puede proporcionarlos. No obstante, en todos los tests realizados el algoritmo siempre ha podido completar la inicialización de la población.

De ahora en adelante llamaremos *bucle genético* al comprendido entre las líneas 2 y 16 que es en el cual son realizadas las operaciones de selección, cruce, mutación

Dato de entrada	Tipo	Descripción
Parcelario del preprocesado	Shapefile	Archivo generado en el preprocesado con información sobre las parcelas.
Población inicial	Entero	Número de individuos que formarán parte de la población.
Tasa de cruce	Doble	Probabilidad de que la operación de cruce se ejecute. Rango $[0,1]$.
Tasa de mutación	Doble	Probabilidad de que la operación de mutación se ejecute. Rango $[0,1]$.
Repetición de fitness para reducción tasa de mutación	Entero	Número de veces que el mejor valor de fitness de la población ha de repetirse entre iteraciones consecutivas para reducir la probabilidad de mutación. Se explica más detalladamente en la Sección 7.5.2.
Reducción tasa mutación	Doble	Valor por el que será multiplicado la tasa de mutación para su reducción. Rango $[0,1]$.
Método cálculo de compacidad	Selección	Método por el que se va a calcular la compacidad de la solución. Opciones: <i>por categorías</i> o <i>por manchas</i> .
Número de threads	Entero	Número de hilos (o <i>threads</i>) en los que se ejecutará el algoritmo paralelo.
Horas de ejecución	Entero	Número de horas durante las que se ejecutará el algoritmo.
Peso para la aptitud	Doble	Peso para el valor de la aptitud en la función objetivo. Rango $[0,1]$.
Peso para la compacidad	Doble	Peso para el valor de la compacidad en la función objetivo. Rango $[0,1]$.
Permitir individuos iguales	Boolean	Indica si se permite que en la misma población existan individuos iguales.
Datos del preprocesado	Serializable	Archivo serializado obtenido del preprocesado.

Tabla 7.1: Datos de entrada de la etapa de preprocesado para la delimitación de núcleos rurales.

<pre> 1 for ($i = 1$ to M) { creación aleatoria de P_i } 2 while ($tiempo_ejec \leq tiempo_max$) { 3 $i = 1$; while ($i \leq M$) { 4 selección P_{i_0} y P_{i_1} donde $1 \leq i_0, i_1, \leq M$ 5 selección aleatoria j_0 y j_1 donde $1 \leq j_0 < j_1 \leq N$ 6 $\tilde{P}_{i_0} = P_{i_{0_1}} \dots P_{i_{0_{j_0}}} P_{i_{1_{j_0+1}}} \dots P_{i_{1_{j_1}}} P_{i_{0_{j_1+1}}} \dots P_{i_{0_N}}$ 7 $\tilde{P}_{i_1} = P_{i_{1_1}} \dots P_{i_{1_{j_0}}} P_{i_{0_{j_0+1}}} \dots P_{i_{0_{j_1}}} P_{i_{1_{j_1+1}}} \dots P_{i_{1_N}}$ 8 for ($a = 1$ to Q) { 9 selección aleatoria k_0 y k_1 donde $1 \leq k_0, k_1 \leq N$ 10 $\tilde{P}_{i_0 k_0} = cat_aleatoria$; $\tilde{P}_{i_1 k_1} = cat_aleatoria$ 11 } 12 $P'_i = \hat{P}_\alpha$ donde $F(\hat{P}_\alpha) =$ 13 $= max\{F(P_{i_0}), F(P_{i_1}), F(\tilde{P}_{i_0}), F(\tilde{P}_{i_1})\}$ 14 $P'_{i+1} = \hat{P}_\beta$ donde $F(\hat{P}_\beta) =$ 15 $= max_{2^{nd}}\{F(P_{i_0}), F(P_{i_1}), F(\tilde{P}_{i_0}), F(\tilde{P}_{i_1})\}$ 16 $i = i + 2$ 17 } 18 $P = P'$ 19 } 20 return $\hat{P}_\gamma$ donde $F(\hat{P}_\gamma) = max\{F(P_0), \dots, F(P_M)\}$ </pre>	} SELECCIÓN } CRUCE } MUTACIÓN } ELECCIÓN
---	---

Figura 7.1: Pseudo-código del algoritmo genético secuencial.

y elección.

Para la operación de selección el algoritmo utiliza la técnica de la ruleta (roulette-wheel) [65] (línea 4). Esta técnica es proporcional respecto al valor de la función de fitness de cada individuo, esto es, los individuos con mayor valor de fitness tienen más posibilidades de salir escogidos. El principal inconveniente de este método es que puede reducir el espacio de búsqueda si existen súper-individuos (individuos con un valor de fitness mucho más alto que el resto) en la población, pero su principal ventaja es que no tiene que ordenar a los individuos según su fitness, operación que puede llegar a ser costosa y que tienen que hacerla métodos basados en *ranking* [28].

Para el operador de cruce fue implementado el método por dos puntos [119] con selección de puntos aleatorios (líneas 5-7). Este método selecciona aleatoriamente

dos posiciones del array de parcelas e intercambia los genes entre dichos puntos. El usuario se encarga de establecer la probabilidad con la que el operador de cruce será ejecutado.

De la misma manera, el usuario también tiene que establecer la probabilidad con la que el operador de mutación será aplicado a los individuos resultantes de las operaciones anteriores. Teóricamente cada gen de un individuo puede mutar o no con la probabilidad establecida, pero para implementar esto se tendrían que recorrer todos los genes de cada individuo y comprobar para cada uno si muta o no muta. Realizar esta operación para cada individuo de la población y en cada iteración es muy costoso, y por ello se ha decidido implementar la mutación de otra manera. El número de genes a ser mutados (Q) puede ser calculado de manera previa multiplicando la probabilidad de mutación establecida por el usuario por la longitud de un individuo y así escoger aleatoriamente Q genes (líneas 8-11). De esta manera se mejora el rendimiento de la operación de mutación y por ende el del algoritmo. A pesar de la importancia que tiene para el resultado final el escoger bien la probabilidad de mutación, no existe una manera definida de cómo seleccionar el mejor valor. La literatura consultada sugiere diferentes valores: desde los 0,0001 de DeJong [37] hasta los $1/L$ de Grefenstette, siendo L el tamaño de los individuos [68]. En la sección 7.5 se explican algunas conclusiones sacadas en este sentido.

En cada iteración del bucle genético se escogen dos⁴ individuos para que formen parte de la nueva población (líneas 12-13). El algoritmo utiliza el llamado *elitismo*, así los cuatro individuos (padres y descendencia) entran en el juego de la elección. Aquellos dos con el mayor valor de fitness serán los que pasen a formar parte de la nueva población. Si por casualidad un individuo escogido fuese igual que algún otro de la nueva población y si el usuario forzó a la unicidad de los individuos seleccionando el parámetro de entrada a tal efecto, el nuevo individuo sería mutado. Si no se encontrara ningún individuo válido después de un determinado número de intentos, se crearía uno nuevo aleatoriamente.

El criterio de parada del algoritmo está basado en tiempo en lugar de depender de la función de fitness (línea 2) que suele ser lo habitual. Esto se hace así porque es muy difícil conocer a priori cual tiene que ser el valor de la función de fitness para

⁴Se escogen dos individuos en vez de uno, como puede ser lo habitual, ya que el reparto de trabajo en las versiones paralelas del algoritmo se utilizan valores pares por sencillez.

que una solución sea aceptable. El usuario por tanto sólo ha de establecer el tiempo que el algoritmo debe ser ejecutado. Para finalizar, el algoritmo genético devolverá el mejor individuo de la última generación disponible (línea 18).

7.3.4. La función de fitness

La función de fitness proporciona un método para evaluar cuantitativamente los individuos de una población. Comparando los valores de fitness el algoritmo puede determinar si un individuo es mejor que otro. La función de fitness utilizada (Ecuación 7.1) contiene dos sumandos: la aptitud que cada parcela tiene a la categoría asignada, y la compacidad, que puede ser por manchas o por categorías como luego se verá más detenidamente. Ambos sumandos tienen unos pesos asociados que permiten al usuario establecer la importancia de cada uno.

$$Fitness = w_a * Aptitud + w_c * Compacidad \quad (7.1)$$

donde w_a y w_c siendo $w_a + w_c = 1$, son los pesos de los sumandos, y son establecidos por el usuario como parámetros de entrada.

La aptitud es calculada utilizando la media de las aptitudes de cada categoría. A su vez, la aptitud de cada categoría es obtenida como la media de las aptitudes de las parcelas asignadas a dicha categoría multiplicada por el área de cada parcela y normalizada por la suma total de las área de las parcelas con la misma categoría. Esto es definido por la siguiente ecuación:

$$Aptitud = \sum_{i=1}^C w_i \left(\frac{\sum_{j=1}^{N_i} ap_{ij} * a_{ij}}{\sum_{j=1}^{N_i} a_{ij}} \right) \quad (7.2)$$

donde C es el número de categorías, w_i es el peso de la categoría i -ésima (parámetro de entrada), N_i es el número de parcelas con la categoría i -ésima asignada, ap_{ij} es la aptitud que la parcela j -ésima tiene para la categoría i -ésima, y a_{ij} es el área de la parcela j -ésima con la categoría i -ésima.

La compacidad puede ser definida de varias formas diferentes. Se proponen dos: una basada en manchas (grupos de parcelas vecinas con la misma categoría) y otra

basada en categorías (grupos de parcelas no necesariamente vecinas con la misma categoría). El usuario puede escoger qué método utilizar.

La compacidad basada en manchas es definida de la segunda manera:

$$Compacidad_{manchas} = 4\Pi \sum_{i=1}^C w_i \left(\frac{\sum_{j=1}^{NP a_i} \frac{a_{ij}}{p_{ij}^2}}{NP a_i} \right) \quad (7.3)$$

donde $NP a_i$ es el número de manchas de la categoría i -ésima, p_{ij} y a_{ij} son el perímetro y el área de la mancha j -ésima con la categoría i -ésima, respectivamente, y C y w_i tienen el mismo significado que en la ecuación 7.2.

Esta ecuación surge debido a que para un área dada, un círculo maximiza la llamada *circularidad* [94], que se define como:

$$Circularidad = 4\Pi \frac{area}{perimetro^2} \quad (7.4)$$

La Ecuación 7.3 tiene un coste computacional alto ya que las manchas de cada individuo han de ser calculadas en cada iteración. La función de compacidad basada en categorías es mucho menos costosa y, como se mostrará en los resultados de los experimentos, produce unos buenos resultados. Se define de la siguiente manera:

$$Compacidad_{categorias} = 4\Pi \sum_{i=1}^C w_i \left(\frac{\sum_{j=1}^{N_i} a_{ij}}{(\sum_{j=1}^{N_i} p_{ij})^2} \right) \quad (7.5)$$

donde C , w_i , N_i , y a_{ij} tienen el mismo significado que en 7.2, y p_{ij} es el perímetro de la j -ésima parcela con la categoría i -ésima asignada.

Las restricciones en el área que cada categoría ha de tener implica: $\min_i \leq \sum_{j=0}^{N_i} a_{ij} \leq \max_i$, donde $\min_i$ y $\max_i$ son la mínima y máxima área permitida para la categoría i -ésima respectivamente, valores estos especificados en los parámetros de entrada.

7.4. Paralelización del algoritmo genético para la asignación de usos del suelo

Los tiempos de ejecución del algoritmo genético son potencialmente muy altos debido a la gran cantidad de parcelas involucradas y a la propia naturaleza del problema. Para hacer un algoritmo práctico y utilizable estos tiempos de ejecución deben ser reducidos. La solución adoptada es la paralelización del algoritmo. Tres soluciones han sido implementadas para este cometido: una basada en el paradigma de memoria compartida (sistemas multi-núcleo); otra basada en el paradigma de paso de mensajes (sistemas clúster); y otra híbrida que combina las anteriores.

7.4.1. Algoritmo genético paralelo en sistemas multi-núcleo

Como ya se ha comentado, actualmente muchos dispositivos electrónicos poseen procesadores multi-núcleo. Es importante por tanto diseñar sistemas que aprovechen al máximo las capacidades que proporciona este tipo de arquitectura.

La paralelización del algoritmo implementado se centra en el bucle genético. Cada thread está encargado de evolucionar un conjunto de individuos utilizando para ello los operadores de selección, cruce, mutación y elección. Como los threads comparten memoria, todos pueden leer de la misma población y escribir en la nueva población.

El número de threads a utilizar es definido por el usuario en los parámetros de entrada. Dado este número y el tamaño de población, se calcula el número de individuos que cada thread ha de generar. La distribución del trabajo debe ser lo más balanceada posible en términos de tiempos de ejecución. La Figura 7.2 muestra un ejemplo de cómo se realiza la distribución del trabajo. Todos los threads pueden seleccionar cualquier individuo de la población pero sólo pueden escribir los nuevos individuos en las posiciones asignadas a cada thread. Esto último se hace para evitar conflictos en la escritura de la nueva generación.

La Figura 7.3 muestra la estructura y etapas del algoritmo genético multi-núcleo.

El paquete *java.util.concurrent* de la plataforma J2EE [75] es el utilizado para

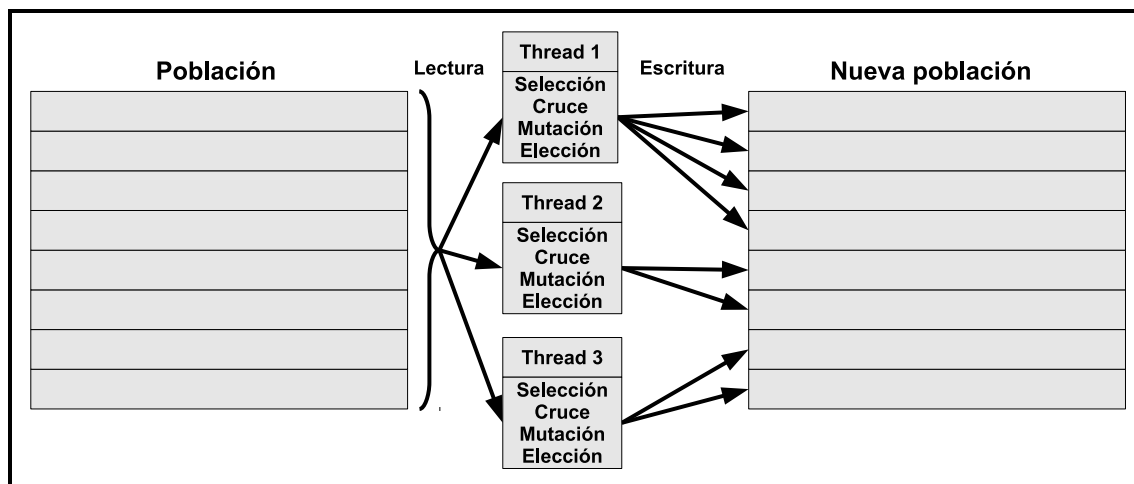


Figura 7.2: Balanceo de carga entre threads esclavos en el algoritmo paralelo multi-núcleo.

lanzar los threads en los diferentes núcleos del procesador. Para cada iteración del bucle genético, un thread actuará como *máster* y será el encargado de lanzar el resto de threads *esclavos* y esperar a que acaben. Cada thread esclavo es independiente del resto, así que la paralelización es sencilla y eficiente. Una vez los threads esclavos acaban su ejecución, la nueva población estará ya creada y cuando el thread maestro recupera el control, solo tendrá que comprobar el criterio de parada.

Con este método de paralelización no hay garantía de obtener mejores individuos que en la misma iteración del algoritmo secuencial. Sin embargo lo que ocurre es que se ejecutan más iteraciones del bucle genético en el mismo espacio de tiempo y por tanto, se llega a mejores soluciones en menos tiempo.

7.4.2. Algoritmo genético paralelo en sistemas clúster

Los clústers son sistemas de memoria distribuida por lo que los procesos que se ejecutan en este tipo de entornos han de ser capaces de enviar y recibir mensajes para intercambiar información. Para implementar estas capacidades se ha utilizado una librería Java de código abierto para el paso de mensajes llamada MPJ Express [116].

Cada proceso es ejecutado en un nodo del clúster y el proceso máster será el encargado de establecer las comunicaciones con los procesos esclavos. Cada uno

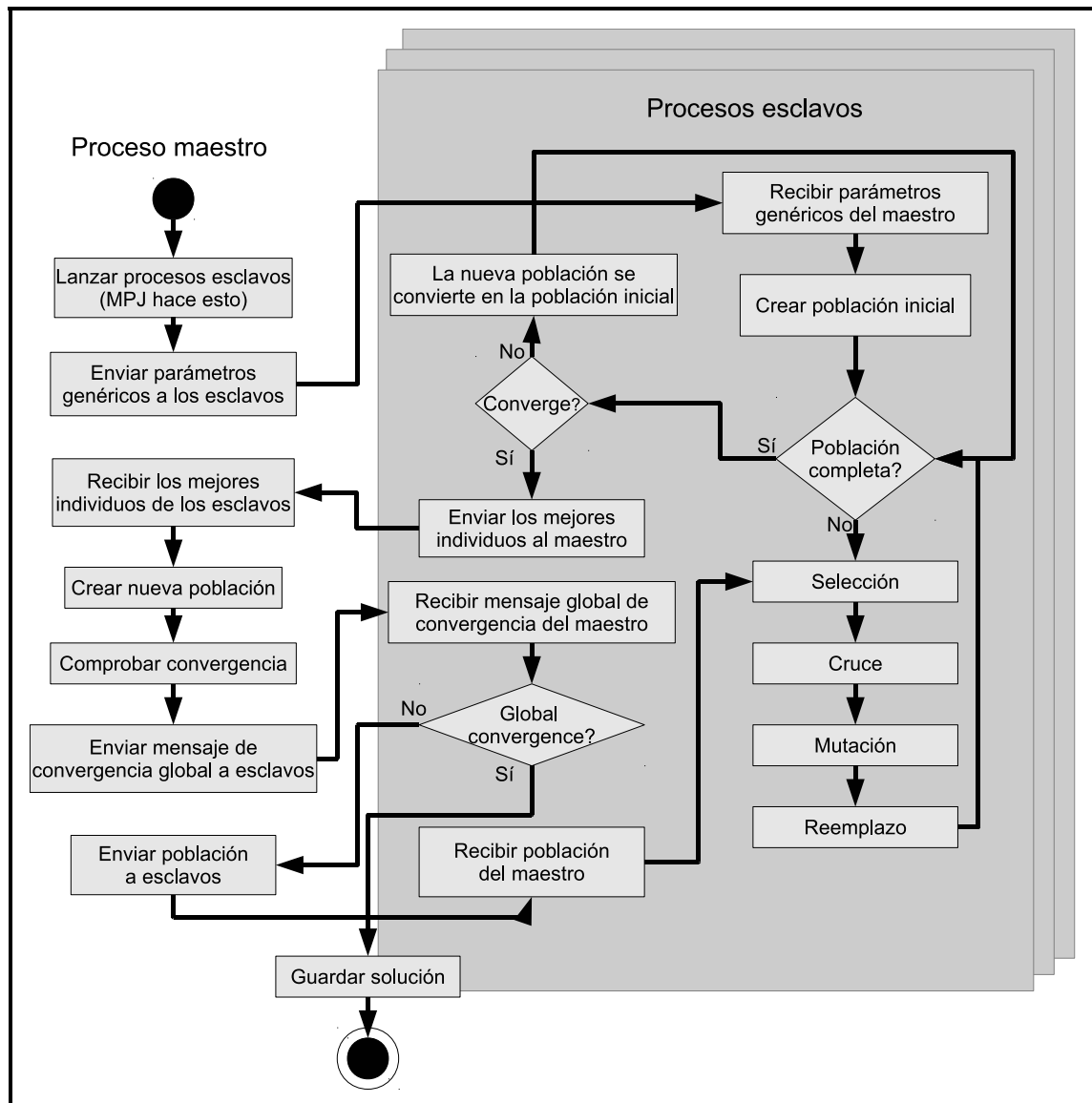


Figura 7.3: Flujo de trabajo del algoritmo paralelo multi-núcleo.

de estos ejecuta el algoritmo genético entero de forma independiente utilizando y evolucionando su propia población. Existe una sincronización periódica que consiste en enviar los mejores individuos de los procesos esclavos al proceso máster. Después de que este último reúna toda la información, crea una nueva población que enviará a los esclavos para que vuelvan a ejecutar sus algoritmos genéticos locales. En cada sincronización el máster también comprobará el criterio de parada. El usuario debe especificar cada cuanto se realiza la sincronización pudiendo ser segundos, minutos u horas. La Figura 7.4 ilustra este tipo de paralelismo.

Esta versión del algoritmo paralelo se aprovecha de que cada proceso esclavo ejecuta el algoritmo genético de manera independiente, y aunque parten de la misma población, cada proceso esclavo la evolucionará a su propia manera debido a la componente aleatoria de los operadores genéticos. Los mejores individuos de cada esclavo son posteriormente enviados al máster que los reunirá en una nueva generación, lo que hará subir la media del fitness de los individuos e incrementar el espacio de búsqueda.

Un aspecto muy interesante es el hecho de que cada nodo esclavo puede leer su propio archivo de configuración local. Esto permite que cada proceso pueda trabajar con sus propios parámetros de entrada, como por ejemplo las probabilidades de ejecutar el cruce y/o la mutación. Esto añade más heterogeneidad a la población y amplía el espacio de búsqueda.

7.4.3. Algoritmo genético paralelo en sistemas clúster con nodos multi-núcleo

Los métodos comentados en las secciones anteriores pueden ser utilizados conjuntamente produciendo un nuevo tipo de paralelización híbrida. La Figura 7.5 muestra todas las comunicaciones entre el proceso máster y los esclavos (versión clúster) y cómo estos lanzan los threads locales (versión multi-núcleo). Este nuevo método es el que mejores resultados produce según las pruebas que se explicarán en la Sección 7.5.

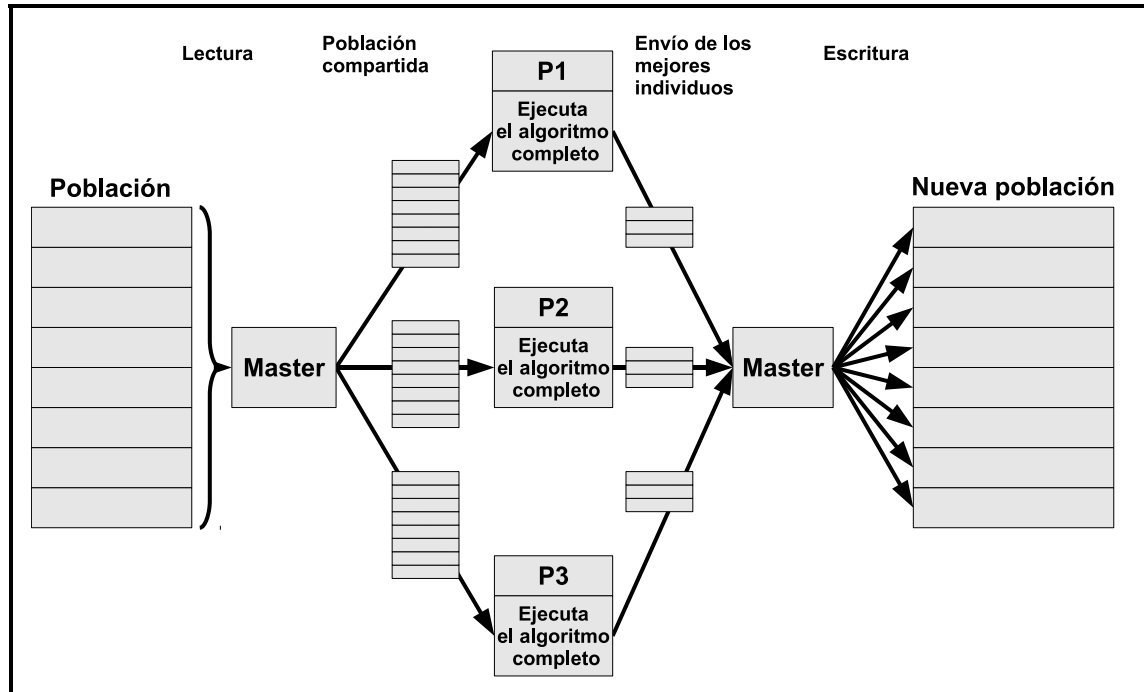


Figura 7.4: Flujo de trabajo del algoritmo genético paralelo para sistemas clúster

7.5. Casos de estudio

Las pruebas se han ejecutado en el municipio gallego de Guitiriz debido principalmente a la gran cantidad de parcelas que contiene, a la heterogeneidad morfológica de sus asentamientos, y a la variedad de los usos de la tierra. Guitiriz cuenta con 138.175 parcelas de las que 52.045 tienen categoría fija. Los individuos utilizados por el algoritmo tendrán por tanto 83.130 genes, de los que 12.232 corresponden a parcelas fijas. Por tanto, las 75.898 parcelas restantes serán las consideradas como candidatas a cambiar su categoría. Para este municipio en concreto se tienen en cuenta 4 categorías no fijas: áreas naturales, agrícola, forestal y urbana.

Todas las pruebas fueron ejecutadas en un clúster cuyos nodos cuentan con dos procesadores Intel Xeon E5520 y 8 GB de RAM. Cada procesador tiene 4 núcleos a 2,27 Ghz con hyper-threading, lo que en la práctica se traduce en que dos threads pueden ser utilizados simultáneamente por cada núcleo.

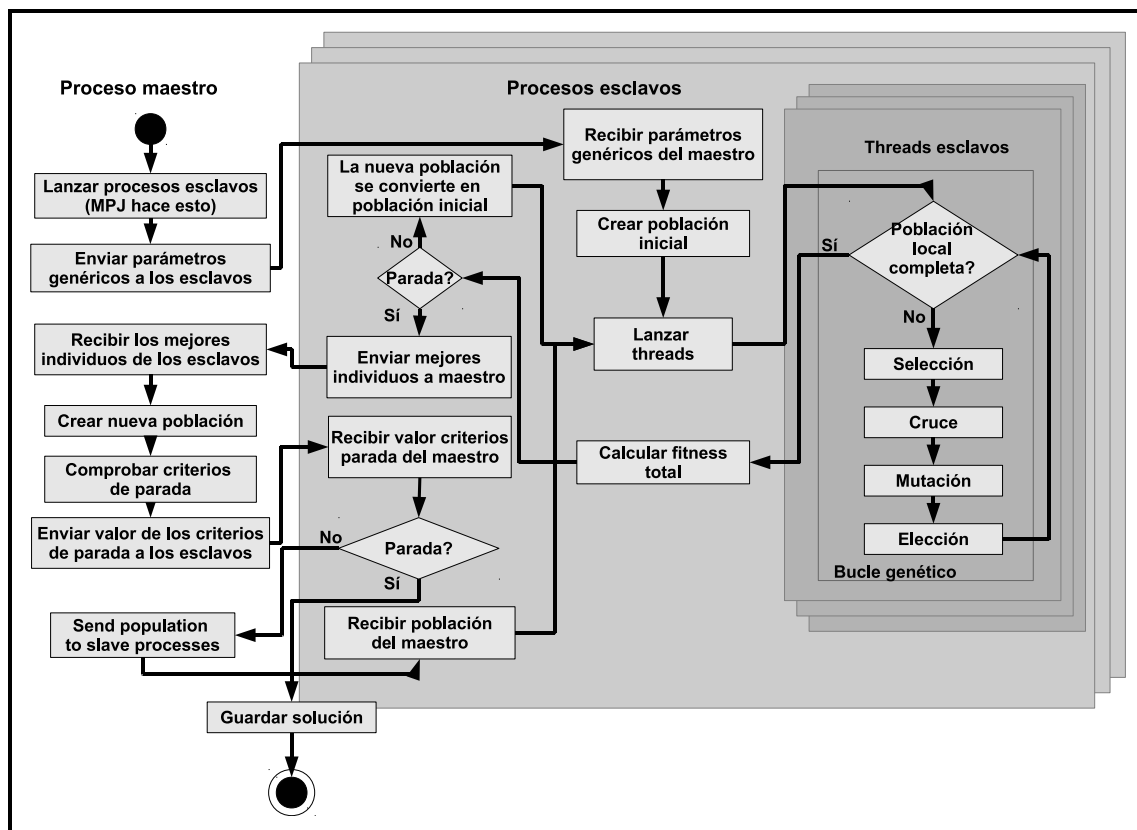


Figura 7.5: Flujo de trabajo del algoritmo genético paralelo híbrido.

7.5.1. Influencia del tamaño de la población

El tamaño de la población se corresponde con el número de individuos utilizados en cada iteración. Cuantos más individuos haya en la población, más lento será el algoritmo. Sin embargo el espacio de búsqueda explorado será más amplio y los individuos serán más heterogéneos, lo que en teoría provoca una mayor probabilidad de encontrar un mejor resultado. Por tanto, encontrar la óptima relación entre el tiempo de ejecución y la amplitud del espacio de búsqueda se convierte en una tarea esencial para el buen rendimiento del algoritmo.

Para ver la influencia del tamaño de la población se ha ejecutado el algoritmo con distintos tamaños de población, concretamente se han utilizado los valores 32, 64, 128 y 256. La Figura 7.6 muestra la media del valor de fitness de cada configuración ejecutada 20 veces cada una durante 3 horas. El peso para la componente de la aptitud de la función de fitness fue de 0.25 y el de la compacidad de 0.75.

Observando la Figura 7.6 podemos comprobar que para poblaciones pequeñas (pocos individuos) el algoritmo ejecuta más iteraciones por unidad de tiempo y consigue mejores resultados en el corto plazo que las ejecuciones que utilizan un mayor número de individuos. En ejecuciones más largas, de 24 horas por ejemplo (no mostradas en el gráfico por temas de legibilidad), se ha comprobado que el valor de la función de fitness va mejorando ligeramente cuantos más individuos participen en el algoritmo. Presumiblemente esto sea debido a lo comentado de que el espacio de búsqueda es mayor y existe más heterogeneidad entre los individuos.

7.5.2. Influencia de la tasa de mutación

Elegir una buena tasa de mutación también es importante para conseguir buenos resultados cuando se utilizan algoritmos genéticos. Una tasa de mutación pequeña hará que los individuos tarden en alcanzar resultados satisfactorios ya que en cada iteración se mutarán pocos genes. Sin embargo, la utilización de una tasa de mutación grande implica hacer muchos cambios en los genes de los individuos, dando lugar a un comportamiento demasiado aleatorio y caótico de la evolución de la población.

La Figura 7.7 muestra la media del valor de la función de fitness en 20 ejecuciones

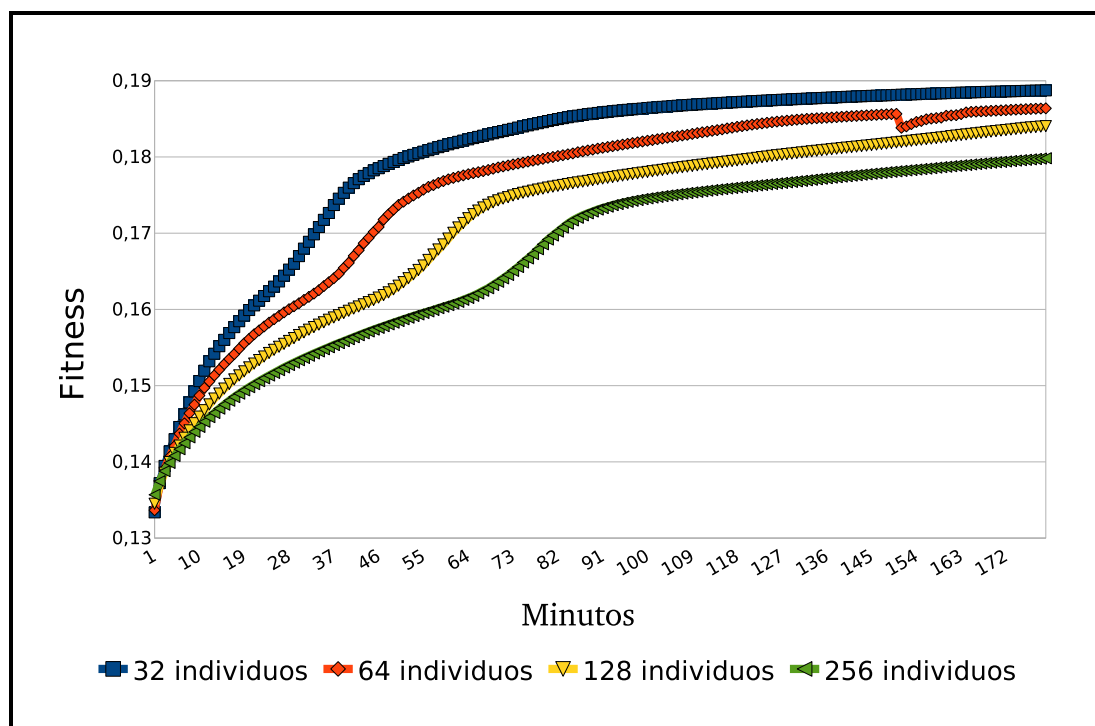


Figura 7.6: Valor de la función de fitness frente al tamaño de la población.

de 1 hora cada una, comparando configuraciones del algoritmo con tasas de mutación fijas⁵ y con los valores recomendados en la Sección 7.3.3. También se muestran configuraciones que utilizan una tasa de mutación adaptativa⁶ que empieza con un valor alto y se va reduciendo (un 75 % en estas pruebas) cada vez que el mejor valor de la función de fitness de toda la población se repite entre varias iteraciones (20 iteraciones en estas pruebas). De esta manera, cada vez que el valor de la función de fitness se estanca, lo que suele indicar que no se crean individuos mejores, se reduce el número de cambios en los genes de los individuos, consiguiendo así un ajuste más fino y una evolución menos aleatoria.

Los resultados indican que las configuraciones del algoritmo con altas tasas de mutación funcionan de forma deficiente, lo que confirma lo comentado anteriormente. Utilizando la tasa de mutación propuesta por Grefenstette ($1/L$ siendo L el tamaño del individuo, que en nuestro caso resulta un $1,3174E^{-5}$) es como se obtienen los mejores resultados. Comparando los resultados entre las tasas de mutación fijas

⁵MR por sus siglas en inglés, *Mutation Rate*

⁶AMR por sus siglas en inglés, *Adaptive Mutation Rate*

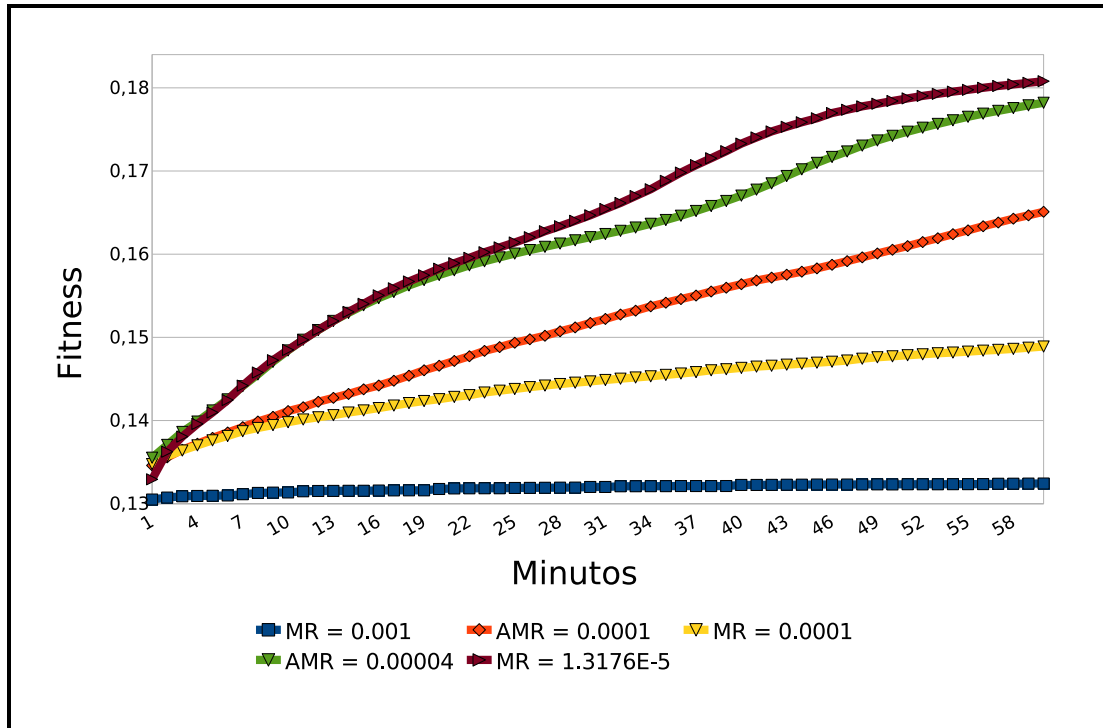


Figura 7.7: Valor de la función de fitness frente a tasas de mutación estáticas y adaptativas.

y adaptativas, a largo plazo los resultados tienden a ser iguales, pero, al menos teóricamente, el espacio de búsqueda explorado es mayor cuando se utilizan los adaptativos. Sabiendo esto, el usuario puede definir este parámetro dependiendo de sus necesidades.

La Figura 7.8 representa la varianza del conjunto de resultados obtenidos en las pruebas de tamaño de población y tasa de mutación. Los valores mostrados son bastante homogéneos, lo que prueba que el uso de la media como medida representativa de los datos es un método de medición válido.

7.5.3. Comparación entre las funciones de fitness

Como ya se ha comentado en secciones anteriores, se han implementado dos funciones para el cálculo de la compacidad: una basada en manchas y otra basada en categorías. Con la intención de estudiar los efectos que ambas funciones tienen sobre el resultado final, se ha creado un parcelario rectangular con 10.496 parcelas

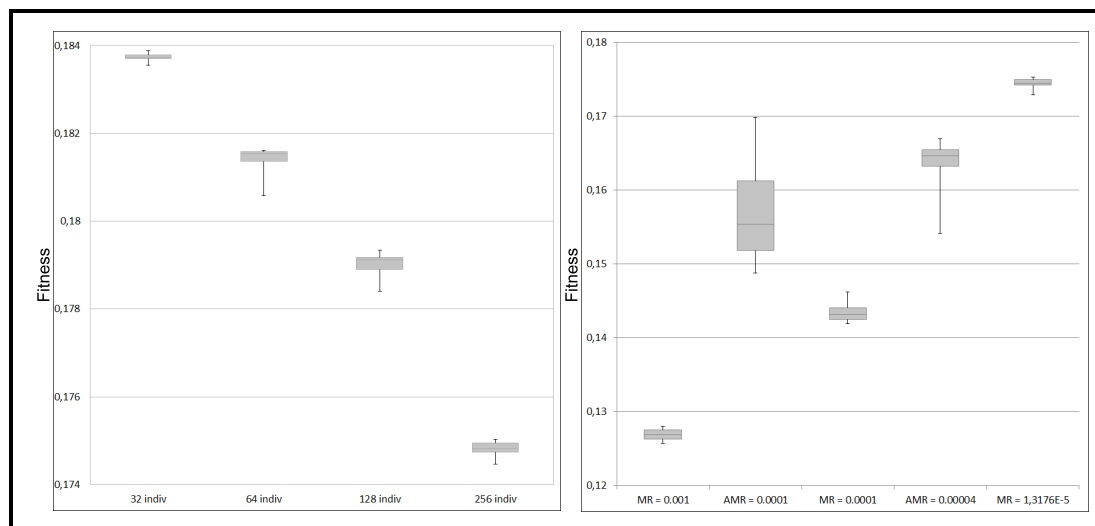


Figura 7.8: Diagramas de caja con los valores de fitness obtenidos en las pruebas del tamaño de la población (A) y la tasa de mutación (B).

de distintos tamaños. Cada parcela tiene únicamente una categoría óptima, por lo que el valor de la aptitud será de 1 cuando se le asigne dicha categoría y de 0 para cuando se le asigne alguna otra. Estas categorías han sido distribuidas de dos maneras distintas: con una distribución aleatoria (DA) y con una distribución compacta (DC). La distribución aleatoria consiste en que a cada parcela se le asigna como óptima una categoría cualquiera, y en la distribución compacta esa asignación se hace por zonas, de manera que parcelas vecinas tengan la misma categoría. La Figura 7.9 muestra la categoría óptima de cada parcela para ambas distribuciones.

Las categorías que se han utilizado para estas pruebas son las mismas cuatro que las se han utilizado en el caso real explicado en la Sección 7.5.5. Las pruebas han sido ejecutadas durante 3 horas con la versión multi-núcleo del algoritmo utilizando 16 threads en una máquina de 8 núcleos con hyperthreading.

Las Tablas 7.2 y 7.3 muestran los valores medios de las 5 ejecuciones que se han realizado para cada prueba. Las columnas indican el valor de los distintos componentes de la función de fitness. Las etiquetas de cada columna significan lo siguiente: *Apt* es el valor de la aptitud, *CC* es el valor de la compacidad por categoría, *CM* el valor de la compacidad por manchas, *NM* es el número de manchas y *Aciertos* el porcentaje de parcelas asignadas a la categoría correcta, es decir, con aptitud igual a 1. Estas etiquetas irán seguidas de *DA* o *DC* para indicar si el valor se re-

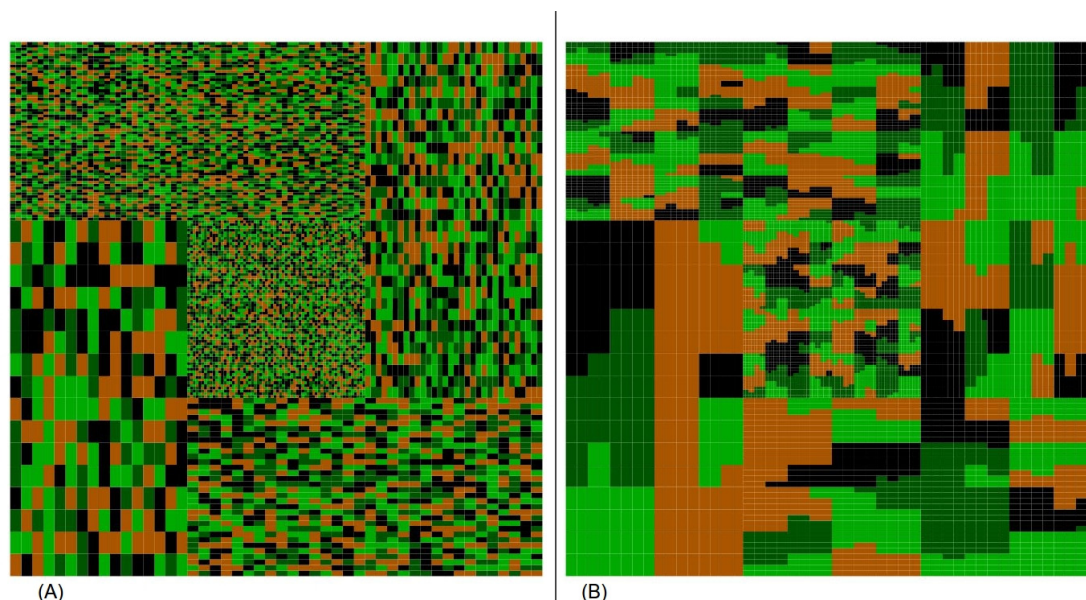


Figura 7.9: Parcelario con distribución aleatoria (A) y distribución compacta (B).

fiere a una ejecución que utiliza la distribución aleatoria o la distribución compacta respectivamente.

Las filas muestran las configuraciones con las que el algoritmo fue ejecutado: la etiqueta *Aptitud* significa que en la función de fitness sólo se tiene en cuenta la componente aptitud; las etiquetas *Comp. Cats.* y *Comp. Manchas* indican que la función de fitness sólo considera la compacidad por categorías y la compacidad por manchas, respectivamente; *Apt:CC X%:Y%* y *Apt:CM X%:Y%* indican que tanto aptitud (al X%) como compacidad (al Y%) son tenidas en cuenta; y *CC+CM X%:Y%* indica que los dos métodos de calcular la compacidad son contemplados, compacidad por categorías al X% y compacidad por manchas al Y%. Esta última configuración fue añadida después de haber ejecutado y estudiado los resultados de las anteriores ejecuciones. Viendo que la configuración CC reduce el número de manchas y que CM ayuda a hacerlas más compactas, se implementó una nueva función de fitness que integrara ambas configuraciones como dos nuevos sumandos para estudiar los efectos de los resultados obtenidos.

Observando los datos de ambas tablas nos damos cuenta de que cuando sólo se premia la aptitud, el algoritmo alcanza el 100% de aciertos (ver en la primera fila de ambas tablas las columnas Aptitud y Aciertos, respectivamente). En casos reales

Configuración	Aptitud		Comp. Cats.		Comp. Manchas	
	DA	DC	DA	DC	DA	DC
Aptitud (Apt)	1	1	0.000633	0.0138	0.6655	0.4620
Comp. Cats. (CC)	0.3005	0.6172	0.0555	0.0494	0.6392	0.5739
Apt:CC 25 %:75 %	0.3229	0.6187	0.0408	0.0505	0.6022	0.5863
Apt:CC 50 %:50 %	1	0.7051	0.000633	0.043	0.6655	0.5691
Apt:CC 75 %:25 %	1	0.9387	0.000633	0.0169	0.6655	0.502
Comp. Manchas (CM)	0.4972	0.6908	0.000816	0.0021	0.7787	0.7795
Apt:CM 25 %:75 %	0.9095	0.95	0.000604	0.0042	0.7509	0.7728
Apt:CM 50 %:50 %	0.9712	0.9716	0.000572	0.006	0.7172	0.7639
Apt:CM 75 %:25 %	0.9992	0.9841	0.000631	0.0078	0.6677	0.7481
CC+CM 50 %:50 %	0.3374	0.7636	0.0191	0.0264	0.6975	0.6330

Tabla 7.2: Comparación de los valores de los distintos componentes de la función objetivo utilizando distribución aleatoria y compacta (I).

Configuración	Número de manchas		Aciertos	
	DA	DC	DA	DC
Aptitud (Apt)	5387	165	100 %	100 %
Comp. Cats. (CC)	143	65	28.14 %	52.07 %
Apt:CC 25 %:75 %	110	68	29.53 %	52.85 %
Apt:CC 50 %:50 %	5386	73	99.99 %	61.41 %
Apt:CC 75 %:25 %	5387	137	100 %	92.73 %
Comp. Manchas (CM)	6244	4030	48.70 %	56.39 %
Apt:CM 25 %:75 %	7065	2747	78.45 %	83.44 %
Apt:CM 50 %:50 %	6934	1769	91.04 %	89.60 %
Apt:CM 75 %:25 %	5445	1153	99.76 %	93.79 %
CC+CM 50 %:50 %	327	144	33.20 %	72.22 %

Tabla 7.3: Comparación de los valores de los distintos componentes de la función objetivo utilizando distribución aleatoria y compacta (y II).

la aptitud de las parcelas no es uno o cero, si no que tomará valores intermedios, por lo que es más complicado llegar a un 100 % de acierto.

Con la distribución aleatoria de las aptitudes óptimas y utilizando solamente la compacidad por categorías (segunda fila), el número de manchas obtenidas en cada ejecución puede cambiar considerablemente (en las pruebas se obtuvieron desde 53 a 213, media de 143). Esto ocurre porque cada ejecución comienza con una solución aleatoria, y los grupos de categorías formados inicialmente difiere entre ejecuciones. Como en este caso la aptitud es ignorada, la evolución de cada individuo puede ser muy diferente y el número de aciertos muy bajo. Si añadimos un 25 % de aptitud, la varianza es reducida (de 73 manchas a 211) y el número de aciertos mejora ligeramente. Con una relación de 50 %-50 % entre aptitud y compacidad por categorías, el algoritmo ya casi alcanza el 100 % de aciertos. Utilizando la misma configuración pero con una distribución compacta de las aptitudes óptimas, los valores de compacidad bajan cuanto más porcentaje de aptitud se utilice en la función de fitness, aunque más aciertos se producen y el número de manchas está más cercano al número de manchas que tiene el parcelario con las categorías óptimas asignadas a las parcelas (que es el mismo que podemos ver en las primera filas de las tablas).

Cuando se utilizan las configuraciones de compacidad por manchas, el algoritmo tiende a crear soluciones con muchas manchas. Si se combinan ambos métodos para el cálculo de la compacidad (última fila) se obtiene una buena relación entre CC, CM y NM ya que la compacidad por categorías tiende a crear menos manchas y la compacidad por manchas ayuda a que la solución sea más compacta. Así pues, si el objetivo es, por ejemplo, obtener el mayor número de aciertos, es más importante asignar la categoría correcta a la parcela que alcanzar un buen grado de compacidad, por lo que es aconsejable subir el porcentaje de la aptitud en la función de fitness. En los resultados obtenidos con una distribución aleatoria de las aptitudes óptimas, el número de manchas obtenidos varía desde 104 hasta 419. Sin embargo con la distribución compacta, esta varianza es reducida de 136 a 160 manchas.

Para observar estas conclusiones de manera gráfica, las figuras 7.10 hasta la 7.19 muestran los resultados obtenidos al ejecutar el algoritmo sobre las distribuciones aleatorias y compactas⁷.

⁷Mientras que las tablas muestran los valores medios de las 5 ejecuciones realizadas, las figuras muestran sólo una de ellas escogida aleatoriamente.

La Figura 7.10 es exactamente igual a la 7.9 ya que sólo se premia la aptitud, no la compacidad, y el número de aciertos en estas ejecuciones es del 100 %. En lo que respecta a la compacidad por categorías, las figuras 7.11 y 7.12 muestran formas más compactas, aunque la que empieza con la distribución compacta parte *con cierta ventaja* y por ello tiene menos manchas y mayor número de aciertos ya que la aptitud no es tenida nada, o casi nada, en cuenta. Cuando la aptitud toma una mayor importancia, los aciertos aumentan en detrimento de la compacidad y del número de manchas, como se puede ver en las figuras 7.13 y 7.14.

Por otra parte, cuando se utiliza la compacidad por manchas (figuras 7.15, 7.16, 7.17, 7.18) los resultados que se obtienen se caracterizan por crear muchas manchas (en el caso de empezar con una distribución aleatoria se crean más manchas que de las que existen en la distribución inicial) pero muy compactas, esto es, con un alto valor en CM.

Estudiando estos resultados y sus características se decidió emplear una configuración que en principio no estaba prevista, que es la utilización de ambas estrategias de medir la compacidad. El resultado se puede ver en la Figura 7.19. Esta configuración es la que consigue mejor relación entre todos los parámetros contemplados. Comparando ambas soluciones, la que fue ejecutada sobre el parcelario con distribución compacta tiene más facilidades en llegar a una solución con pocas manchas y con muchos aciertos, lo que por norma general es lo que se pretende.

Estas pruebas fueron ejecutadas sobre parcelarios con distribución de categorías aleatorias y compactas y, aunque los parcelarios reales suelen tener esta última distribución, los resultados expuestos en este capítulo serán de gran ayuda a los usuarios finales del algoritmo ya que podrán decidir qué pesos asignar a cada parámetro de la función de fitness dependiendo del objetivo deseado: gran grado de compacidad o aptitud, pocas manchas o muchos aciertos.

7.5.4. Rendimiento del algoritmo paralelo

Para conocer si el rendimiento del algoritmo paralelo es suficientemente bueno, hemos tenido en cuenta dos métricas: el speed-up (S) y la eficiencia (E). El speed-up es una medida para saber cuánto de rápido es un algoritmo paralelo con respecto

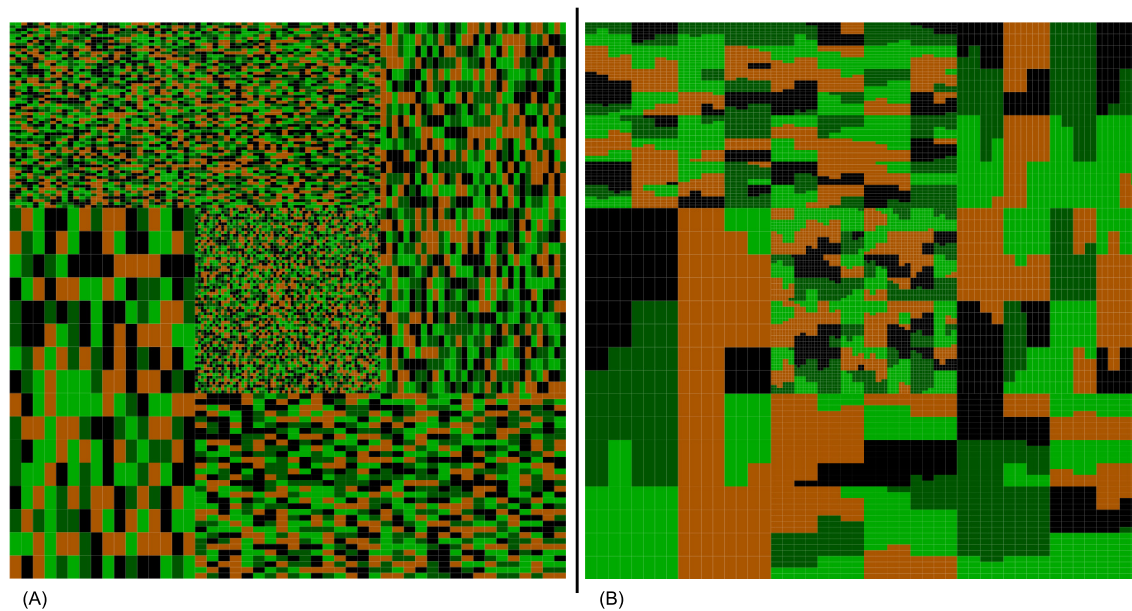


Figura 7.10: Solución obtenida para la configuración *Apt*: distribución aleatoria (A) y distribución compacta (B).

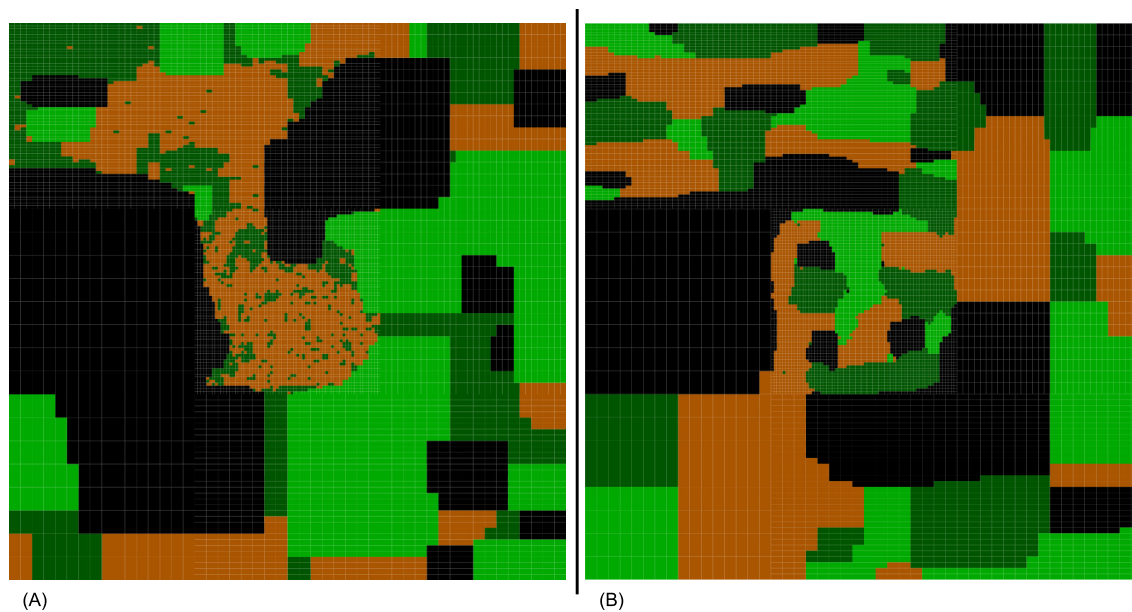


Figura 7.11: Solución obtenida para la configuración *CC*: distribución aleatoria (A) y distribución compacta (B).

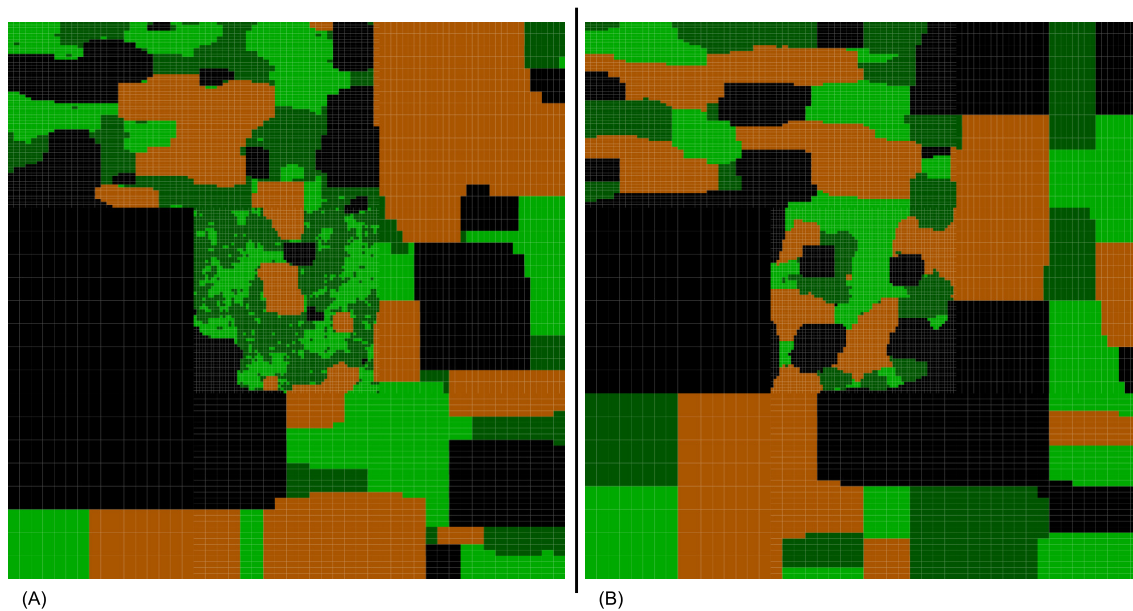


Figura 7.12: Solución obtenida para la configuración *Apt:CC 25 %-75 %*: distribución aleatoria (A) y distribución compacta (B).

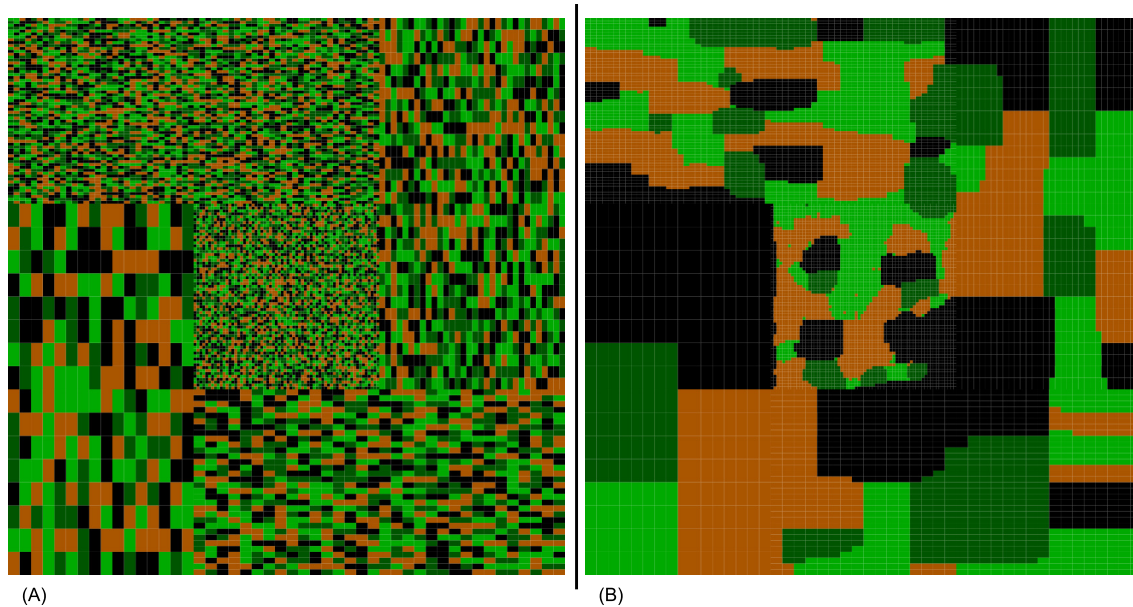


Figura 7.13: Solución obtenida para la configuración *Apt:CC 50 %-50 %*: distribución aleatoria (A) y distribución compacta (B).

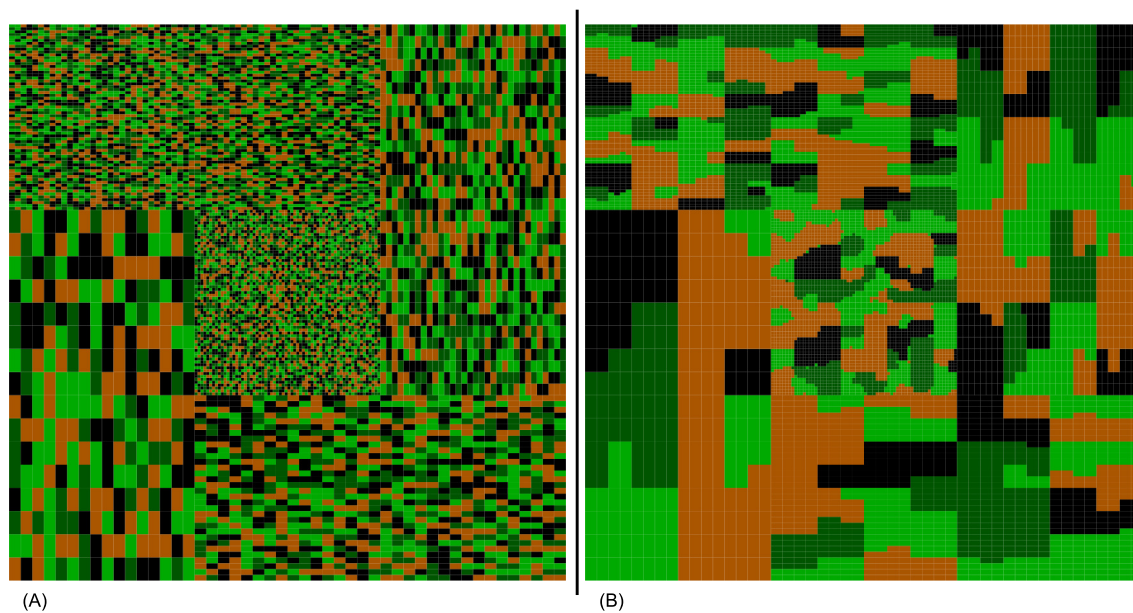


Figura 7.14: Solución obtenida para la configuración *Apt:CC* 75 %-25 %: distribución aleatoria (A) y distribución compacta (B).

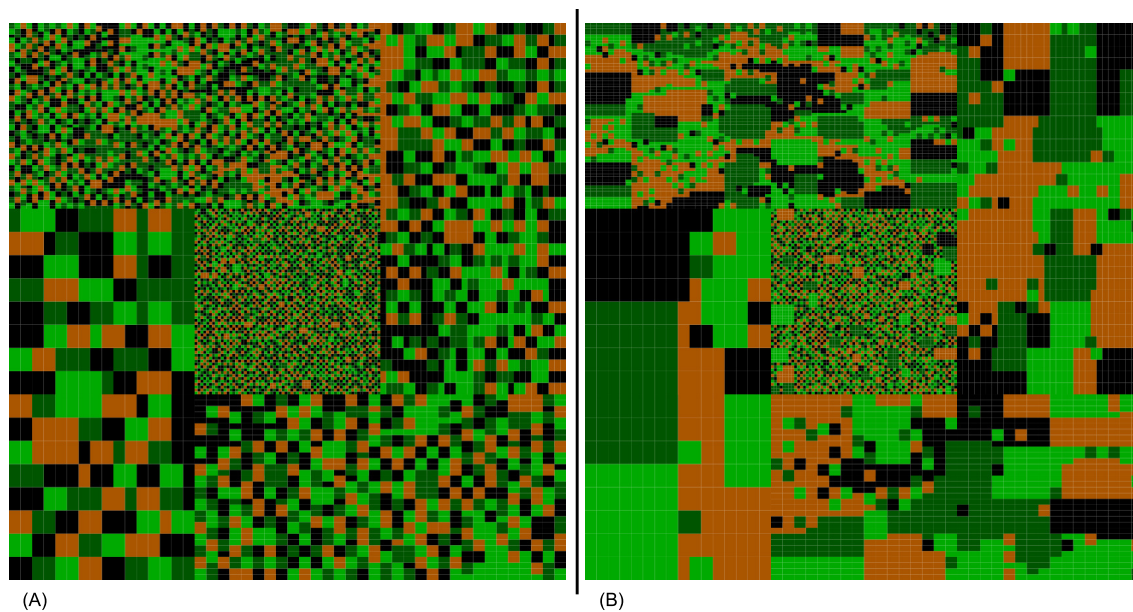


Figura 7.15: Solución obtenida para la configuración *CM*: distribución aleatoria (A) y distribución compacta (B).

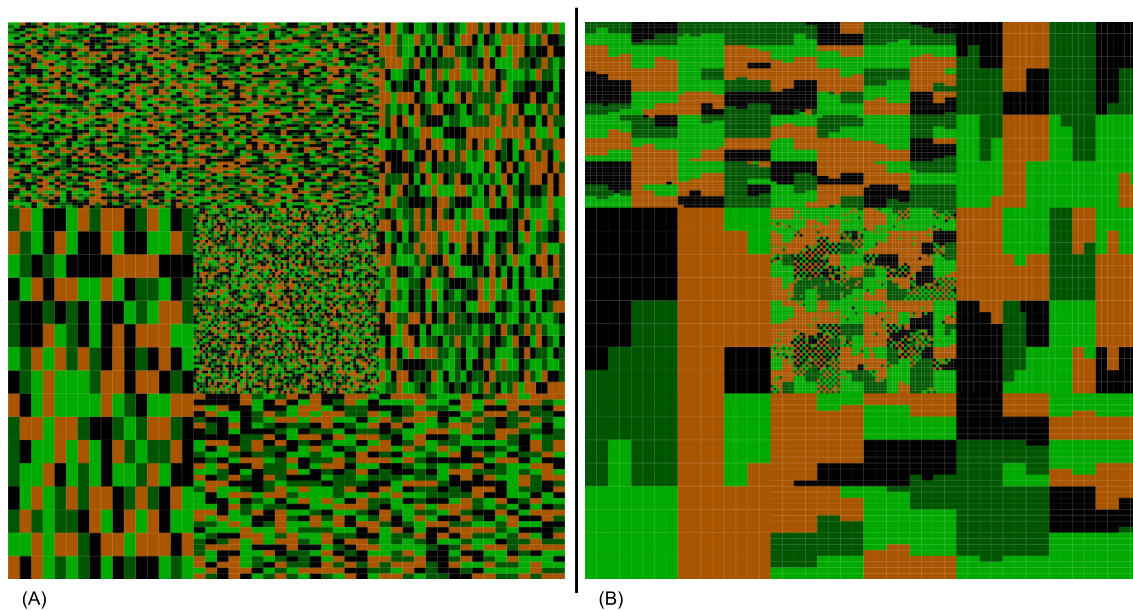


Figura 7.16: Solución obtenida para la configuración *Apt:CM 25 %-75 %*: distribución aleatoria (A) y distribución compacta (B).

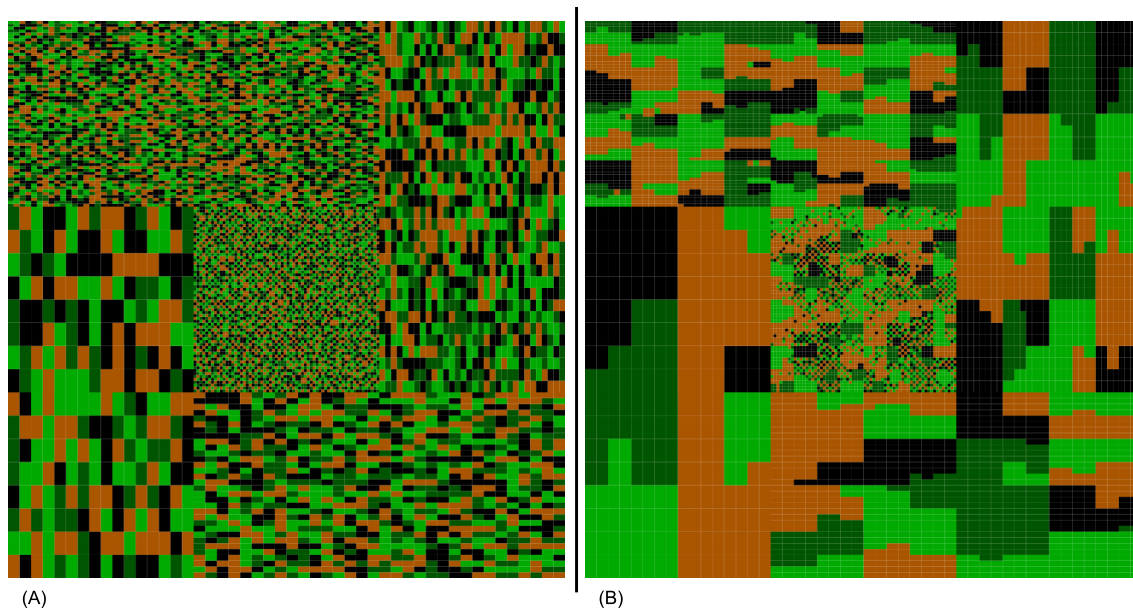


Figura 7.17: Solución obtenida para la configuración *Apt:CM 50 %-50 %*: distribución aleatoria (A) y distribución compacta (B).

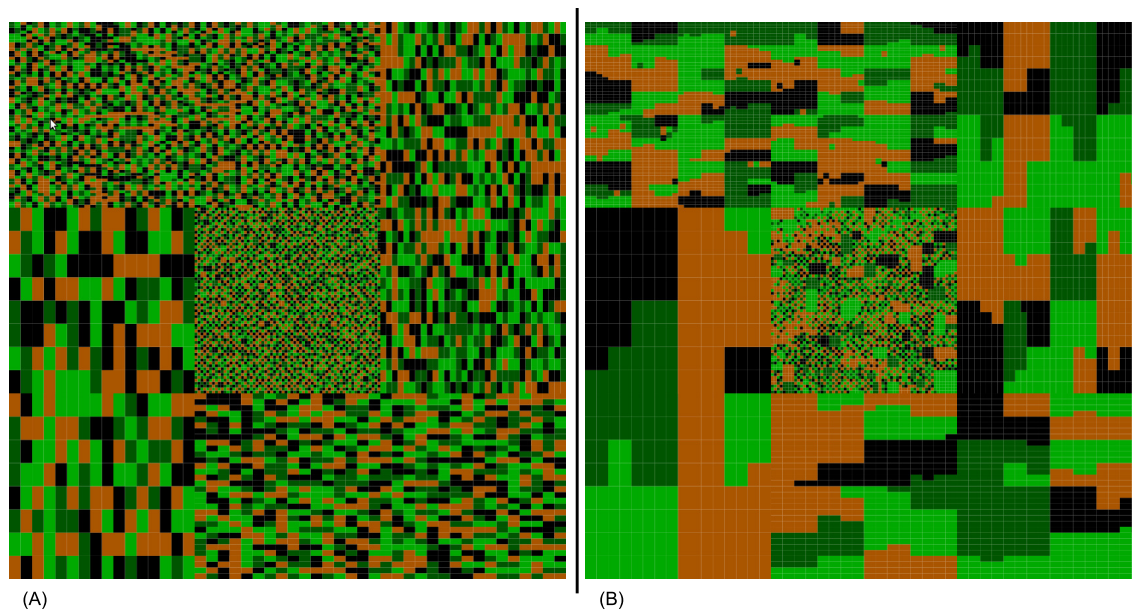


Figura 7.18: Solución obtenida para la configuración *Apt:CM* 75 %-25 %: distribución aleatoria (A) y distribución compacta (B).

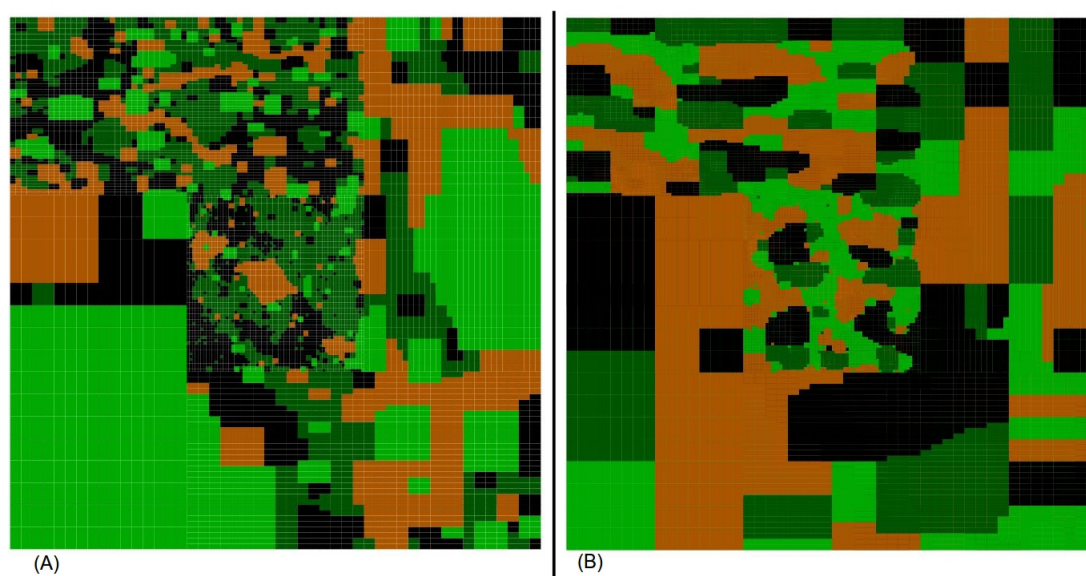


Figura 7.19: Solución obtenida para la configuración *CC+CM*:50 %-50 %: distribución aleatoria (A) y distribución compacta (B).

a su versión secuencial (ver Ecuación 7.6). Por su parte, la eficiencia se obtiene dividiendo el speed-up entre el número de procesos paralelos creados, dando lugar a una medida que nos da idea de lo bien o mal que están siendo utilizados los recursos (ver Ecuación 7.7). La formulación clásica de speed-up y eficiencia es la siguiente:

$$S_p = \frac{T_1}{T_p} \quad (7.6)$$

$$E_p = \frac{S_p}{p} \quad (7.7)$$

siendo p el número de procesos paralelos utilizados, T_1 el tiempo de ejecución del algoritmo secuencial, y T_p el tiempo de ejecución del algoritmo paralelo sobre p procesos paralelos.

La Figura 7.20 muestra la evolución del speed-up para las versiones paralelas multi-núcleo con respecto a la versión secuencial (ejecución en un único thread deshabilitando el hyper-threading), y la Figura 7.21 muestra el speed-up (A) y la eficiencia (B) en la iteración 20.000 del bucle genético. Los valores son la media de las 20 ejecuciones realizadas por prueba.

Como se puede observar en la Figura 7.21, las ejecuciones paralelas tienen un speed-up super-lineal mientras el número de threads utilizados por el algoritmo genético paralelo sea menor o igual al número de núcleos que tiene la máquina (4 en este caso). Esto es debido a una mejor utilización de la jerarquía de memoria ya que cada thread debe lidiar con menos datos, y aunque el speed-up todavía aumenta cuando se utilizan un número de threads superior al número de núcleos disponibles, la eficiencia decrece. A pesar de que el hyper-threading disponible en los procesadores E5520 permite ejecutar más de un thread simultáneamente por núcleo, los threads comparten algunos recursos, lo que implica que el rendimiento real del hyper-threading aporte sólo entre un 30 % y un 40 % respecto a la versión *single-threaded*.

El speed-up de la versión del algoritmo paralelizado para clúster no es mostrado ya que dicha paralelización busca alcanzar los mayores valores de fitness en tiempos de ejecución más cortos. El speed-up, en este caso, lo marca el nodo más lento del clúster.

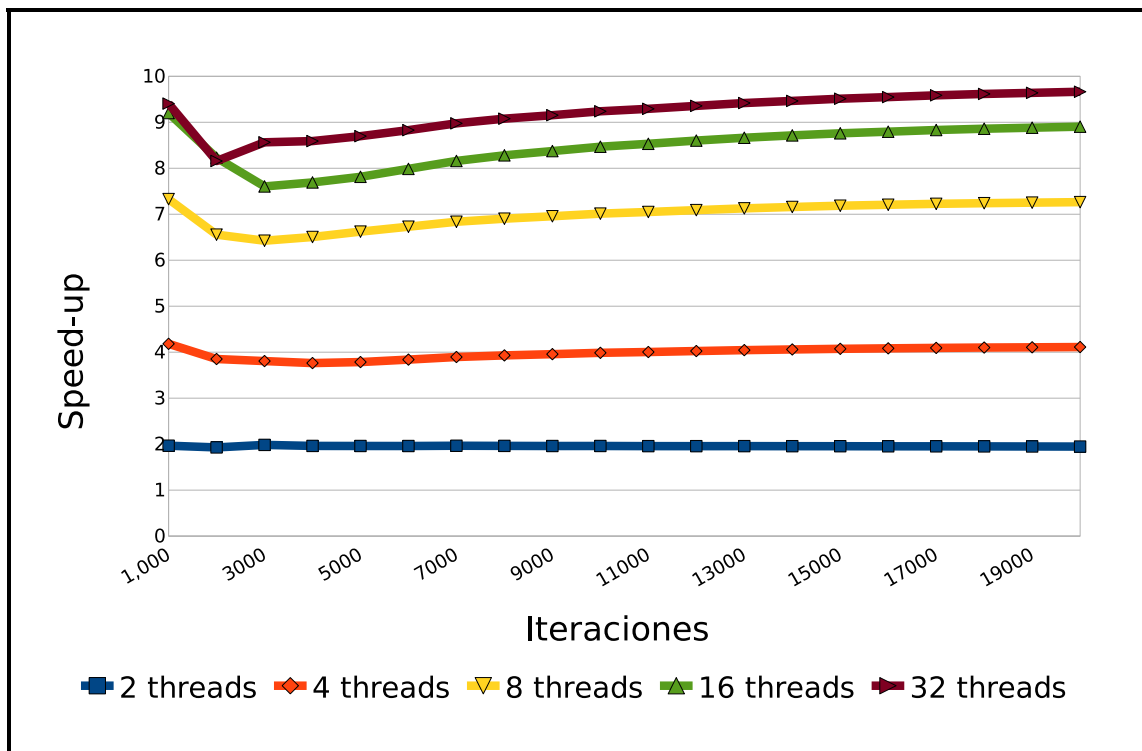


Figura 7.20: Speed-up para las versiones paralelas multi-núcleo con respecto a la versión secuencial.

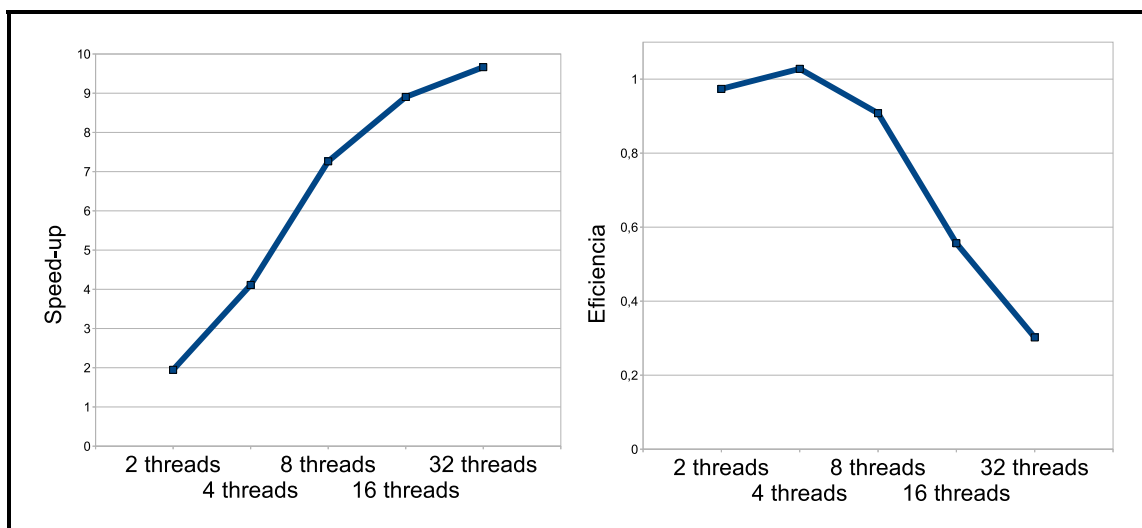


Figura 7.21: Speed-up (A) y eficiencia (B) en la iteración 20.000 de la versión paralela multi-núcleo con respecto a la versión secuencial.

La Figura 7.22 muestra los valores medios de fitness de la versión híbrida con respecto a los de la versión multi-núcleo. Para estas pruebas se han utilizado 16 threads por nodo. Las ejecuciones de la versión híbrida han sido realizadas sobre dos, cuatro y ocho nodos del clúster con diferentes tiempos de sincronización⁸, dando todas ellas mejores resultados que la versión paralela multi-núcleo.

En la gráfica se ha realizado un zoom para poder comprobar en detalle los valores de la función de fitness en los últimos momentos de ejecución del algoritmo. Si se observa sólo la versión paralela híbrida, se puede ver que el mejor resultado fue obtenido utilizando la configuración de 2 nodos con 15 minutos entre sincronizaciones, y la de 4 nodos con 6 minutos. Las curvas del valor de fitness de las configuraciones de 2 nodos tienden a estancarse. Por contra, las curvas de las configuraciones de 4 nodos (especialmente la que utiliza 6 minutos entre sincronizaciones) tienen una pendiente más pronunciada. Sin embargo, las diferencias entre los valores de fitness de las distintas configuraciones son pequeñas. Las ejecuciones con 8 nodos consiguen un fitness menor que las otras, pero en ejecuciones a largo plazo pueden alcanzar mejores resultados. Como se ha mencionado antes, el criterio de parada está basado en tiempo, y estas pruebas fueron ejecutadas durante 3 horas, tiempo que se puede considerar como corto plazo teniendo en cuenta el tamaño de los individuos.

7.5.5. Simulación de un caso real sobre el municipio de Guitiriz

La Figura 7.23 muestra la solución propuesta por el algoritmo para el municipio gallego de Guitiriz. Como ya se ha comentado, este municipio tiene 138.175 parcelas, 52.045 de las cuales tienen una categoría fija. Sabiendo que el parcelario de Guitiriz tiene cierta compacidad en sus aptitudes óptimas y siguiendo las conclusiones sacadas de las tabla 7.2 y 7.3, las pruebas fueron ejecutadas utilizando una configuración de Apt:CC 50 %:50 % para conseguir una buena relación entre el número de manchas y aciertos. Los distintos colores de la figura nos muestran que las manchas tienen la regularidad poligonal que se esperaba. Las manchas blancas se corresponden con categorías fijas, como orillas de ríos o carreteras. Las otras manchas incluyen el resto

⁸Se recuerda que la sincronización consiste en el envío de los mejores individuos por parte de los procesos esclavos al máster.

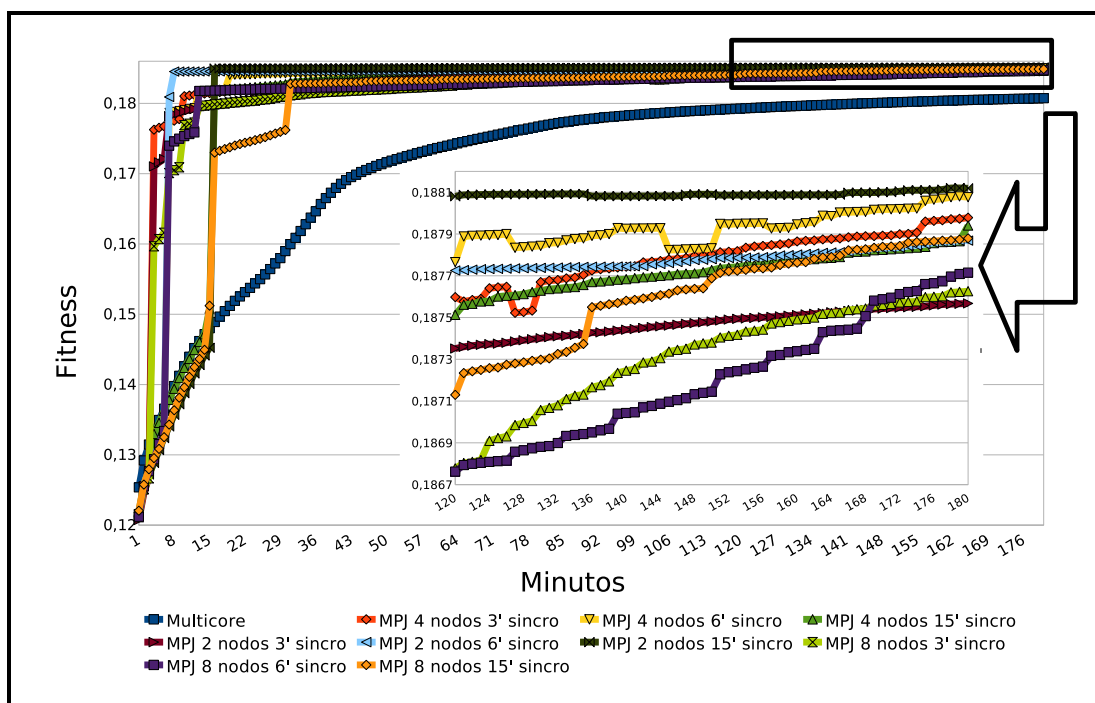


Figura 7.22: Valor de la función de fitness para las versiones paralelas multi-núcleo e híbrida del algoritmo genético.

de parcelas, tanto fijas como no fijas.

Por su parte, la Figura 7.24 compara la solución generada por el algoritmo genético (A) y la propuesta por expertos en problemas de planificación (B). La solución (A) tiene más manchas blancas (categorías fijas) porque los expertos no siempre tienen en cuenta todas las categorías fijas. Aparte de esta consideración, las dos soluciones tienen bastante similitudes. La solución propuesta por el algoritmo genético paralelo híbrido puede ser un muy buen punto de inicio para el experto encargado de la planificación de un municipio.

7.6. Funcionalidades adicionales

Algunas funcionalidades han sido incluidas en el algoritmo para facilitar su uso a los usuarios y para ayudarles y proteger sus datos ante posibles fallos del sistema.

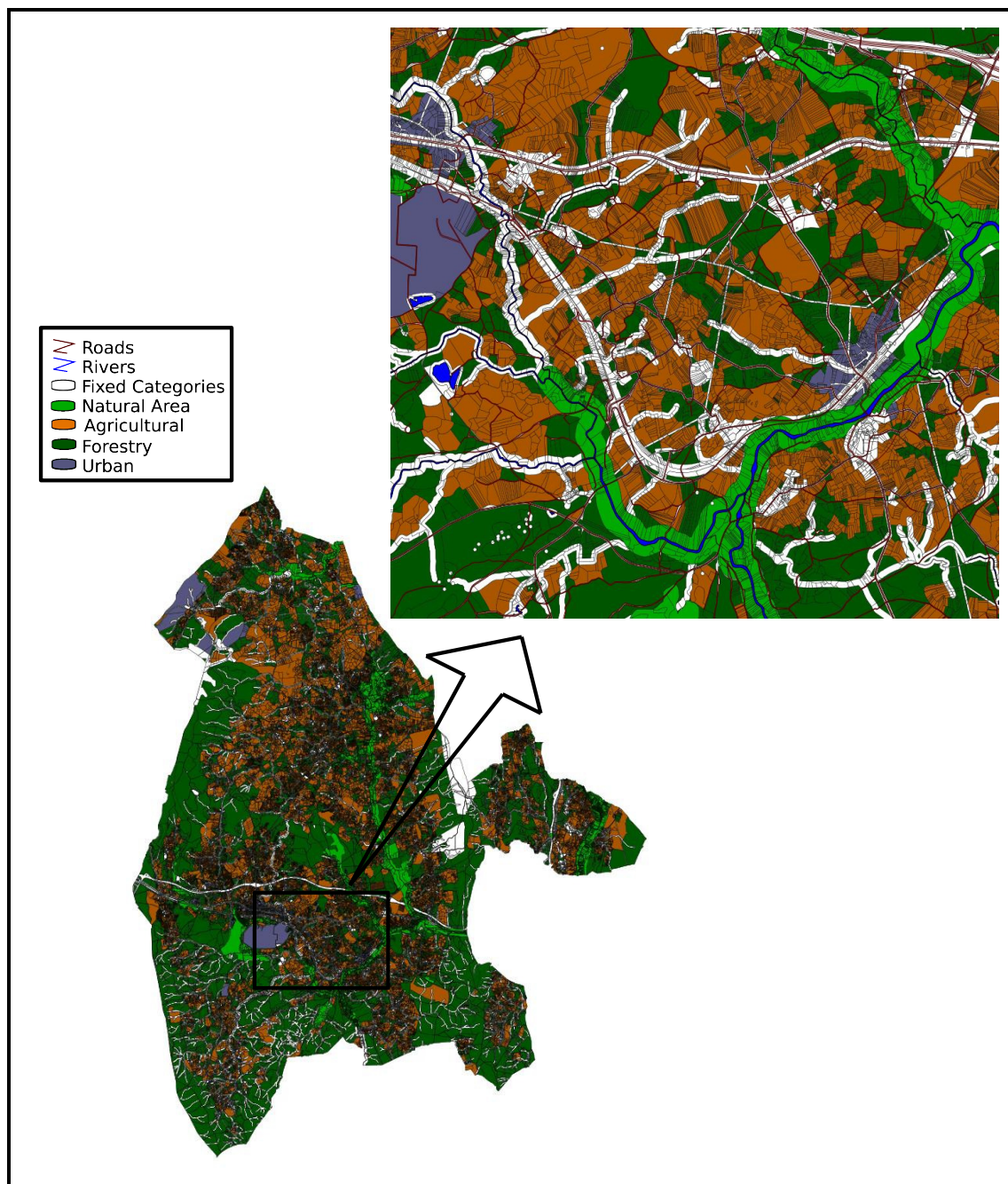


Figura 7.23: Resultado propuesto por el algoritmo en un caso real.

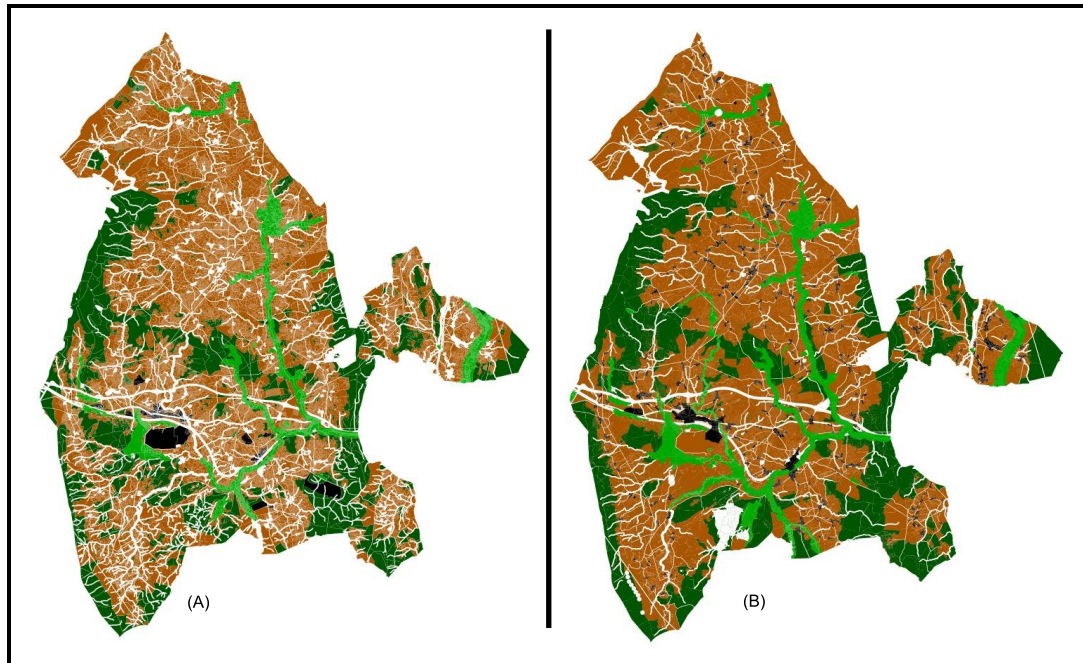


Figura 7.24: Comparación entre la solución propuesta por el algoritmo (A) y la solución propuesta por expertos (B).

7.6.1. Checkpointing

Debido a las características del problema a resolver, las ejecuciones pueden ser muy largas. Si ocurriese algún fallo en la máquina donde se ejecuta el algoritmo se podrían perder datos que tanto tiempo han costado calcular. Para evitar esto se ha implementado un mecanismo para que ante caídas del sistema, se pueda continuar la ejecución del algoritmo desde un estado anterior próximo al estado en el que estaba cuando sucedió el incidente.

Todos los individuos de la población, incluido su valor de la función de fitness y algunos datos más de control, son guardados en disco periódicamente. De esta manera, si el algoritmo detiene su ejecución por cualquier causa, este puede ser relanzado con los últimos datos guardados. Dependiendo de la versión del algoritmo que se esté ejecutando (secuencial o paralela), el checkpointing se implementa de diferentes formas, aunque en todas ellas el usuario es quien especifica cada cuanto tiempo se guardan los datos. Cada vez que se finaliza una iteración del bucle genético se comprueba si es momento de guardar los datos. En las versiones clúster e híbrida

del algoritmo paralelo, es el nodo máster el encargado de hacer esta comprobación y de guardar los datos en caso necesario, y lo hará justo después de cada sincronización.

Esta funcionalidad tiene además una ventaja añadida, y es que no sólo es útil ante caídas del sistema, sino que también permite la modificación de la configuración del algoritmo. Cuando se relanza una ejecución parada se puede hacer con parámetros de entrada distintos a los originales. Por ejemplo, se puede disminuir o aumentar la tasa de mutación, ejecutar el algoritmo sobre más threads, e incluso utilizar más o menos nodos de un clúster. De esta manera se pueden hacer cambios dinámicos en la configuración del algoritmo, consiguiendo así un espacio de búsqueda más amplio. También se pueden hacer pruebas en los parámetros de entrada y si unos resultados no son los esperados, volver atrás y probar con otros.

Como la cantidad de información a guardar puede ocupar espacio, estos archivos son comprimidos por el propio mecanismo de checkpointing.

7.6.2. Interfaces de usuario

Para implementar el algoritmo se ha utilizado una herramienta Java de código abierto para análisis geoespacial llamada SEXTANTE [98], cuyo principal objetivo es ser una plataforma de desarrollo de geoalgoritmos tanto para la implementación como para el uso de los mismos. SEXTANTE puede ser utilizado en software SIG de escritorio como uDig, openJUMP o gvSIG. Dispone para ello de una interfaz gráfica pero también permite ejecutar los algoritmos mediante línea de comandos, funcionalidad esta muy útil para entornos que no dispongan de entorno gráfico. Así pues, el algoritmo genético secuencial y la versión paralela multi-núcleo pueden ser ejecutados a través de la interfaz gráfica y la versión paralela clúster mediante la línea de comandos, ya que en esos entornos normalmente no te permiten lanzar aplicaciones gráficas. La Figura 7.25 muestra la interfaz gráfica del algoritmo genético sobre gvSIG.

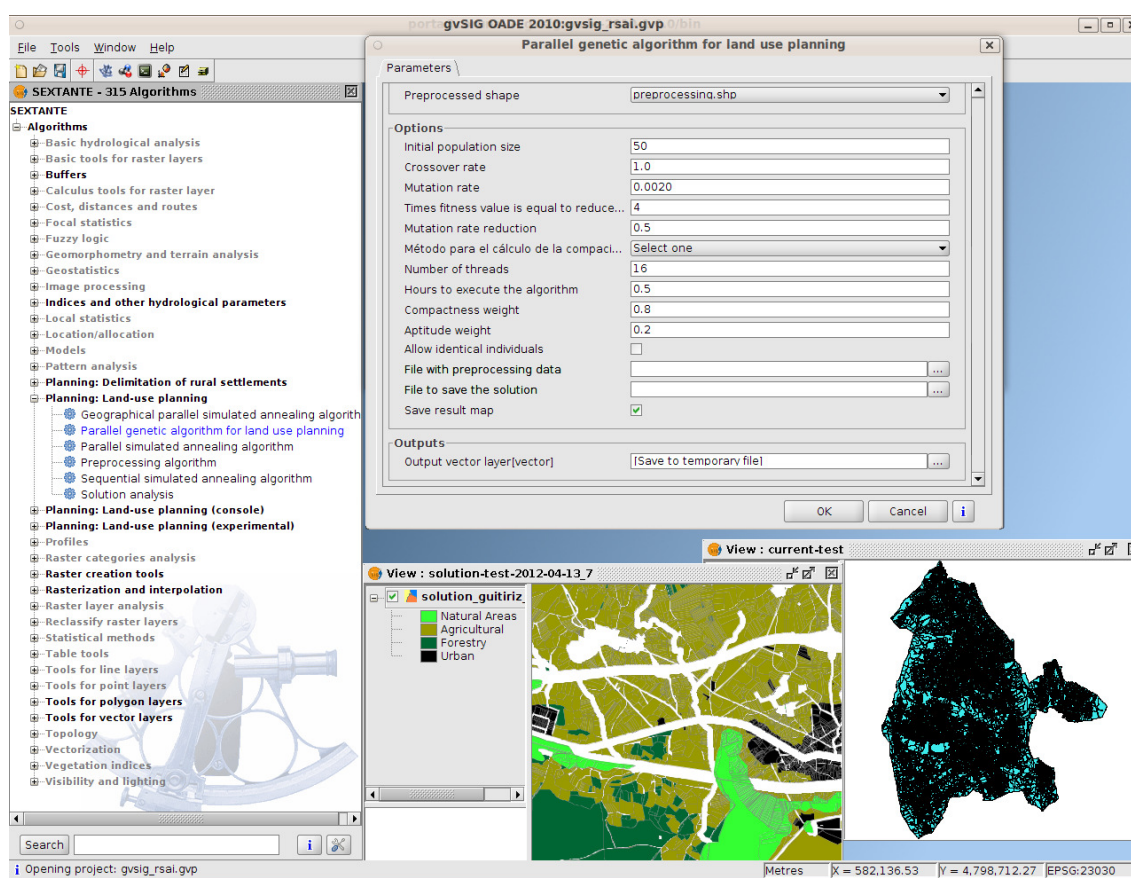


Figura 7.25: Interfaz gráfica de gvSIG con SEXTANTE.

7.7. Resumen

Todo el proceso de desarrollo (diseño, implementación y pruebas incluidas) del algoritmo genético descrito, ha permitido sacar varias conclusiones. La primera es que los algoritmos genéticos paralelos son una buena elección para resolver problemas de planificación territorial donde sean muchas las posibles combinaciones entre categorías y parcelas. Otra conclusión es que la paralelización multi-núcleo del algoritmo alcanza un speed-up lineal cuando los recursos físicos (núcleos de procesador) lo permiten, y que la paralelización clúster permite la ejecución paralela de varias instancias del algoritmo con intercambio periódico de datos entre sí. Una versión paralela híbrida también fue implementada para el aprovechamiento máximo de clústers con nodos multi-núcleo.

Para probar el buen comportamiento del algoritmo se han ejecutado varias pruebas utilizando distintos tipos de parcelario, distintos parámetros de entrada, distintos pesos para los componentes de la función de fitness, y en general cambiando cualquier parámetro que pudiese influir en la solución final.

Los resultados de las pruebas también han servido para demostrar la eficacia del algoritmo y para elaborar las tablas 7.2 y 7.3, en las que los expertos pueden fijarse para elegir los parámetros de entrada que ayuden a alcanzar los objetivos marcados. Las soluciones aportadas por el algoritmo suponen un muy buen punto de partida para los procesos de ordenación.

Otra conclusión es que utilizando tipos de datos básicos y estructuras de datos no complejas, Java es un lenguaje eficiente en cuanto a consumo de recursos y con un amplio soporte para la programación paralela.

Este proyecto ha dado lugar a las publicaciones y participaciones en congresos referenciadas en [105, 108, 124].

Capítulo 8

Aportaciones y trabajo futuro

La tierra es un factor estratégico de suma importancia en la prosperidad de una determinada región. Este hecho cobra especial importancia en aquellas zonas donde la tierra sirve para algo más que para construir en ella, en aquellas zonas donde el trabajo de la tierra es vital para obtener el máximo aprovechamiento de todo el potencial económico que esta puede dar. Estamos hablando del medio rural, entornos donde la tierra no sólo es clave para la subsistencia de sus habitantes sino que también ha de ser considerada como un motor económico capaz de impulsar el crecimiento de una región.

Hasta ahora muchas de las actividades relacionadas con la Administración de Tierras se venían haciendo de una manera que dista mucho de ser la óptima. Muchas planificaciones se hacían (y se hacen) con herramientas no preparadas para el manejo de información con carácter espacial. En procesos de planificación o demarcación es común encontrar a personal trabajando con programas CAD, más enfocados al diseño, en vez de utilizar software SIG, que es el ideal para gestionar la información geográfica que este tipo de procesos requieren. Otra situación habitual es que mucha de la información georreferenciada de la disponen los organismos, tanto públicos como privados, está almacenada en simples hojas de cálculo en vez de estar guardada en sistemas de persistencia adecuados, como pueden ser las bases de datos con extensión espacial.

Es interesante por tanto dotar a los órganos responsables de la gestión de la tierra (en general la Administración de Tierras) de herramientas con capacidades

espaciales que optimicen procesos que hasta ahora no se hacían de manera eficiente. Muchos de estos procesos son tan complejos que el software SIG genérico no es suficiente y se requieren herramientas diseñadas específicamente para la resolución de dichos problemas.

En esta tesis se han desarrollado una serie de herramientas que están enfocadas a la optimización de procesos característicos del medio rural, como son la gestión de elementos con información geográfica aplicada a los mercados de tierras y los procesos de planificación y ordenación territorial.

8.1. Gestión de elementos con información geográfica

Muchos de los procesos que tienen que ver la gestión de elementos con información espacial (como parcelas, edificaciones, entorno, etc.) tienen necesidades similares, por lo que al abstraerlas es posible la creación de una herramienta genérica que pueda agruparlas y resolverlas. Por otra parte, la participación de los propietarios de las parcelas involucradas en cualquier tipo de gestión es indispensable, por lo que la herramienta genérica debe estar accesible para todos los actores participantes en esos procesos. Teniendo esto en cuenta, en el Capítulo 3 se propone una arquitectura Web-SIG que puede ser utilizada como base para la implementación de sistemas de información que gestionen información geográfica y que estén accesibles vía web.

Aplicando esta idea se creó SITEGAL, un innovador sistema de información Web-SIG, pieza central del Banco de Tierras de Galicia y explicado con detalle en el Capítulo 4. Abanderado de la evolución hacia el e-Government de Galicia, SITEGAL gestiona información relacionada con parcelas, propietarios, e información georreferenciada de los elementos que están en su entorno. También gestiona los procesos de alquiler y arrendamiento entre propietarios que quieren sacar un beneficio económico de las tierras que tiene paradas, y agricultores que quieren trabajarlas para ampliar sus explotaciones o crear unas nuevas. El objetivo es dinamizar un mercado de tierras que en los últimos años se ha estancado de manera preocupante en la comunidad gallega. SITEGAL, presentado en sociedad en noviembre de 2007 y que a día de hoy aún se le siguen añadiendo funcionalidades, ha sido implementa-

do utilizando únicamente software libre y empleando servicios estándares como los marcados por la OGC.

SITEGAL está en continua evolución y son varias las ideas para desarrollar en un futuro: (1) permitir la inserción de lotes de parcelas para que los propietarios puedan insertar en los fondos del sistema varias parcelas en un sólo paso; (2) añadir parcelas de propietarios desconocidos; y (3) extender la herramienta para incorporar funcionalidades que permitan la gestión de los procesos de concentración parcelaria.

8.2. Procesos de demarcación y ordenación territorial

Los procesos de planificación y ordenación territorial son arduos, laboriosos y en los que están en juego muchos intereses. La Parte III de la tesis se centra en la delimitación de núcleos rurales, y en la asignación de categorías del suelo a parcelas en procesos de ordenación.

En el Capítulo 6 se explica el funcionamiento de un algoritmo para la delimitación de núcleos rurales. Este algoritmo proporciona soluciones aplicando la normativa vigente en los procesos de demarcación y también soluciones obtenidas tras aplicar la nueva metodología propuesta. El algoritmo desarrollado es de tipo *greedy iterativo basado en población* (PBIG), una clase de algoritmo de búsqueda heurística que amplía a los *greedy iterativos* (IG) y cuya utilización es relativamente novedosa.

Mediante la búsqueda de soluciones de manera heurística y manejando un conjunto de soluciones (por eso se le denomina *basado en población*), este algoritmo es capaz de proporcionar resultados válidos y de suficiente calidad para el usuario, que son ideales para ser considerados como un excelente punto de partida para un posterior refinamiento de las soluciones por parte de los expertos en planificación. Con la nueva metodología propuesta, más adaptativa a la morfología de los núcleos, se han alcanzado los mejores resultados en las pruebas realizadas sobre núcleos rurales con diferentes características. El algoritmo, que es muy configurable, permite al usuario explorar múltiples vías utilizando las normas actualmente vigentes y/o aplicando la nueva metodología. Además todos los resultados obtenidos pueden ser guardados y

recuperados posteriormente para realizar comparaciones o ver la evolución de una delimitación concreta.

Como trabajo futuro respecto al algoritmo para la demarcación de núcleos rurales es la implementación de nuevos métodos para el cálculo de la densidad edificatoria (como los nuevos métodos propuestos por la Instrucción 4/2011 expuesta en la Sección 6.1.1), tarea muy interesante y de gran utilidad para los expertos. Además, como este algoritmo utiliza métodos para la validación y evaluación de las soluciones, estos podrían ser externalizados para permitir la creación de herramientas capaces de validar y evaluar diferentes soluciones, que pueda comparar las modificaciones realizadas a mano por los expertos de manera rápida y simple, y que incluso detecte qué restricciones fallan en delimitaciones no válidas, tarea que normalmente es muy complicada de hacer de manera visual. Otra cuestión interesante sería dotar al algoritmo de la capacidad de delimitar múltiples zonas rurales a la vez. Esto no es trivial, ya que la delimitación de una zona puede afectar directamente a las vecinas. Otra buena práctica sería poner este algoritmo disponible a través de Internet y evitar así que el usuario tenga que tenerlo instalado en su máquina local. Una de las muchas ventajas de proporcionar esta funcionalidad es que las actualizaciones del algoritmo tendrían efecto inmediato para todos los usuarios. Además, utilizando estándares de la OGC como WMS, WFS o WPS, el usuario ni siquiera tendría que tener los datos en local, sino que podrían estar en servidores remotos. Por último, para aumentar las capacidades computacionales del algoritmo, se podría utilizar la computación de altas prestaciones (HPC, de *High Performance Computer*) como se hace en [105]. La utilización de la HPC podría ayudar a generar más soluciones en el mismo periodo de tiempo y así explorar un espacio de búsqueda más amplio.

En el Capítulo 7 se ha detallado el desarrollo de un algoritmo genético para la asignación de categorías del suelo a parcelas. Este algoritmo proporciona una potente herramienta que puede ser utilizada sobre un conjunto muy grande de parcelas y con información asociada a su entorno, ya que está pensado para ser aplicado a nivel municipal. Como el algoritmo desarrollado tiene como objetivo ser una herramienta útil para los expertos en planificación, esta tiene que ser capaz de proporcionar soluciones válidas en tiempos aceptables. Es por esto por lo que muchos de los esfuerzos de diseño e implementación fueron dedicados a su paralelización. El algoritmo puede trabajar sobre cuatro entornos diferentes: entornos monoprocesador (ejecu-

ción secuencial); entornos multi-núcleo (ejecución paralela en memoria compartida); entornos clúster (ejecución paralela en memoria distribuida bajo el paradigma de paso de mensajes); y entornos híbridos (sistemas clúster con nodos multi-núcleo). Aparte del beneficio en cuanto al rendimiento, la paralelización en entornos clúster también ofrece un beneficio en la calidad de las soluciones, ya que se puede jugar con parámetros como el tiempo de sincronización o hacer que la ejecución de cada nodo trabaje con configuraciones distintas, lo que aporta más heterogeneidad a la población.

El algoritmo es altamente configurable por el usuario y esta característica puede ser un arma de doble filo. Por un lado permite al usuario experto controlar multitud de variables y conseguir así mejores soluciones, pero por otro lado un usuario no tan experto puede perderse entre tanta combinación de parámetros. Es por esto por lo que también se aportan unas tablas en las que se muestran resultados obtenidos al ejecutar el algoritmo sobre unos parcelarios ficticios pero representativos, y que son de ayuda a la hora de configurar el algoritmo en base al tipo de parcelario sobre el que se va a aplicar y a los objetivos que se quieren conseguir.

Uno de los retos pendientes de este algoritmo tiene que ver con la detección y posterior unión de manchas con la misma categoría pero que están separadas por carreteras o ríos, lo que afecta directamente a la compacidad de las manchas. Otras cuestiones interesantes a tener en cuenta serían: favorecer la asignación de usos agrícolas en las parcelas cercanas a explotaciones agrícolas ya existentes para favorecer su expansión; y la creación de corredores que conectasen áreas naturales protegidas.

Cabe destacar también que estos dos algoritmos propuestos pueden ser ejecutados de dos maneras diferentes: por línea de comandos, útil en entornos sin interfaz gráfica como clústers; y embebidos en software SIG genérico, como puede ser gvSIG, lo que facilita tanto su uso como su distribución a los usuarios.

8.3. Epílogo

En tiempos de crisis económica como en los que nos encontramos, todas las actividades que potencialmente puedan suponer una reactivación de la economía deben

ser estudiadas y aprovechadas. El campo siempre ha sido un recurso de mucha riqueza hasta que la industrialización y la migración hacia la ciudad redujo su impacto económico. Volver a prestarle atención al medio rural es una buena idea que puede ayudar a paliar los efectos de la crisis. La modernización de los procesos, la mejora de los procedimientos, la utilización de nuevas tecnologías, el aprovechamiento de la información y servicios públicos, la promoción del e-Government, el involucramiento de organismos públicos y privados, y otros muchos aspectos más, deben ser aplicados al y adoptados por el medio rural para conseguir una gestión eficiente de la tierra, con la consecuente obtención de riqueza asociada.

En esta tesis se han expuesto algunas herramientas que apuntan en esta dirección. Todas son herramientas basadas en software libre y cuya aplicación supondrá una mejora en aquellos procesos para los que han sido diseñadas. Acercar los avances tecnológicos al medio rural para disminuir la brecha digital que existe entre campo y ciudad ayudará a que el primero recupere la importancia (en todos los aspectos) que en las últimas décadas ha perdido respecto al segundo.

Apéndice A

Publicaciones de la tesis

A.1. Artículos en revistas

- J. Porta, J. Parapar, P. García, G. Fernández, J. Touriño, R. Doallo, F. Onega, I. Santé, P. Díaz, D. Miranda, R. Crecente. **Web-GIS tool for the management of rural land markets. Application to the Land Bank of Galicia (NW Spain)**. Earth Science Informatics. 2013. (En segunda revisión, publicación en JCR 2011)
- V. Barbosa, I. Santé, R. Crecente, J. Porta, J. Parapar, R. Doallo. **Validação de um algoritmo para a delimitação dos assentamentos rurais na Galiza, NO, Espanha**. Ciência Rural, 2013. (En revisión, publicación en JCR 2011).
- J. Porta, J. Parapar, R. Doallo, V. Barbosa, I. Santé, R. Crecente, C. Díaz. **A population-based iterated greedy algorithm for the delimitation and zoning of rural settlements**. Computers, Environment and Urban Systems, 39(0):12-26, 2013. Publicación en JCR 2011, factor de impacto 1,795.
- J. Porta, J. Parapar, R. Doallo, , F. F. Rivera, I. Santé, R. Crecente. **High performance genetic algorithm for land use planning**. Computers, Environment and Urban Systems, 37(0):45-58, 2013. Publicación en JCR 2011, factor de impacto 1,795.

- J. Porta, J. Parapar, P. García, G. Fernández, J. Touriño, F. Ónega, P. Díaz, D. Miranda, R. Crecente. **Sistema de Información del Banco de Tierras de Galicia**. RISTI: Revista Ibérica de Sistemas y Tecnologías de Información, (9):24-41, 2012.

A.2. Congresos nacionales e internacionales

- J. Porta, J. Parapar, R. Doallo, V. Barbosa, I. Santé, R. Crecente, C. Díaz. **Evolutionary algorithm for the delimitation of rural settlements**. In 9th World Congress of Regional Science Association International, Timisoara, Romania, 2012.
- J. Porta, J. Parapar, G. L. Taboada, R. Doallo, F. F. Rivera, I. Santé, M. R. Suárez, M. Boullón, R. Crecente. **A Java-based Parallel Genetic Algorithm for the Land Use Planning Problem**. In The 2011 Genetic and Evolutionary Computation Conference (GECCO), pages 213-214, Dublin, Ireland, 2011.
- V. Barbosa, I. Santé, R. Crecente, C. Redondo, J. Porta, J. Parapar. **Criterios de identificación y delimitación de núcleos rurales. Aplicación a los núcleos de Guitiriz**. In Congreso Internacional Las Reservas de la Biosfera como Estrategia Territorial de Sostenibilidad, 2011.
- M. Suárez, I. Santé, F. F. Rivera, R. Crecente, M. Boullón, J. Porta, J. Parapar, R. Doallo. **A Parallel Algorithm Based On Simulated Annealing For Land Use Zoning Plans**. In The 2011 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA), pages 360-366, Las Vegas, Nevada, EEUU, 2011.
- F. Ónega, D. Sarmiento, D. Miranda, R. Crecente, J. Porta, J. C. Toucedo, J. Parapar, J. Touriño, R. Doallo. **An open source-based Web-GIS system for the creation of forest management units**. In ForestSat2010: Operational tools in forestry using remote sensing techniques, pages 259-262, 2010.
- R. Doallo, J. Touriño, J. Parapar, R. Crecente, F. Ónega, D. Miranda, J. Porta, J. C. Toucedo, M. Coladas. **Information systems for the Land Bank of**

Galicia (Bantegal) and the Forest Management Units (UXFOR). In Congreso Workshop On Land Tenure and Land Consolidation (Seminario de tenencia de la tierra y concentración parcelario). FAO (Organización de las Naciones Unidas para la Agricultura y la Alimentación), 2009.

- R. Doallo, J. Touriño, J. Parapar, R. Crecente, F. Ónega, D. Miranda, J. Porta, M. Coladas. **SITEGAL, un sistema de información para la gestión de tierras en Galicia.** In Jornadas de Sistemas de Información Geográfica, GISDay. LaboraTe (Laboratorio do Territorio) de la Universidad de Santiago de Compostela y REDIX (Rede Galega de Investigación en Información Xeográfica), 2007.

Apéndice B

Encuesta a técnicos de los posibles impedimentos para la movilidad de tierras

Restricciones	Tipo	Puntuación
Las fincas infrautilizadas no son aptas para el cultivo por cuestiones de suelo, clima o pendiente.	RF	2.12
Las fincas infrautilizadas son pequeñas y/o irregulares.	RF	4.19
Las fincas infrautilizadas tienen un acceso inadecuado.	RF	3.19
Las fincas infrautilizadas están lejos de los lugares donde hay demanda.	RF	2.69
No se conoce a los propietarios/as de las fincas.	CT	2.19
Los/as propietarios/as no residen en la zona y es difícil contactar con ellos.	CT	2.56
Existen problemas legales con las fincas: herencias por resolver, problemas con los límites, ...	CT	2.53
Los/as propietarios/as no quieren arrendar por desconfianza de perder derechos sobre la parcela.	CT	4.06
Los/as propietarios/as no quieren arrendar por miedo a tener algún otro problema durante la vigencia del contrato (cobro, daños, etc.).	CT	3.12

Restricciones	Tipo	Puntuación
Los/as propietarios/as no quieren arrendar de manera formal (con <i>papeles</i> y esto es un límite para los agricultores).	CT	3.94
Los/as propietarios/as no quieren vender porque esperan utilizar en el futuro.	RPO	3.60
Los/as propietarios/as no quieren vender porque esperan futuras revalorizaciones.	RPO	3.19
Los/as propietarios/as no quieren vender porque consideran que la tierra es un patrimonio seguro.	RPO	3.62
Los/as propietarios/as no quieren vender por motivos de apego emocional o tradición de mantener un patrimonio familiar.	RPO	3.94
Los trámites para la venta y sus costes de notario, registro, etc., son un obstáculo.	RPO	2.13
Los agricultores, aún queriendo tierras, no hacen una búsqueda activa.	CT	3.06
Los agricultores, aún buscando tierras, no tienen fuentes de información adecuadas.	CT	2.56

Tabla B.1: Puntuaciones (1-5) de las restricciones que dificultan la movilidad geográficas (RF: Restricciones Físicas, CT: Costes de Transacción, RPO: Reparos de la Parte Ofertante).

Bibliografía

- [1] AERTS, J. C. J. H., EISINGER, E., HEUVELINK, G. B. M., AND STEWART, T. Using linear integer programming for multi-site land-use allocation. *Geographical Analysis* 35, 2 (2003), 148–169. Citado en pp: 133
- [2] AERTS, J. C. J. H., VAN HERWIJNEN, M., JANSSEN, R., AND STEWART, T. J. Evaluating spatial design techniques for solving land-use allocation problems. *Journal of Environmental Planning and Management* 48, 1 (2005), 121–142. Citado en pp: 76, 133
- [3] ALCANTARA, C., KUEMMERLE, T., PRISHCHEPOV, A. V., AND RADELOFF, V. C. Mapping Abandoned Agriculture with Multi-Temporal MODIS Satellite Data. *Remote Sensing of Environment* 124 (2012), 334–347. Citado en pp: 20
- [4] ARENTZE, T. A., BORGER, A. W. J., MA, L., AND TIMMERMAN, H. J. P. An agent-based heuristic method for generating land-use plans in urban planning. *Environment and Planning B: Planning and Design* 37 (2010), 463–482. Citado en pp: 133
- [5] ARGE, L., DE BERG, M., HAVERKORT, H. J., AND YI, K. The Priority R-tree: a practically efficient and worst-case optimal R-tree. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data* (New York, NY, USA, 2004), SIGMOD '04, ACM, pp. 347–358. Citado en pp: 82
- [6] AROZARENA, A., GARCÍA, L., AND VILLA, G. The Digital Imaging Cartography (DIC) as Basic Cartographic Information System. The Spanish National Plan for Land Observation (PNOT): Collaborative and Decentralized

- Model of Production. In *ICC 2009. 24th International Cartographic Conference*. (Santiago, Chile, 2009). URL: http://icaci.org/documents/ICC_proceedings/ICC2009/html/refer/2_1.pdf. Citado en pp: 64
- [7] BÄCK, T. *Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms*. Oxford University Press, Oxford, UK, 1996. Citado en pp: 76
- [8] BAKOS, J. Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science* 43, 12 (1997), 1676–1692. Citado en pp: 25
- [9] BALLESTÍN, F., SCHWINDT, C., AND ZIMMERMANN, J. Resource leveling in make-to-order production: modeling and heuristic solution method. *International Journal of Operations* 4, 1 (2007), 50–62. Citado en pp: 79
- [10] BALLING, R. J., TABER, J. T., BROWN, M. R., AND DAY, K. Multiobjective urban planning using genetic algorithm. *Journal of Urban Planning and Development* 125, 2 (1999), 16–99. Citado en pp: 76, 77
- [11] BANDEIRA, P., AND SUMPSI, J. Las políticas de tierra en los países en desarrollo. In *V Congreso de Economía Agraria. Agricultura, alimentación y espacio rural en transición*, AEEA (2004). Citado en pp: 16, 21
- [12] BANSKI, J., AND WESOŁOWSKA, M. Transformations in Housing Construction in Rural Areas of Poland's Lublin Region - Influence on the Spatial Settlement Structure and Landscape Aesthetics. *Landscape and Urban Planning* 94, 2 (2010), 116–126. Citado en pp: 86
- [13] BARBOSA, V. Criterios de delimitación y planeamiento aplicados a Núcleos Rurales. Master's thesis, Departamento de Ingeniería Agroforestal. Grupo de investigación GI-1934 TB (Territorio-Biodiversidade/LaboraTe). Universidad de Santiago de Compostela. Escuela Politécnica Superior de Lugo. España, 2010. Citado en pp: 90
- [14] BARBOSA, V., CRECENTE, R., AND SANTÉ, I. The Rural Connection As a Re-construction of a Territorial Structure: Sustainable Spatial Criteria to Delimit Rural Settlements Following an Algorithm. *The International Journal of the Constructed Environment* 2, 2 (2011), 111–129. Citado en pp: 110, 119

- [15] BARBOSA, V., CRECENTE, R., AND SANTÉ, I. The rural connection as a re-construction of a territorial structure: sustainable spatial criteria to delimit rural settlements following an algorithm. *The International Journal of the Constructed Environment* 2, 2 (2012), 111–129. Citado en pp: 131
- [16] BARBOSA, V., SANTÉ, I., AND CRECENTE, R. A evolução da lei do solo nos núcleos rurais da Galiza, Espanha. In *7th International Conference on Virtual Cities and Territories* (2011). Citado en pp: 131
- [17] BARBOSA, V., SANTÉ, I., CRECENTE, R., PORTA, J., PARAPAR, J., AND DOALLO, R. Validação de um algoritmo para a delimitação dos assentamentos rurais na Galiza, NO, Espanha. *Ciência Rural* (2013). En revisión. Citado en pp: 131
- [18] BARBOSA, V., SANTÉ, I., CRECENTE, R., REDONDO, C., PORTA, J., AND PARAPAR, J. Criterios de identificación y delimitación de núcleos rurales. Aplicación a los núcleos de Guitiriz. In *Congreso Internacional Las Reservas de la Biosfera como Estrategia Territorial de Sostenibilidad* (2011). Citado en pp: 131
- [19] BARBOSA, V., SANTÉ, I., FERREIRA, J. A., AND CRECENTE, R. Análisis de los factores que influyen en la capacidad de los núcleos rurales para la fijación de población en el medio rural. In *Congreso Internacional Las Reservas de la Biosfera como Estrategia Territorial de Sostenibilidad* (2011). Citado en pp: 131
- [20] BARBOSA BRANDAO, V., CORBELLE RICO, E., AND FERREIRA NETO, J. A. El entorno inmediato del núcleo en el planeamiento rural. Delimitación como área de transición para aplicar la concentración parcelaria. *Territorios a Exame: Traballos de Ordenación Territorial* (2012). Citado en pp: 125
- [21] BAYFIELD, N. G., CONROY, J., BIRNIE, R. V., GEDDES, A., MIDGLEY, J. L., SHUCKSMITH, M. D., AND ELSTON, D. Current Awareness, Use and Perceived Priorities for Rural Databases in Scotland. *Land Use Policy* 22, 2 (2005), 153–162. Citado en pp: 65
- [22] BENEDETTINI, S., BLUM, C., AND ROLI, A. A Randomized Iterated Greedy Algorithm for the Founder Sequence Reconstruction Problem. In *Learning and*

- Intelligent Optimization*, C. Blum and R. Battiti, Eds., vol. 6073 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2010, pp. 37–51. Citado en pp: 78
- [23] BENNETT, R., A., R., WILLIAMSON, I. P., AND WALLACE, J. On the Need for National Land Administration Infrastructures. *Land Use Policy* 29, 1 (2012), 208–219. Citado en pp: 23
- [24] BHATTA, G. D., AND DOPPLER, W. Farming Differentiation in the Rural-urban Interface of the Middle Mountains, Nepal: Application of Analytic Hierarchy Process (AHP) Modeling. *Journal of Agricultural Science* 2 (2010), 37–51. Citado en pp: 87
- [25] BOUAMAMA, S., BLUM, C., AND BOUKERRAM, A. A population-based iterated greedy algorithm for the minimum weight vertex cover problem. *Applied Soft Computing* 12, 6 (2012), 1632–1639. Citado en pp: 79, 113
- [26] BROOKES, C. J. A genetic algorithm for designing optimal patch configurations in GIS. *International Journal of Geographical Information Science* 15, 6 (2001), 539–559. Citado en pp: 133
- [27] BUSCAFINCAS. <http://buscafincas.com>. Visitado en mayo de 2013. Citado en pp: 32
- [28] BÄCK, T. *Evolutionary Algorithms in Theory and Practice*. Oxford Univ. Press, 1996. Citado en pp: 141
- [29] C. BARCIELA, M.A. LÓPEZ, J. M. La intervención del Estado en la agricultura durante el siglo XX. *Revista AYER. Asociación de Historia Contemporánea-Marcial Pons, Ediciones de Historia* 0, 6 (1996), 51–96. Citado en pp: 10
- [30] CLEMATIS, A., MINETER, M., AND MARCIANO, R. High performance computing with geographical data. *Parallel Computing* 29 (2003), 1275–1279. Citado en pp: 134
- [31] COASE, R. H. The nature of the firm. *Economica* 4 (1937), 386–405. Citado en pp: 16, 21

- [32] CORBELLE, E., AND CRECENTE, R. Urbanization, Afforestation, Abandonment. Recent Changes in Galician Landscape. *Economic Galician Journal (in Galician: Revista Galega de Economía)* (2012). In press. Citado en pp: 21
- [33] CORMEN, T., LEISERSON, C., RIVEST, R., AND STEIN, C. *Introduction to Algorithms*. MIT Press, 2001. Citado en pp: 78, 82
- [34] CRECENTE, R. *Land Consolidation in Galicia: Validation as an Instrument for Rural Planning*. PhD thesis, Department of Agricultural and Forestry Engineering, University of Santiago de Compostela, Spain, 1998. Citado en pp: 11
- [35] CROMLEY, R. G., AND HANINK, D. M. Coupling land use allocation models with raster GIS. *Journal of Geographical Systems* 1 (1999), 137–153. Citado en pp: 133
- [36] DAHLMAN, C. The Problem of Externality. *Journal of Law and Economics* 22, 1 (1979), 141–162. Citado en pp: 21
- [37] DE JONG, K. *An analysis of the behavior of a class of genetic adaptive systems*. PhD thesis, University of Michigan, Ann Arbor, MI, USA, 1975. Citado en pp: 142
- [38] DEININGER, K., SAVASTANO, S., AND CARLETTO, C. Land Fragmentation, Cropland Abandonment, and Land Market Operation in Albania. *World Development* 40, 10 (2012), 2108–2122. Citado en pp: 21
- [39] DICKINSON, R. E. Rural Settlements in the German Lands. *Annals of the Association of American Geographers* 39, 4 (1949), 239–263. Citado en pp: 85
- [40] DIJK, T. *Dealing with Central European Land Fragmentation. A Critical Assessment on the Use of Western European Instruments*. Eburon, Delft (The Netherlands), 2003. Citado en pp: 57
- [41] DIJK, T. v., AND KOPEVA, D. Land Banking and Central Europe: Future Relevance, Current Initiatives, Western European Past Experience. *Land Use Policy* 23, 3 (2006), 286–301. Citado en pp: 23, 24

- [42] DOALLO, R., TOURIÑO, J., PARAPAR, J., CRECENTE, R., ÓNEGA, F., MIRANDA, D., PORTA, J., AND COLADAS, M. SITEGAL, un sistema de información para la gestión de tierras en Galicia. In *Jornadas de Sistemas de Información Geográfica, GISDay* (2007), LaboraTe (Laboratorio do Territorio) de la Universidad de Santiago de Compostela y REDIX (Rede Galega de Investigación en Información Xeográfica). Citado en pp: 70
- [43] DOALLO, R., TOURIÑO, J., PARAPAR, J., CRECENTE, R., ÓNEGA, F., MIRANDA, D., PORTA, J., TOUCEDO, J. C., AND COLADAS, M. Information systems for the Land Bank of Galicia (Bantegal) and the Forest Management Units (UXFOR). In *Congreso Workshop On Land Tenure and Land Consolidation (Seminario de tenencia de la tierra y concentración parcelario)* (2009), FAO (Organización de las Naciones Unidas para la Agricultura y la Alimentación). Citado en pp: 70
- [44] DOUGLAS D. NEBERT, TECHNICAL WORKING GROUP CHAIR, G., Ed. *The SDI Cookbook*. Developing Spatial Data Infrastructures, 2004. URL: <http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf>. Citado en pp: 27
- [45] DUH, J. D., AND BROWN, D. G. Knowledge-informed Pareto simulated annealing for multi-objective spatial allocation. *Computers, Environment and Urban Systems* 31 (2007), 235–281. Citado en pp: 76, 133
- [46] EASTMAN, J., JIN, W., KYEM, P. A. K., AND TOLEDANO, J. Raster procedures for multi-criteria/multi-objective decisions. *Photogrammetric Engineering & Remote Sensing* 61, 5 (1995), 539–547. Citado en pp: 133
- [47] EBERLIN, R., VAN HOLST, F., AND ÓNEGA LÓPEZ, F. Land banking in Europe: Instruments for Rural Development. In *3rd International LANDNET Workshop on Land Market Development and Land Consolidation* (Budapest, Hungary, 2012). Citado en pp: 57
- [48] ELDRANDALY, K. A GEP-based spatial decision support system for multisite land use allocation. *Applied Soft Computing* 10 (2010), 694–702. Citado en pp: 76

-
- [49] ENEMARK, S. Underpinning Sustainable Land Administration Systems. In *Sixteenth United Nations Regional Cartographic Conference for Asia and the Pacific* (2003). Citado en pp: 25
- [50] ENEMARK, S., WILLIAMSON, I., AND WALLACE, J. Building Modern Land Administration Systems in Developed Economies. *Journal of Spatial Science* 50, 2 (2005). Citado en pp: 25
- [51] ENRÍQUEZ, R. O., AND RODRÍGUEZ, S. G. Urban Deconcentration in the Territorial System of Galicia: A Conceptual and Empirical Approach. *Boletín de la Asociación de Geógrafos Españoles* 44 (2007), 383–386. Citado en pp: 118
- [52] EUROPEAN UNION. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). *Summaries of EU legislation. Environment. General provisions* (2007). URL: http://europa.eu/legislation_summaries/environment/general_provisions/128195_en.htm. Citado en pp: 27
- [53] FAO ET AL. *Guidelines for Land-Use Planning*. Development Documents Series. Food and Agriculture Organization of the United Nations, 1993. Citado en pp: 74
- [54] FARM SELLER. <http://www.farmseller.com>. Visitado en mayo de 2013. Citado en pp: 32
- [55] FAZ ONLINE. <http://fazonline.com.br>. Visitado en mayo de 2013. Citado en pp: 32
- [56] FEDER, H., AND FEENY, D. Land Tenure and Property Rights: Theory and Implications for Development Policy. *The World Bank Economic Review* 5 (1991), 135–154. Citado en pp: 21
- [57] FENG, C. M., AND LIN, J. J. Using a genetic algorithm to generate alternative sketch maps for urban planning. *Computers, Environment and Urban Systems* 23 (1999), 91–108. Citado en pp: 76

- [58] FENG, W., WANG, N., WANG, C., LI, G., AND ZHANG, C. Study on Characteristics of Rural Settlements in the Northeast Loess Plateau of China. In *IEEE International Geoscience and Remote Sensing Symposium IGARSS* (Barcelona, Spain, 2007), IEEE, Ed. Citado en pp: 85, 86
- [59] FERNÁNDEZ AGUERRI, M. Retos y problemática dela mujer en el ámbito rural. El papel de las agrupaciones de este sector de población en el desarrollo rural. *Jornada Temática sobre Políticas de relevo generacional e incorporación de la mujer al mundo rural* (2002). Madrid. Citado en pp: 1
- [60] FERREIRA, J. A., CONDESSA, B., CASTRO E ALMEIDA, J., AND PINTO, P. Urban Settlements Delimitation in Low-Density Areas - An Application to the Municipality of Tomar (Portugal). *Landscape and Urban Planning* 97 (2010), 156–157. Citado en pp: 85, 86
- [61] FERREIRA-NETO, J. A., CARNEIRO DOS SANTOS-JUNIOR, E., FRA-PALEO, U., MIRANDA-BARRÓS, D., AND CÉSAR DE OLIVEIRA-MOREIRA, M. Optimal Subdivision of Land in Agrarian Reform Projects: An Analysis Using Genetic Algorithms. *Ciencia e Investigación Agraria* 38, 2 (2011), 169–178. Citado en pp: 77
- [62] FINCAS TURÓ. <http://www.turofincas.com>. Visitado en mayo de 2013. Citado en pp: 32
- [63] FUENTES, J. Methodological Bases for Documenting and Reusing Vernacular Farm Architecture. *Journal of Cultural Heritage* 11, 2 (2010), 119–129. Citado en pp: 96
- [64] GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989. Citado en pp: 76, 77
- [65] GOLDBERG, D. E., AND DEB, K. A comparative analysis of selection schemes used in genetic algorithms. In *Foundations of Genetic Algorithms* (1991), Morgan Kaufmann, pp. 69–93. Citado en pp: 141
- [66] GONÇALVES, L., CORDERO, M., CRECENTE, R., AND MIRANDA, D. GIS Proprietary Versus Free: a Comparative Study Using Geomedia Professional

- and Grass. *International Journal of Geoinformatics* 2, 2 (2006), 48–57. Citado en pp: 39
- [67] GOVERNMENT SERVICE FOR LAND AND WATER MANAGEMENT OF THE NETHERLANDS. Land abandonment, biodiversity and the cap. In *Land Abandonment and Biodiversity in Relation to the 1st and 2nd Pillars of the EU's Common Agricultural Policy; Outcome of an International Seminar in Sigulda* (2004). Citado en pp: 20
- [68] GREFFENSTETTE, J. J. Optimization of Control Parameters for Genetic Algorithms. *Systems, Man and Cybernetics, IEEE Transactions on* 16, 1 (1986), 122–128. Citado en pp: 142
- [69] GROSSMAN, D., AND KATZ, Y. Rural Settlement Patterns in Eretz-Israel. *Geografiska Annaler. Series B, Human Geography* 74, 1 (1992), 57–73. Citado en pp: 85
- [70] HERR, R. *España contemporánea*. Biblioteca Clásica. Marcial Pons Historia, 2004. Citado en pp: 9
- [71] HOLZKAMPER, A., AND SEPPELT, R. A generic tool for optimizing land-use patterns and landscape structures. *Environmental Modelling & Software* 22 (2007), 1801–1804. Citado en pp: 76
- [72] HOUSING MAPS. <http://www.housingmaps.com>. Visitado en mayo de 2013. Citado en pp: 32
- [73] I SUBIRANA, J. C. Infraestructura de Datos Espaciales (IDE). Definición y Desarrollo Actual en España. *Scripta Nova. Revista Electrónica de Geografía y Ciencias Sociales* (2004). Citado en pp: 27
- [74] JANSSEN, R., VAN HERWIJNEN, M., STEWART, T. J., AND AERTS, J. C. J. H. Multiobjective decision support for land-use planning. *Environment and Planning B: Planning and Design* 35 (2008), 740–756. Citado en pp: 76, 133
- [75] JAVA.COM. URL: <http://download.oracle.com/javase/6/docs/api/java/util/concurrent/package-summary.html>, 2011. Visitado en mayo de 2013. Citado en pp: 145

- [76] JOHNSON, M., SINGH, I., AND STEARNS, B. *Designing Enterprise Applications with the J2EE Platform, 2nd Edition*. Prentice Hall PTR, 2002. URL: http://java.sun.com/blueprints/guidelines/designing_enterprise_applications_2e/. Citado en pp: 42
- [77] KAI, C., BO, H., QING, Z., AND SHENGXIAO, W. Land use allocation optimization towards sustainable development based on genetic algorithm. In *Geoinformatics, 2009 17th International Conference on* (2009), pp. 1–5. Citado en pp: 76, 80, 133
- [78] KAUFMANN, J. E STEUDLER, D. Catastro 2014. Visión para un sistema catastral futuro. In *FIG* (1998). Citado en pp: 19
- [79] KOCH, R. *The 80/20 Principle: The Secret to Achieving More with Less*. Crown Business, 1999. Citado en pp: 1
- [80] LAND AND RANCH SALES. <http://www.landandranchsales.com>. Visitado en mayo de 2013. Citado en pp: 32
- [81] LERISE, F. Centralised Spatial Planning Practice and Land Development Realities in Rural Tanzania. *Habitat International* 24, 2 (2000), 185–200. Citado en pp: 85, 86
- [82] LEUTENEGGER, S., LOPEZ, M., AND EDGINGTON, J. STR: a simple and efficient algorithm for R-tree packing. In *Data Engineering, 1997. Proceedings. 13th International Conference on* (1997), pp. 497–506. Citado en pp: 81
- [83] LIORIT, D. SAFER’s Role in Guiding the Rural Land Market: Instruments, Organisation and Experiences. In *3rd International LANDNET Workshop on Land Market Development and Land Consolidation* (Budapest, Hungary, 2012). Citado en pp: 57
- [84] LOZANO, M., MOLINA, D., AND GARCÍA-MARTÍNEZ, C. Iterated greedy for the maximum diversity problem. *European Journal of Operational Research* 214, 1 (2011), 31–38. Citado en pp: 79
- [85] LÓPEZ-IGLESIAS, E. Galician Agricultural Sector on the Threshold of XXI Century. *Economic Galician Journal (in Galician: Revista Galega de Economía)* 9, 1 (2000), 167–196. Citado en pp: 21

- [86] MACDONALD, D., CRABTREE, J., WIESINGER, G., DAX, T., STAMOU, N., FLEURY, P., LAZPITA, J. G., AND GIBON, A. Agricultural Abandonment in Mountain Areas of Europe: Environmental Consequences and Policy Response. *Journal of Environmental Management* 59, 1 (2000), 47–69. Citado en pp: 20
- [87] MASSER, I., RAJABIFARD, A., AND WILLIAMSON, I. Spatially enabling governments through SDI implementation. *International Journal of Geographical Information Science* 0, 1 (2007), 1–16. Citado en pp: 27
- [88] MATTHEWS, K. B., SIBBALD, A. R., AND CRAW, S. Implementation of a spatial decision support system for rural land use planning: integrating geographic information system and environmental models with search and optimisation algorithms. *Computers and Electronics in Agriculture* 23, 1 (1999), 9–26. Citado en pp: 76
- [89] MATTHEWS, K. B., SIBBALD, A. R., AND CRAW, S. Implementation of a spatial decision support system for rural land use planning: integrating geographic information system and environmental models with search and optimisation algorithms. *Computers and Electronics in Agriculture* 23, 1 (1999), 9–26. Citado en pp: 80
- [90] MINETER, M. J., AND DOWERS, S. Parallel processing for geographical applications: a layered approach. *Journal of Geographical Systems* 1 (1999), 61–74. Citado en pp: 134
- [91] MOLEN, P. v. D. Some Macroeconomic Aspects of Land Ownership. *2nd FIG Regional Conference, FIG* (2003). Citado en pp: 23
- [92] MOLEN, P. v. D. Some Microeconomic Aspects of Land Ownership. *3rd FIG Regional Conference for Asia and the Pacific: Surveying the Future - Contributions to Economic, Environmental and Social Development, FIG* (2004). Citado en pp: 23
- [93] MOLEN, P. v. D. E ÖSTERBERG, T. Land tenure and land administration for social and economic development in (Western) Europe. In *UN-FIG Conference on Land Tenure and Cadastral Infrastructure for Sustainable Development, FIG* (1999). Citado en pp: 26

-
- [94] MONTERO, R. S., AND BRIBIESCA, E. State of the art of compactness and circularity measures. *International Mathematical Forum* 4, 25-28 (2009), 1305–1335. Citado en pp: 111, 144
- [95] MUILU, T., AND RUSANEN, J. Rural Definitions and Short-Term Dynamics in Rural Areas of Finland in 1989-97. *Environment and Planning A* 36, 8 (2004), 1499–1516. Citado en pp: 85
- [96] MUKERJI, A. Rural Settlements of the Chandigarh Siwalik Hills (India): A Morphogenetic Analysis. *Geografiska Annaler. Series B, Human Geography* 58, 2 (1976), 95–115. Citado en pp: 85
- [97] NEAPOLITAN, R., AND NAIMIPOUR, K. *Foundations of Algorithms*. Jones & Bartlett Learning, 2010. Citado en pp: 78
- [98] OLAYA, V. <http://www.sextantegis.com>, 2011. Visitado en mayo de 2013. Citado en pp: 170
- [99] ÓNEGA, F., SARMIENTO, D., MIRANDA, D., CRECENTE, R., PORTA, J., TOUCEDO, J. C., PARAPAR, J., TOURIÑO, J., AND DOALLO, R. An open source-based Web-GIS system for the creation of forest management units. In *ForestSat2010: Operational tools in forestry using remote sensing techniques* (2010), pp. 259–262. Citado en pp: 70
- [100] ÓNEGA LÓPEZ, F., PUPPIM DE OLIVEIRA, J., AND CRECENTE-MASEDA, R. Planning Innovations in Land Management and Governance in Fragmented Rural Areas: Two Examples from Galicia (Spain). *European Planning Studies* 18, 5 (2010), 755–773. Citado en pp: 21
- [101] PAN, Q.-K., WANG, L., AND ZHAO, B.-H. An improved iterated greedy algorithm for the no-wait flow shop scheduling problem with makespan criterion. *The International Journal of Advanced Manufacturing Technology* 38 (2008), 778–786. Citado en pp: 78
- [102] PORTA, J. Aplicación para la gestión de un banco de tierras basada en herramientas web y SIG libres. Master's thesis, Departamento de Electrónica y Sistemas. Universidad de A Coruña, 2007. Citado en pp: 22, 28

- [103] PORTA, J., PARAPAR, J., DOALLO, R., BARBOSA, V., SANTÉ, I., CRECENTE, R., AND DÍAZ, C. Evolutionary Algorithm for the Delimitation of Rural Settlements. In *9th World Congress of Regional Science Association International* (Timisoara, Romania, 2012). Citado en pp: 131
- [104] PORTA, J., PARAPAR, J., DOALLO, R., BARBOSA, V., SANTÉ, I., CRECENTE, R., AND DÍAZ, C. A Population-Based Iterated Greedy Algorithm for the Delimitation and Zoning of Rural Settlements. *Computers, Environment and Urban Systems* 39, 0 (2013), 12 – 26. Citado en pp: 131
- [105] PORTA, J., PARAPAR, J., DOALLO, R., RIVERA, F. F., SANTÉ, I., AND CRECENTE, R. High performance genetic algorithm for land use planning. *Computers, Environment and Urban Systems* 37, 0 (2013), 45–58. Citado en pp: 172, 176
- [106] PORTA, J., PARAPAR, J., GARCÍA, P., FERNÁNDEZ, G., TOURIÑO, J., DOALLO, R., ÓNEGA, F., SANTÉ, I., DÍAZ, P., MIRANDA, D., AND CRECENTE, R. Web-GIS tool for managing rural land markets: application to the Land Bank of Galicia (NW Spain). *Transactions in GIS* (2013). En revisión. Citado en pp: 70
- [107] PORTA, J., PARAPAR, J., GARCÍA, P., FERNÁNDEZ, G., TOURIÑO, J., ÓNEGA, F., DÍAZ, P., MIRANDA, D., AND CRECENTE, R. Sistema de Información del Banco de Tierras de Galicia. *RISTI: Revista Ibérica de Sistemas y Tecnologías de Información* 0, 9 (2012), 24–41. Citado en pp: 70
- [108] PORTA, J., PARAPAR, J., TABOADA, G. L., DOALLO, R., RIVERA, F. F., SANTÉ, I., R.SUÁREZ, M., BOULLÓN, M., AND CRECENTE, R. A Java-based Parallel Genetic Algorithm for the Land Use Planning Problem. In *The 2011 Genetic and Evolutionary Computation Conference (GECCO)* (Dublin, Ireland, 2011), pp. 213–214. Citado en pp: 172
- [109] REGNAULD, N., AND REVELL, P. Automatic Amalgamation of Buildings for Producing Ordnance Survey® 1: 50 000 Scale Maps. *Cartographic Journal* 44, 3 (2007), 239–250. Citado en pp: 110
- [110] RENWICK, A., JANSSON, T., VERBURG, P. H., REVOREDO-GIHA, C., BRITZ, W., GOCHT, A., AND MCCracken, D. Policy Reform and Agricul-

- tural Land Abandonment in the EU. *Land Use Policy* 30, 1 (2013), 446–457. Citado en pp: 20
- [111] RIBAS, I., COMPANYS, R., AND TORT-MARTORELL, X. An iterated greedy algorithm for the flowshop scheduling problem with blocking. *Omega* 39, 3 (2011), 293–301. Citado en pp: 78
- [112] RODRIGUEZ, F., BLUM, C., LOZANO, M., AND GARCÍA-MARTÍNEZ, C. Iterated Greedy Algorithms for the Maximal Covering Location Problem. In *Evolutionary Computation in Combinatorial Optimization*, J.-K. Hao and M. Middendorf, Eds., vol. 7245 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2012, pp. 172–181. Citado en pp: 79
- [113] RODRÍGUEZ, F. P., AND ALBORECA, A. R. MPC 2.0©, Software for the Application of AHP Method in Multiple-Criteria Decision-Making. *Recursos Rurais* 7 (2011), 27–33. Citado en pp: 87
- [114] RUIZ, R., AND STÜTZLE, T. A simple and effective iterated greedy algorithm for the permutation flowshop scheduling problem. *European Journal of Operational Research* 177, 3 (2007), 2033–2049. Citado en pp: 78
- [115] SACO, A. Desarrollo rural y despoblación en Galicia: escenarios y desarrollos de futuro. *Revista de Estudios sobre Despoblación y Desarrollo Rural* 0, 9 (2010), 11–30. Citado en pp: 12
- [116] SHAFI, A. <http://mpj-express.org>, 2013. Visitado en mayo de 2013. Citado en pp: 36, 146
- [117] SMAILES, P. J., AND MOLYNEUX, J. The Evolution of an Australian Rural Settlement Pattern: Southern New England, NSW. *Transactions of the Institute of British Geographers* 36 (1965), 31–54. Citado en pp: 85
- [118] SOLUTIONS, V. URL: <http://www.vividsolutions.com/jts/jtshome.htm>, 2012. Visitado en mayo de 2013. Citado en pp: 81
- [119] SPEARS, W. M., AND DE JONG, K. An analysis of multipoint crossover. In *Proc. of Workshop of the Foundations of Genetic Algorithms. Bloomington, USA.* (1991), pp. 301–315. Citado en pp: 141

- [120] STEWART, T. J., JANSSEN, R., AND HERWIJNEN, M. A genetic algorithm approach to multiobjective land use planning. *Computers & Operations Research* 31 (2004), 2293–2313. Citado en pp: 76, 133
- [121] STOIAN, D., AND HENKEMANS, A. B. Between Extractivism and Peasant Agriculture: Differentiation of Rural Settlements in the Bolivian Amazon. *International Tree Crops Journal* 10, 4 (2000), 299–320. Citado en pp: 85, 86
- [122] STRADER, T., AND SHAW, M. Characteristics of Electronic Markets: Electronic Commerce. *Decision Support Systems* 21, 3 (1997), 185–198. Citado en pp: 25
- [123] SUÁREZ, L. *Historia general de España y América: La Segunda República y la Guerra*. Ediciones Rialp, 1986. Citado en pp: 9
- [124] SUÁREZ, M., SANTÉ, I., RIVERA, F. F., CRECENTE, R., BOULLÓN, M., PORTA, J., PARAPAR, J., AND DOALLO, R. A Parallel Algorithm Based On Simulated Annealing For Land Use Zoning Plans. In *The 2011 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA)* (Las Vegas, Nevada, EEUU, 2011), pp. 360–366. Citado en pp: 136, 172
- [125] TABOADA, G. L., TOURIÑO, J., AND DOALLO, R. Java for High Performance Computing: Assessment of Current Research and Practice. In *7th International Conference on the Principles and Practice of Programming in Java, PPPJ 2009. ACM International Conference Proceeding Series* (Calgary, Alberta, Canada, 2009), pp. 30–39. Citado en pp: 36, 83
- [126] TABOADA, G. L., TOURIÑO, J., AND DOALLO, R. F-MPJ: scalable Java message-passing communications on parallel systems. *The Journal of Supercomputing* 60, 1 (2012), 117–140. Citado en pp: 36
- [127] TOYAMA, F., SHOJI, K., AND MIYAMICHI, J. An iterated greedy algorithm for the node placement problem in bidirectional manhattan street networks. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation* (New York, NY, USA, 2008), GECCO '08, ACM, pp. 579–584. Citado en pp: 79

- [128] TUFFERY, P., GUYON, F., AND DERREUMAUX, P. Improved greedy algorithm for protein structure reconstruction. *Journal of Computational Chemistry* 26, 5 (2005), 506–513. Citado en pp: 78
- [129] TURNOCK, D. The Planning of Rural Settlement in Romania. *The Geographical Journal* 157, 2 (1991), 251–264. Citado en pp: 86
- [130] UNION NATIONS ECONOMIC COMMISSION FOR EUROPE. Land administration guidelines. Tech. rep., United Nations, New York and Geneve, 1996. Citado en pp: 19
- [131] VALIENTE, F. *El marco político de la desamortización en España*. Ariel quincenal. Ariel, 1971. Citado en pp: 9
- [132] WALLACE, J., AND WILLIAMSON, I. Building Land Markets. *Land Use Policy* 23 (2006), 123–135. Citado en pp: 26
- [133] WALLACE, J., WILLIAMSON, I., RAJABIFARD, A., AND BENNETT, R. Spatial Information Opportunities for Government. *Spatial Sciences Institute Special Edition on Spatial Information and Spatial Data Infrastructures* 51, 1 (2006). Citado en pp: 25, 26
- [134] WALLACE, J., AND WILLIAMSON, I. P. Developing Cadastres to Service Complex Property Markets. *Computers, Environment and Urban Systems* 30, 5 (2006), 614–626. Citado en pp: 23
- [135] WILLIAMSON, I. Land administration best practice providing the infrastructure for land policy implementation. *Land Use Policy* 18 (2001), 297–307. Citado en pp: 26
- [136] XIAO, N., BENNETT, D. A., AND ARMSTRONG, M. P. Interactive evolutionary approaches to multiobjective spatial decision making: A synthetic review. *Computers, Environment and Urban Systems* 31 (2001), 232–252. Citado en pp: 76
- [137] XIBAO, X., JIANMING, Z., AND XIAOJIAN, Z. Integrating GIS, cellular automata and genetic algorithm in urban spatial optimization. A case study of Lanzhou. In *Proc of SPIE* (1995), vol. 6420. 64201U-1 - 64201U-10. Citado en pp: 76, 133

- [138] XIN, H., AND ZHI-XIA, Z. Application of genetic algorithm to spatial distribution in urban planning. In *IEEE International Symposium on Knowledge Acquisition and Modeling Workshop* (2008), pp. 1026–1029. Wuhan, China. Citado en pp: 76, 77
- [139] XUNTA DE GALICIA. Ley 9/2002, del 30 de diciembre, de Ordenación Urbanística y Protección del Medio Rural de Galicia. *Diario Oficial de Galicia* 252 (2002), 18025–18094. Citado en pp: 88, 135
- [140] XUNTA DE GALICIA. Ley 3/2007, de 9 de abril, de prevención y defensa contra los incendios forestales de Galicia. *Boletín Oficial de Estado (BOE)* (2007). URL: http://www.boe.es/diario_boe/txt.php?id=BOE-A-2007-10021. Citado en pp: 68
- [141] XUNTA DE GALICIA. Ley 7/2007, de 21 de mayo, de medidas administrativas y tributarias para la conservación de la superficie agraria útil y del Banco de Tierras de Galicia. *Boletín Oficial de Estado (BOE)* (2007). URL: http://www.boe.es/diario_boe/txt.php?id=BOE-A-2010-10707. Citado en pp: 58
- [142] XUNTA DE GALICIA. Ley 2/2010, del 25 de marzo, de Medidas Urgentes de Modificación de la ley 9/2002 de Ordenación Urbanística y Protección del Medio Rural de Galicia. *Diario Oficial de Galicia* 61 (2010), 4639–4666. Citado en pp: 88, 135
- [143] ZEVENBERGEN, J. Land registration, transaction cost and the land market. *22nd Urban and Regional Data Management Symposium. Seminar : Land Markets and Land Consolidation in Central Europe* (2000). Citado en pp: 16, 21
- [144] ZHANG, H. H., ZENG, Y. N., AND BIAN, L. Simulating multi-objective spatial optimization allocation of land use based on the integration of multi-agent system and genetic algorithm. *International Journal of Environmental Research* 4, 4 (2010), 765–776. Citado en pp: 76