



---

**UNIVERSIDAD DE VALLADOLID**

MÁSTER  
GESTIÓN FORESTAL BASADA EN CIENCIA DE DATOS  
FOREST MANAGEMENT BASED ON DATA SCIENCE  
(DATAFOREST)

**Asignación automática de parcelas del Inventario  
Forestal Nacional a municipios usando Datos  
Abiertos Enlazados**

Autor: Natalia Crespo Lera

Tutores: Guillermo Vega Gorgojo - Irene Ruano Benito

Junio 2023



# Índice general

<b>1. INTRODUCCIÓN</b>	<b>1</b>
<b>2. OBJETIVOS</b>	<b>9</b>
<b>3. MATERIAL Y MÉTODOS</b>	<b>11</b>
3.1. Modelos de datos geospaciales . . . . .	11
3.1.1. Modelos de datos vectoriales . . . . .	12
3.1.2. Modelo semántico de datos geospaciales . . . . .	15
3.2. Descripción de los datos fuente . . . . .	22
3.2.1. Parcelas del IFN3 . . . . .	22
3.2.2. Recintos municipales españoles . . . . .	24
3.3. Procesado geoespacial . . . . .	27
3.3.1. Conversión de las geometrías de los municipios a RDF	28
3.3.2. Integración de datos . . . . .	29
3.3.2.1. Integración mediante SPARQL . . . . .	29
3.3.2.2. Integración mediante un SIG . . . . .	30
3.4. Comparación y validación de los resultados . . . . .	31
3.5. Aplicación de los resultados . . . . .	32
3.6. Publicación de los resultados como LOD . . . . .	33
3.7. Diagrama de flujo resumen de todo el proceso . . . . .	33
<b>4. RESULTADOS</b>	<b>35</b>
4.1. Conversión de los recintos municipales a RDF . . . . .	35
4.2. Asignación de las parcelas del IFN3 a los recintos municipales	37
4.2.1. Asignación mediante SPARQL . . . . .	38
4.2.1.1. Virtuoso . . . . .	38
4.2.1.2. Fuseki . . . . .	40
4.2.2. Asignación mediante un SIG . . . . .	41
4.2.3. Resultados de las asignaciones en RDF . . . . .	41
4.3. Comparación y validación de los resultados obtenidos . . .	42
4.4. Aplicación de los resultados . . . . .	45
4.4.1. Estadísticas a partir de los resultados obtenidos . . .	45

4.4.2. Mapas de especies dominantes por municipio . . . .	46
4.4.2.1. Parques Nacionales . . . . .	50
4.5. Publicación de los resultados en el <i>dataset Cross-Forest</i> . .	54
<b>5. DISCUSIÓN</b>	<b>55</b>
<b>6. CONCLUSIONES</b>	<b>61</b>
<b>7. AGRADECIMIENTOS</b>	<b>65</b>
<b>Referencias</b>	<b>67</b>
<b>Anexos</b>	<b>71</b>
<b>A. Lista de prefijos e IRI</b>	<b>73</b>
<b>B. Cambios en los códigos INE de los municipios</b>	<b>75</b>
<b>C. <i>Script</i> en R para generar el dígito de control de los municipios</b>	<b>77</b>
<b>D. Consulta SPARQL-Generate geometrías</b>	<b>79</b>
<b>E. Despliegue de un contenedor Docker de Virtuoso</b>	<b>81</b>
<b>F. Configuración de una instancia en Fuseki con GeoSPARQL</b>	<b>83</b>
<b>G. Consulta SPARQL-Generate QGIS</b>	<b>85</b>
<b>H. Asignación parcelas-municipios con Virtuoso</b>	<b>87</b>
<b>I. Asignación parcelas-municipios con Fuseki</b>	<b>89</b>
<b>J. Asignación parcelas-municipios con QGIS</b>	<b>91</b>
<b>K. <i>Script</i> en R comparación y validación</b>	<b>93</b>
<b>L. <i>Script</i> de R de especies dominantes por municipio</b>	<b>99</b>
<b>M. Mapas Especies dominantes Parques Nacionales</b>	<b>103</b>

## Resumen del TFM

El Inventario Forestal Nacional es una base de datos crucial para el estudio y gestión de los ecosistemas forestales nacionales. La información contenida en la tercera edición del inventario (IFN3) se ha incluido en la base de datos *Cross-Forest*, un repositorio integrado de bases de datos forestales de España y Portugal basada en *Linked Open Data* (LOD). LOD es un conjunto de buenas prácticas para publicar datos en la Web de manera estandarizada, fomentando la interconectividad entre las fuentes de datos mediante formatos y protocolos desarrollados por el W3C, como RDF, un modelo estandarizado para el intercambio de datos en la Web, y SPARQL, un lenguaje de consultas para datos RDF. Asimismo, el OGC ha creado GeoSPARQL como un modelo de representación y consulta de datos geospaciales en RDF.

El objetivo principal de este trabajo es identificar de forma automática los municipios donde se encuentran las parcelas del IFN3. Esto es especialmente relevante puesto que los municipios son la entidad administrativa básica en España con competencias en la gestión de los montes públicos. Además, muchos estudios utilizan esta división territorial como base para la realización de comparaciones y estadísticas. Con este propósito convertimos los 8 217 municipios, proporcionados por el Instituto Geográfico Nacional, a RDF. Posteriormente realizamos la asignación de las parcelas a los municipios mediante diferentes procedimientos, comparando la implementación GeoSPARQL en dos almacenes de tripletas, Virtuoso y Fuseki, con un Sistema de Información Geográfica como QGIS.

Tras la validación cruzada entre las tres asignaciones obtuvimos un 100 % de coincidencias entre Fuseki y QGIS y más del 70 % de falsos positivos empleando Virtuoso. Además, también cotejamos nuestros resultados con la información de los códigos municipales asociados a la parcelas en el IFN3, lo que demostró que la información disponible era contradictoria y poco fiable. Finalmente, a modo de muestra de la potencialidad de los resultados obtenidos al integrarlos con la base de datos *Cross-Forest*, llevamos a cabo mapas de las especies dominantes por municipio, teniendo en cuenta el área basimétrica ( $m^2/ha$ ) y el número de pies/ha. Con este trabajo, donde se identifica el municipio de la parcela del IFN3, se aporta una información valiosa para la gestión forestal a nivel municipal, comarcal y regional.

## Palabras clave

Inventario Forestal Nacional, parcelas, municipios, QGIS, Datos Abiertos Enlazados, GeoSPARQL

**Abstract**

The National Forest Inventory is a crucial database for the study and management of national forest ecosystems. The information contained in the third edition of the inventories (NFI3) has been included in the Cross-Forest database, an integrated repository of forest databases from Spain and Portugal based on Linked Open Data (LOD). LOD is a set of best practices for publishing data on the Web in a standardized manner, promoting interconnectivity between data sources through formats and protocols developed by the W3C, such as RDF, a standardized model for data exchange on the Web, and SPARQL, a query language for RDF data. Additionally, the OGC has created GeoSPARQL as a model for representing and querying geospatial data in RDF.

The main objective of this work is to automatically identify the municipalities where the plots from the NFI3 are located. This is especially relevant since municipalities are the primary administrative entities responsible for public forest management in Spain. Besides, many studies rely on this territorial division for comparisons and statistical analyses. To achieve this purpose we converted the 8,217 municipalities, provided by the Spanish National Geographic Institute, to RDF. Subsequently, we employed different procedures to automatically assign the plots to the municipalities, comparing the GeoSPARQL implementation in two triplestores, Virtuoso and Fuseki, with a Geographic Information System such as QGIS.

Through cross-validation of the three assignments, we achieved a 100 % match between Fuseki and QGIS and over 70 % false positives using Virtuoso. Additionally, we compared our results with the information of the municipal codes associated with the plots in the NFI3, which confirmed that the available information was contradictory and unreliable. Finally, to demonstrate the potential of the obtained results when integrated with the Cross-Forest database, we created maps showing the dominant species per municipality, considering the basal area ( $m^2/ha$ ) and the number of trees per hectare. This work, where the municipality of the NFI3 plots is identified, provides valuable information for forest management at the municipal, regional, and county levels.

**Keywords**

Spanish National Forest Inventory, plots, municipalities, QGIS, Linked Open Data, GeoSPARQL

# Índice de figuras

1.1. Información disponible y comparación de los códigos INE de los municipios asociados a parcelas en el IFN3. . . . .	4
3.1. Diagrama de flujo que resume la metodología empleada. . .	11
3.2. Ejemplo parcela del IFN3 en Forest Explorer . . . . .	13
3.3. Extracto de la ontología del IFN del <i>dataset Cross-Forest</i> . .	16
3.4. Ejemplo consulta SPARQL para contar el número de árboles de las parcelas del IFN3. . . . .	17
3.5. Esquema de 5 estrellas propuesto por Tim Berners-Lee . . .	19
3.6. Ejemplo grafo . . . . .	20
3.7. Clases y propiedades de GeoSPARQL . . . . .	20
3.8. Malla 1x1km de las parcelas del IFN3 . . . . .	23
3.9. Recintos municipales de España de la base de datos del IGN de 2023. . . . .	26
3.10. Consulta SPARQL municipio de Palencia . . . . .	27
3.11. Diagrama de la conversión de municipios a RDF . . . . .	28
3.12. Diagrama resumen de la metodología . . . . .	34
4.1. Ejemplo formato GeoJSON generado . . . . .	36
4.2. Ejemplo formato <i>Turtle</i> generado . . . . .	36
4.3. Consulta SPARQL CONSTRUCT para la descarga de las parcelas . . . . .	37
4.4. Consulta SPARQL CONSTRUCT para la descarga de los municipios . . . . .	37
4.5. Consulta SPARQL en Virtuoso para la asignación de parcelas a municipios . . . . .	38
4.6. Asignación en Virtuoso de una parcela a varios municipios .	40
4.7. Consulta SPARQL en Fuseki para la asignación de parcelas a municipios . . . . .	40
4.8. Ejemplo en formato <i>Turtle</i> de la asignación parcela-municipio	42
4.9. Consulta SPARQL para obtener el número de parcelas del IFN3 en cada municipio (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A). . . . .	45

4.10. Consulta SPARQL para obtener la media de parcelas del IFN3 por municipio en España (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A). . . . .	46
4.11. Consulta SPARQL para saber cuántos municipios españoles no tienen parcelas del IFN3 (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A). . . . .	46
4.12. Consulta SPARQL G media por especie y por municipio . . . . .	47
4.13. Consulta SPARQL N media por especie y por municipio . . . . .	47
4.14. Mapa de especies dominantes por municipio (m <sup>2</sup> /ha) . . . . .	48
4.15. Mapa de especies dominantes por municipio (pies/ha) . . . . .	49
4.16. Consulta SPARQL número de parcelas del IFN3 en Picos de Europa . . . . .	50
M.1. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y el número de pies/ha y parcelas del IFN3 del Parque Nacional de Tablas de Daimiel . . . . .	103
M.2. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Aigüestortes y Lago de San Mauricio . . . . .	104
M.3. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Aigüestortes y Lago de San Mauricio . . . . .	104
M.4. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Cabañeros . . . . .	105
M.5. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Cabañeros . . . . .	105
M.6. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional del Archipiélado de Cabrera . . . . .	106
M.7. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional del Archipiélado de Cabrera . . . . .	106
M.8. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Caldera de Taburiente . . . . .	107
M.9. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Caldera de Taburiente . . . . .	107

M.10. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Doñana . . . . .	108
M.11. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Doñana . . . . .	108
M.12. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Garajonay . . . . .	109
M.13. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Garajonay . . . . .	109
M.14. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de las Islas Atlánticas . . . . .	110
M.15. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de las Islas Atlánticas . . . . .	110
M.16. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Monfragüe . . . . .	111
M.17. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Monfragüe . . . . .	111
M.18. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Ordesa y Monte Perdido . . . . .	112
M.19. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Ordesa y Monte Perdido . . . . .	112
M.20. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Picos de Europa . . . . .	113
M.21. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Picos de Europa . . . . .	113
M.22. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Sierra de Guadarrama . . . . .	114

M.23. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Sierra de Guadarrama . . . . .	114
M.24. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Sierra de las Nieves . . . . .	115
M.25. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Sierra de las Nieves . . . . .	115
M.26. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional de Sierra Nevada . . . . .	116
M.27. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Sierra Nevada . . . . .	116
M.28. Mapa de especies dominantes por municipio según el área basimétrica (m <sup>2</sup> /ha) y parcelas del IFN3 del Parque Nacional del Teide . . . . .	117
M.29. Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional del Teide . . . . .	117

# Índice de tablas

3.1.	Discrepancias en las parcelas del IFN3 . . . . .	24
3.2.	Municipio de Palencia en el <i>dataset Cross-Forest</i> . . . . .	27
4.1.	Ejemplo de integración del polígono de los límites del municipio de Palencia con sus propiedades. . . . .	38
4.2.	Municipios asignados por Virtuoso a la parcela 22-0139-A-1 utilizando la consulta de la Figura 4.5 (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A). . . . .	39
4.3.	Asignación mediante QGIS de una parcela a su municipio . . . . .	41
4.4.	Comparación de la asignación automática de parcela-municipio . . . . .	43
4.5.	Parcelas sin municipio . . . . .	44
4.6.	Comparación de la asignación parcela-municipio del IFN3 . . . . .	45
4.7.	Especies dominantes por municipio de los Parques Nacionales . . . . .	51
A.1.	Prefijos y sus IRI correspondientes utilizados en este trabajo por orden de aparición . . . . .	73
B.1.	Cambios en los códigos INE de los municipios desde 2007 hasta 2023 . . . . .	75



## Capítulo 1

# INTRODUCCIÓN

En los últimos años, los datos abiertos (*open data*) han experimentado un crecimiento exponencial en todo el mundo, lo que ha favorecido un incremento en la transparencia, la accesibilidad y el valor de la información (Kitchin, 2014). Además, se estima que alrededor del 80 % de los datos abiertos contienen información geoespacial (Sieber, 2019). Los datos geospaciales se definen como cualquier dato que haga referencia, directa o indirectamente, a una localización o zona geográfica específica sobre la superficie terrestre (BOE, 2007). Las dos formas más comunes de representar datos geospaciales son mediante 1) el formato vectorial, que plasma la información geoespacial a través de puntos, líneas o polígonos (datos discretos); y 2) el formato ráster, que representa los datos de forma continua como una cuadrícula de celdas o píxeles (Lillesand et al., 2008).

Los datos geospaciales son complejos y normalmente no solo contienen información acerca de la posición de un objeto, sino que también incluyen atributos vinculados a este, por lo que sirven de referencia para información de cualquier ámbito (de la Fuente, 2020; Kügeler, 2022). Concretamente en los datos vectoriales, en los que nos vamos a centrar en este trabajo, cada fenómeno discreto se denomina *feature* (entidad espacial) que puede llevar asociada una serie de características y se representa mediante una geometría (Sayar et al., 2014; W3C, 2017). A pesar de tener una estructura definida, la información geoespacial es conocida por componer grandes volúmenes de datos extremadamente heterogéneos, lo que ha dado lugar a que la interoperabilidad de este tipo de datos se convierta en uno de los problemas más importantes a resolver por la comunidad de información geoespacial (Breunig et al., 2020; Brodeur et al., 2003).

La heterogeneidad en los datos geospaciales ha propiciado la creación de organizaciones a nivel mundial que fijen estándares y recomendaciones para facilitar su gestión e integración, como son la OGC (*Open Geospatial*

*Consortium*<sup>1</sup>), el W3C (*World Wide Web Consortium*<sup>2</sup>) o la serie de Normas ISO (*International Organization for Standardization*<sup>3</sup>) 19100 sobre información geográfica. En Europa, la Directiva europea INSPIRE (BOE, 2007) ha sentado las bases para la interoperabilidad de los datos geospaciales con el propósito de solucionar los problemas sintácticos, estructurales y las heterogeneidades semánticas que existen entre las bases de datos abiertas que contienen información espacial. INSPIRE no solo establece un marco para la interoperabilidad de los datos geoespaciales a nivel europeo, sino que también proporciona directrices para la creación de metadatos y servicios de datos geográficos. Estos servicios de datos deben cumplir con ciertas especificaciones técnicas incluyendo el uso de modelos de datos semánticos para asegurar la interoperabilidad y el intercambio de información abierta en la Web (BOE, 2007).

La interoperabilidad de los datos geoespaciales es esencial en múltiples ámbitos, especialmente en la implementación de políticas medioambientales y la gestión de recursos naturales a nivel nacional, europeo o mundial. En este sentido, la información geoespacial desempeña un papel fundamental en el monitoreo, análisis y visualización de ecosistemas forestales a lo largo del tiempo; proporcionando datos valiosos para la planificación y la toma de decisiones informadas en la gestión forestal, la conservación de la biodiversidad y la investigación en dichos ecosistemas (Brosofske et al., 2013; Guerra-Hernández & Pascual, 2021; Meddens et al., 2022; Otsu et al., 2011; Saeedeh, 2015).

En España, una de las bases de datos geospaciales sobre los ecosistemas forestales nacionales más importante y compleja es el Inventario Forestal Nacional (IFN). El IFN es un proyecto multiobjetivo a largo plazo que empezó con el primer inventario en 1966, y tras casi 60 años de recorrido, está terminando su cuarta edición (MITECO, s.f.-b). La finalidad del IFN es proporcionar información actualizada, aproximadamente cada 10 años, sobre el estado y evolución de los ecosistemas forestales españoles mediante un muestreo sistemático de recogida de datos en campo (Alberdi et al., 2017). La información contenida en el IFN está organizada a nivel provincial y se asocia a alrededor de 90 000 parcelas permanentes y georreferenciadas (localizadas mediante un rejón metálico enterrado en el centro de estas) a las que se intenta seguir su progreso a lo largo del tiempo (Alberdi et al., 2016).

Los datos de las diferentes ediciones del IFN se encuentran publicados en

---

<sup>1</sup><https://www.ogc.org/>

<sup>2</sup><https://www.w3.org/>

<sup>3</sup><https://www.iso.org/home.html>

la página Web del Ministerio para la Transición Ecológica y el Reto Demográfico (MITECO<sup>4</sup>). La información está disponible en una gran cantidad de ficheros agrupados por provincias que varían en formato según la edición del inventario (archivo de base de datos dBASE en la segunda, *Microsoft Access* en la tercera). Esto dificulta en gran medida el seguimiento de las parcelas a lo largo del tiempo y su integración en un solo archivo que reflejaría la información del conjunto del país. Actualmente, la última edición del IFN de la que tenemos información completa es la tercera edición (IFN3), llevada a cabo entre 1997 y 2007, de la que contamos con un total de 100 archivos entre los datos recogidos en campo y los obtenidos mediante Sistemas de Información Geográfica (SIG).

Además, si bien los datos del IFN se encuentran clasificados por provincias, es importante tener en cuenta que la unidad más pequeña de organización del territorio que tiene competencias en gestión forestal son los municipios. Según la Ley 7/1985 Reguladora de las Bases del Régimen Local, el municipio es “la entidad local básica de la organización territorial del Estado. Tiene personalidad jurídica y plena capacidad para el cumplimiento de sus fines” (BOE, 1985). Por otro lado, en la Ley de Montes “se revitaliza el papel de las Administraciones locales en la política forestal, concediéndoles una mayor participación en la adopción de decisiones que inciden directamente sobre sus propios montes, reconociendo con ello su papel como principales propietarios forestales públicos en España... Son los propietarios de los montes los que primero y más directamente se responsabilizan de su gestión sostenible” (BOE, 2003).

Para conocer el territorio que comprenden los municipios españoles, el Instituto Geográfico Nacional<sup>5</sup> (IGN) es el responsable de la publicación de la cartografía oficial de los límites y recintos municipales y su información asociada. En la base de datos espacial del IGN, que se creó en 2009 y que se actualiza de forma periódica, se incluyen tanto las delimitaciones territoriales inscritas en el Registro Central de Cartografía según lo descrito en el Real Decreto 1545/2007<sup>6</sup>, de 23 de noviembre, por el que se regula el Sistema Cartográfico Nacional<sup>6</sup>, como los códigos de los municipios del Instituto Nacional de Estadística (INE<sup>7</sup>).

Sin embargo, en las distintas ediciones del IFN los códigos de los municipios solo se encuentran parcialmente disponibles. Además, de acuerdo con

---

<sup>4</sup><https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/informacion-disponible/ifn3.aspx>

<sup>5</sup><https://www.ign.es>

<sup>6</sup>[https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2007-20556](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2007-20556)

<sup>7</sup><https://www.ine.es/>

Tragsatec, la empresa que colabora con el MITECO en la elaboración del IFN, no es posible asegurar que la asociación de las parcelas a un municipio determinado en los inventarios sea precisa. Las principales razones son que esta información no ha sido verificada y que la identificación de los municipios a los que pertenecen las parcelas de forma exacta, tanto en gabinete como en campo, muchas veces era compleja con los medios disponibles en la época en la que se llevó a cabo esta asignación (1997-2007).

En el IFN3 existen dos Tablas que cuentan con los códigos INE de los municipios asociados a las parcelas. Una de ellas, *PCDatosMap*, se realizó antes de los muestreos de campo (en gabinete) y la información que contiene es residual de la capa cartográfica de la que disponían en la época en la que se llevó a cabo el inventario. La otra tabla, *PCParcelas*, comprende los datos de los códigos municipales relativos a la asignación realizada por el personal de campo durante el apeo de las parcelas. La información contenida en estas dos tablas en algunos casos es contradictoria y no existe para todas las parcelas del inventario (Figura 1.1).

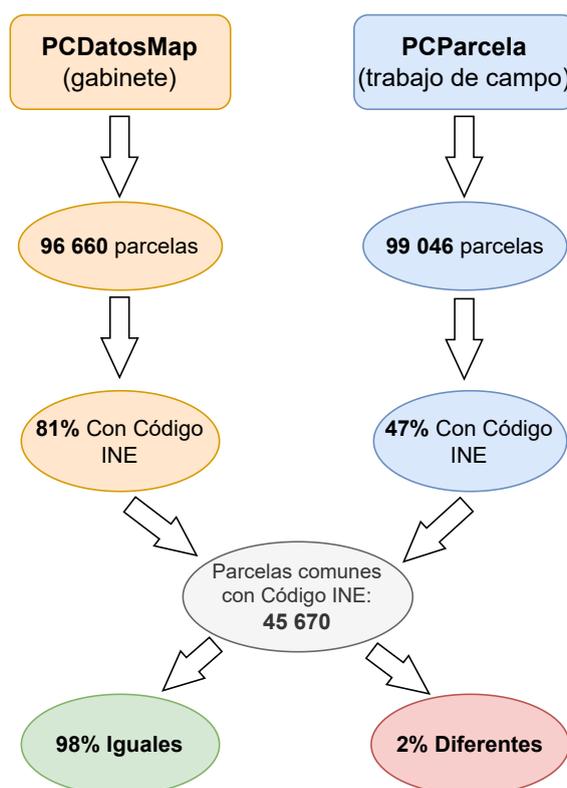


Figura 1.1: Información disponible y comparación de los códigos INE de los municipios asociados a parcelas en el IFN3.

No obstante, la información relativa a los municipios en los que se encuentran las parcelas del IFN es muy útil para la gestión forestal a nivel local

y para muchos estudios que utilizan los términos municipales como base para el análisis y la comparación de áreas forestales en temas tan amplios como las áreas reforestadas (Marey-Pérez & Rodríguez-Vicente, 2009), la biodiversidad (Guadilla-Sáez et al., 2019), la captura de carbono (Gil et al., 2011; Lorenzo-Sáez et al., 2022) o los incendios forestales (Martínez et al., 2009). Los resultados de estos estudios contribuyen a una mejora en la planificación y gestión forestal a futuro y también a un mejor manejo de los recursos naturales y su conservación. Además, desde un punto de vista socioeconómico, la asignación de las parcelas del inventario al municipio al que pertenecen también posibilita la realización de análisis sobre la contribución del sector forestal a la economía local y regional (Ortuño-Pérez & Martín-Fernández, 2006; Sánchez-García et al., 2015).

Por todo ello, en el dominio forestal es cada vez más importante la integración de datos abiertos geospaciales, ya que, en la actualidad, nos encontramos con grandes cantidades de bases de datos que funcionan independientemente cuya integración enriquecería en gran medida los ya existentes. A pesar de todo, la integración de datos espaciales es difícil debido en gran medida a que este tipo de información es típicamente manejada y procesada principalmente por SIG, como QGIS<sup>8</sup> o ArcGIS<sup>9</sup>, que utilizan sistemas de gestión de bases de datos relacionales extremadamente eficientes y optimizadas para el procesamiento centralizado de consultas geoespaciales (Longley et al., 2005). Sin embargo, los SIG no fueron desarrollados teniendo en cuenta la naturaleza abierta y descentralizada de los datos Web y no resuelven el problema de la interoperabilidad de los datos. Por lo que la integración y homogeneización de conjuntos de datos externos en estos sistemas es un proceso complejo y que, generalmente, consume mucho tiempo (Zhang et al., 2008). Por este motivo, para la integración de los límites municipales y las parcelas del IFN de toda España, es especialmente útil emplear una estructura semántica común.

Las tecnologías de la Web Semántica, estandarizadas por el W3C, pretenden resolver el problema de la interoperabilidad mediante la creación de la llamada “Web de Datos”. La Web Semántica representa la información mediante grafos, una forma sencilla de representar relaciones entre entidades haciendo que estas sean más fácilmente procesables por las máquinas. De esta manera se facilita la búsqueda, reutilización e intercambio de información entre bases de datos. A los datos abiertos que se comparten siguiendo los principios de los datos enlazados se denominan datos abiertos enlazados

---

<sup>8</sup><https://qgis.org/es/site/>

<sup>9</sup><https://www.arcgis.com/index.html>

(*Linked Open Data* o LOD) (Pascal, 2021). Tomando en cuenta la potencialidad de los LOD, surge el proyecto Europeo *Cross-Forest*<sup>10</sup>, que finalizó recientemente y en el que participó el grupo de investigación donde se ha desarrollado este Trabajo de Fin de Máster (TFM), con la idea de crear un conjunto de datos abiertos enlazados y de uso libre para la integración de bases de datos forestales.

LOD es un conjunto de principios y buenas prácticas para publicar y conectar datos estructurados en la Web, lo que permite una mejor interoperabilidad e integración de las bases de datos (Bizer et al., 2009). Esto se logra utilizando un conjunto de tecnologías de la Web Semántica y los estándares desarrollados por el W3C, como RDF (*Resource Description Framework*), RDFS (*RDF Schema*), OWL (*Web Ontology Language*) y SPARQL (*SPARQL Protocol and RDF Query Language*). Estos estándares permiten expresar la información de manera semántica mediante tripletas (sujeto-predicado-objeto), que se almacenan en *triplestores* o almacenes de tripletas, y enlazar los datos a través de identificadores únicos (IRI) y vocabularios u ontologías comunes (Berners-Lee, 2006). Por estos motivos, LOD presentan numerosas ventajas para resolver la gran heterogeneidad que tienen los datos geospaciales.

Para facilitar la consulta y manejo de datos geospaciales, la OGC ha desarrollado la especificación GeoSPARQL (OGC, 2012b), una ontología con un conjunto de funciones espaciales para su uso integrado con el lenguaje de consultas SPARQL. GeoSPARQL soporta la representación de *features* mediante diferentes tipos de geometrías y funciones geospaciales como `st_within`<sup>11</sup>, mejorando la interoperabilidad y la integración de los datos geospaciales (OGC, 2012b).

Muchos almacenes de tripletas, como por ejemplo Virtuoso<sup>12</sup> (almacén utilizado en la base de datos *Cross-Forest* en su versión de libre uso) o Fuseki<sup>13</sup> (almacén de libre uso cada vez más popular) dan soporte a GeoSPARQL permitiendo realizar consultas SPARQL sobre datos geoespaciales anotados semánticamente. Sin embargo, el soporte a GeoSPARQL puede variar mucho de una implementación a otra, lo que es necesario tener en cuenta a la hora de utilizar diferentes almacenes de tripletas para realizar consultas geospaciales (Jovanovik et al., 2021).

---

<sup>10</sup><https://crossforest.eu/>

<sup>11</sup>`st_within` es una función estandarizada por la OGC que también utilizan los SIG, y que sirve para consultar si una entidad espacial o *feature* esta contenida en otra, lo cual es muy útil para, como en el problema planteado en este trabajo, se pretende localizar parcelas dentro de términos municipales.

<sup>12</sup><https://virtuoso.openlinksw.com/>

<sup>13</sup><https://jena.apache.org/documentation/fuseki2/>

La información espacial contenida en el conjunto de datos *Cross-Forest*, anteriormente mencionado, esta publicada actualmente en un punto SPARQL<sup>14</sup> utilizando la versión de libre uso del almacén de tripletas Virtuoso (versión 07.20.3230) que permite gestionar y consultar grafos RDF utilizando el lenguaje de consultas SPARQL. Este conjunto de datos publicados como LOD contiene información de España y Portugal relativa a los inventarios forestales nacionales más actuales y los mapas de ocupación y usos del suelo. En el conjunto de datos *Cross-Forest*, que de aquí en adelante se hará referencia a él con el término “*dataset Cross-Forest*”, se incluye, concretamente, la tercera edición del Inventario Forestal Nacional Español (IFN3), el Mapa Forestal Español a escala 1:50.000 (MFE50), el Sexto Inventario Forestal Nacional Portugues (IFN6) y la Cartografía de Ocupación de Suelos de Portugal (COS2018) (Fierro-Garcia et al., 2022).

Además, en el *dataset Cross-Forest* también se incluye la información (provincia, nombre, código INE, dígito de control) de los 8 131 municipios oficiales españoles inscritos en el Registro de Entidades Locales que fueron integrados dentro del proyecto Erasmus+ VirtualForests<sup>15</sup>. Sin embargo, no se incluyen las geometrías asociadas a estos municipios que definen los límites territoriales de los mismos. Asimismo, el *dataset Cross-Forest* puede visualizarse fácilmente empleando el mapa interactivo ofrecido por la herramienta *Web Forest-Explorer*<sup>16</sup> (Vega-Gorgojo et al., 2022).

Uno de los propósitos del *dataset Cross-Forest* es seguir un proceso de mejora continua mediante la integración de nuevos conjuntos de datos procedentes de diferentes bases de datos abiertas que enriquezcan los ya existentes. De esta manera, se pretende promover aún más las ventajas que ofrecen los LOD para facilitar el uso y mejorar la calidad de las bases de datos forestales. Actualmente, el grupo de investigación en el que se ha llevado a cabo este TFM está realizando el proyecto LOD.For.Trees que tiene como uno de sus objetivos la integración de la segunda y cuarta edición del IFN.

En el proceso de curado del *dataset Cross-Forest* se ha visto la necesidad de mejorar los datos de los municipios ya existentes incluyendo las geometrías de los límites municipales. De esta manera, sería posible realizar la asignación automática de las posiciones de las parcelas del IFN3, ya incluidas en la base de datos, a sus municipios correctos. Esta nueva información facilitaría la realización de los estudios y gestión de las áreas forestales por

---

<sup>14</sup><https://crossforest.gsic.uva.es/sparql>

<sup>15</sup><https://virtualforests.eu>

<sup>16</sup>[forestexplorer.gsic.uva.es/](https://forestexplorer.gsic.uva.es/)

municipio, como se ha presentado anteriormente, permitiendo llevar a cabo estadísticas, mapas y planes de gestión a nivel municipal.

## Capítulo 2

# OBJETIVOS

La falta de información fiable de la localización de las parcelas de la tercera edición del Inventario Forestal Nacional (IFN3) en sus respectivos términos municipales y la utilización de esta información en diferentes estudios y planes de gestión, ha hecho que se crea necesario desarrollar un procedimiento para realizar esta asignación de forma automática y de la manera más precisa posible. Con esto pretendemos incrementar y complementar la gran cantidad de datos que componen la base de datos del IFN3 y que, como hemos mencionado en el Capítulo 1, se encuentra incluida en el *dataset Cross-Forest* como LOD, lo que facilita la integración e interoperabilidad de la información disponible. Junto con esto, se pretende hacer este resultado público y fácilmente accesible, siguiendo los principios LOD.

Para ello, se definen los siguientes objetivos:

1. **Diseñar un flujo de trabajo para convertir los recintos municipales a RDF**, actualmente publicados en la base de datos geoespacial del Instituto Geográfico Nacional (IGN) en formato *shapefile*.
2. **Asignar de forma automática las parcelas del IFN3 a sus municipios** mediante diferentes procedimientos, utilizando por un lado GeoSPARQL a partir de LOD, y por otro lado, un Sistema de Información Geográfica.
3. **Comparar y validar las diferentes aproximaciones** en la asignación de las parcelas a sus respectivos municipios y cotejarlas con la información que se encuentra disponible en el IFN3.
4. **Aplicar la integración de los municipios con el resto de información forestal** en la obtención de estadísticas o métricas a nivel municipal.
5. **Publicar como LOD en el *dataset Cross-Forest*** tanto los datos en formato RDF de las geometrías de los municipios españoles, como la asignación de las parcelas del IFN3 a sus códigos municipales.

El fin último de este TFM es mejorar los datos del IFN3 y posibilitar la utilización de esta información en formato LOD para favorecer el análisis y la gestión sostenible de los ecosistemas forestales nacionales a nivel local. Asimismo, se pretende mostrar la potencialidad de los LOD en la integración de diferentes bases de datos geospaciales y las diversas posibilidades que brinda en la realización de estadísticas y métricas a nivel municipal con los datos forestales disponibles.

## Capítulo 3

# MATERIAL Y MÉTODOS

En la presente Capítulo se describe la metodología utilizada para realizar la conversión de las geometrías de los recintos municipales españoles a datos abiertos enlazados (LOD), su asignación a las parcelas de la tercera edición del Inventario Forestal Nacional (IFN3), su validación y por último su integración en el *dataset Cross-Forest*. Con el propósito de proporcionar una visión general del proceso empleado, se incluye en la Figura 3.1 un diagrama de flujo que ilustra las diferentes etapas llevadas a cabo para lograr los objetivos planteados. En las siguientes secciones se profundizará en las fases representadas en dicho diagrama de flujo.

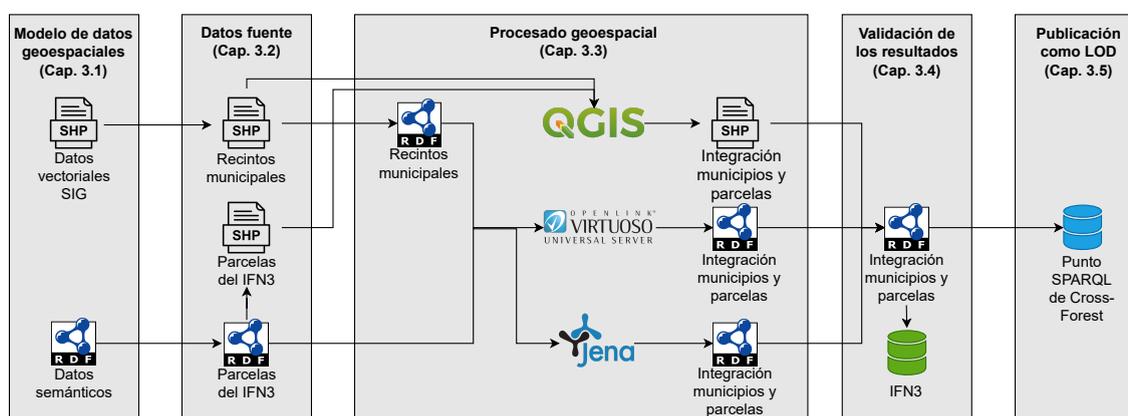


Figura 3.1: Diagrama de flujo que resume la metodología empleada.

### 3.1. Modelos de datos geospaciales

Los Sistemas de Información Geográfica (SIG) utilizan principalmente dos modelos lógicos para representar los datos geospaciales dependiendo de las características particulares del tipo de información que se quiere representar (Longley et al., 2005):

- **Modelos vectoriales:** representan objetos geospaciales mediante tres estructuras básicas (puntos, líneas y polígonos) que se almacenan en forma de coordenadas de los vértices de cada elemento geométrico y que pueden estar asociados a unos atributos. Este tipo de datos son útiles para el estudio más preciso de objetos discretos, como las parcelas del Inventario Forestal Nacional (IFN).
- **Modelos ráster:** son imágenes digitales que representan la superficie terrestre mediante una red formada por celdas o píxeles. Cada píxel tiene un valor que indica una propiedad específica como la pendiente, la temperatura... Los modelos ráster son útiles para el análisis de superficies continuas y pueden simular la realidad tridimensional de forma más fiable.

Por otro lado, la Web Semántica soporta todo tipo de datos geospaciales en RDF, tanto vectoriales como ráster. Sin embargo, en este TFM nos vamos a centrar en los modelos de datos vectoriales. Esto es debido a que las bases de datos a integrar, el IFN3 y los recintos municipales de la base de datos geospacial del IGN, solo cuentan con este formato de datos geospaciales en la actualidad.

### 3.1.1. Modelos de datos vectoriales

Los datos vectoriales son un modelo de datos ampliamente utilizado para representar y almacenar los datos geospaciales de manera estructurada. Para ello utilizan tres elementos principales: *feature*, *geometry* y *spatial object*. El *World Wide Web Consortium* (W3C), que ha definido una serie de estándares para el intercambio y la representación de datos geoespaciales en la Web, describe los tres componentes principales que forman la estructura básica de los datos geospaciales vectoriales de la siguiente manera (Taylor et al., 2017):

- ***Feature* (entidad geospacial):** representa un objeto geográfico individual. En nuestro caso concreto se podría ejemplificar como un árbol. Cada *feature* puede estar asociada con una serie de atributos que describen características específicas de la misma, como la especie o el diámetro del árbol. Estos datos se suelen almacenar en una Tabla de atributos que se vincula a su geometría a través de un identificador único para cada *feature*. Además, una misma *feature* puede estar asociada a una o más geometrías diferentes.
- ***Geometry*:** es la forma y ubicación física del *feature*. Las geometrías básicas son puntos, líneas y polígonos. En el espacio bidimensional, un punto se describe mediante un par de coordenadas únicas, mientras

que una línea o un polígono se representan como una lista ordenada de vértices que conforman la geometría.

- **Spatial object:** es la combinación de una *feature* y su geometría. Por ejemplo, un objeto espacial puede ser un árbol representado mediante un punto, con un par de coordenadas únicas, que se encuentra, o no, asociado a unos atributos. Un objeto espacial a su vez puede contener un conjunto de objetos espaciales, como es el caso de una parcela del IFN y los árboles que contiene. Las parcelas del IFN tienen una geometría que se asocia a características propias de la entidad parcela (área basimétrica, número de pies...). Sin embargo, cada parcela a su vez puede contener otros objetos espaciales, árboles, con otros atributos asociados (especie, diámetro, altura) (Figura 3.2).



Figura 3.2: Ejemplo de una parcela del IFN3 con sus atributos y árboles asociados. Fuente: <https://forestexplorer.gsic.uva.es/explorer/>.

Además, existen varios formatos de datos vectoriales disponibles para almacenar y compartir información espacial. Cada formato tiene su propio conjunto de características, lo que los hace adecuados para diferentes usos y aplicaciones en el ámbito geoespacial. Los formatos de datos vectoriales más comunes son:

- **Shapefile**<sup>1</sup> es un formato propietario desarrollado por ESRI<sup>2</sup> a principios de los años 90 cuya especificación es abierta. Hoy en día es el más

<sup>1</sup><https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf>

<sup>2</sup><https://www.esri.es/es-es/home>

utilizado en la comunidad SIG para el almacenamiento de datos vectoriales. Un *Shapefile* se compone de varios ficheros binarios que son leídos por los SIG como si fueran uno solo. Los archivos básicos son tres: el archivo .shp que almacena la geometría de la entidad, el archivo .shx contiene un índice de los objetos espaciales que mejora el acceso a los datos y el archivo .dbf que se presenta en formato de base de datos dBASE donde se almacenan los atributos. Los *Shapefile* están diseñados para almacenar datos vectoriales sencillos, cada archivo solo admite un tipo de geometría a la vez, es decir, no se puede almacenar un punto y un polígono en el mismo archivo. Además, este formato tiene algunas deficiencias como la limitación en el tamaño de los archivos (2GB) o en la definición de los atributos (ESRI, s.f.).

- **GeoPackage**<sup>3</sup> es un formato de archivo de datos geospaciales de código abierto que admite tanto datos vectoriales como ráster. Fue introducido en 2 014 como una alternativa al formato de archivo *Shapefile*, aunque actualmente todavía no es tan ampliamente utilizado. A diferencia del *Shapefile*, el *GeoPackage* se basa en los estándares desarrollados por la OGC<sup>4</sup> y utiliza una base de datos SQLite para almacenar la información en un formato binario complejo permitiendo guardar múltiples capas de datos en un solo archivo facilitando su manejo. Cada capa de datos representa una vista cartográfica temática de la información geoespacial, como parcelas (polígonos), árboles (puntos), ríos (líneas), etc. (OGC, 2021).
- **GeoJSON**<sup>5</sup> es un formato abierto utilizado para almacenar y representar entidades geospaciales sencillas de tipo vectorial basándose en la sintaxis de JSON (*JavaScript Object Notation*). *GeoJSON* se propuso en 2 008 siendo un formato específico para representar objetos espaciales (*Point*, *LineString*, *Polygon*, *MultiPoint*, *MultiLineString*, *MultiPolygon*, *GeometryCollection*) y sus atributos no espaciales. La gramática utilizada para definir estos datos geospaciales se basa en el estándar WKT. *GeoJSON* es comúnmente utilizado en aplicaciones Web de mapas, por ser un formato de texto sencillo, ampliamente extendido y compatible con otros formatos, programas o herramientas (Butler & Daly, 2016).
- **WKT**<sup>6</sup> (*Well-Known Text*) es una codificación de caracteres ASCII estandarizada por la OGC en 1 999. WKT se utiliza para describir objetos geospaciales vectoriales de forma simple. Su uso está muy extendido

---

<sup>3</sup><https://www.geopackage.org/>

<sup>4</sup><https://www.ogc.org/>

<sup>5</sup><https://geojson.org/>

<sup>6</sup><https://www.ogc.org/standard/wkt-crs/>

en el ámbito de la geoinformática siendo la base de otros formatos más conocidos como KML o GeoJSON, mencionado anteriormente.

- **GML** (*Geography Markup Language*)<sup>7</sup> es el estándar abierto para la representación de datos geospaciales con sintaxis XML (*eXtensible Markup Language*) definido por la OGC y lanzado por primera vez en el 2 000. A diferencia de WKT, GML puede describir información más compleja, como la topología y la semántica de los datos, sin embargo suelen ser ficheros de gran tamaño. La importancia del GML radica en que sirve como un modelo común en el ámbito informático para la transferencia y manejo de datos espaciales entre diversos sistemas que hacen uso de este tipo de datos, como los SIG. Se utiliza principalmente para describir datos vectoriales pero también sirve para representar datos ráster.

### 3.1.2. Modelo semántico de datos geospaciales

Los modelos semánticos de datos son una forma de representar la información de manera estructurada. Posibilitan la definición del significado de los datos en el contexto de sus relaciones con otros formando grafos (Figura 3.3). Las bases de datos semánticas conforman la Web Semántica, una extensión de la *World Wide Web* (WWW) que busca facilitar el procesamiento de los datos por parte de las máquinas permitiendo a los sistemas informáticos comprender mejor la información y realizar tareas complejas como el razonamiento y la inferencia (Pascal, 2021).

En el contexto de la Web Semántica es esencial contar con ontologías o vocabularios para definir la terminología de un dominio específico de interés, así como un lenguaje para anotar los datos de acuerdo con la ontología y un lenguaje de consultas para recuperar la información. Los datos semánticos se almacenan en gestores de bases de datos especializados conocidos como almacenes de tripletas, que incluyen un motor de procesamiento de consultas. El W3C ha desarrollado una serie de estándares para la Web Semántica, que incluyen RDF (*Resource Description Framework*) como lenguaje de modelado de datos, OWL (*Web Ontology Language*) como lenguaje para la construcción de ontologías o vocabularios y SPARQL (*Protocol and RDF Query Language*) como lenguaje de consultas de datos RDF:

- **RDF** es una especificación para representar información estructurada en la Web mediante un conjunto de tripletas sujeto-predicado-objeto (W3C, 2014). Estas tripletas se utilizan para describir recursos en la Web y sus

---

<sup>7</sup><https://www.ogc.org/standard/gml/>

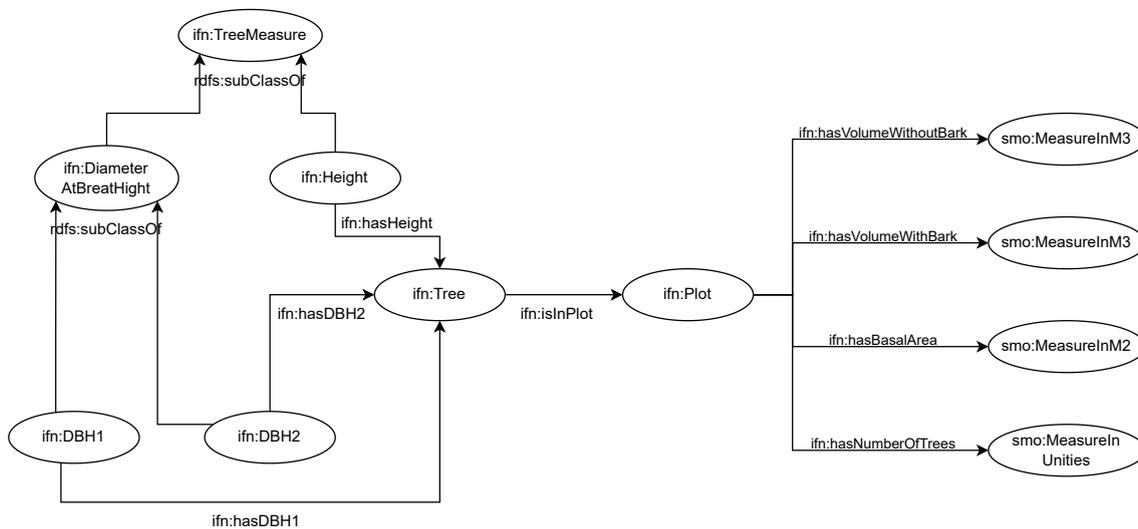


Figura 3.3: Extracto de la ontología del IFN del *dataset Cross-Forest* con sus clases y propiedades principales donde se puede observar las clases de medidas de altura y diámetro asociados a un árbol que a su vez esta asociado a una parcela. La parcela se interconecta mediante una sintaxis definida (propiedades como `ifn:hasBasalArea`, `ifn:hasNumberOfTrees`...) con otras clases asociadas (`smo:MeasureInM2`, `smo:MeasureInUnities`...) pertenecientes a la ontología de posiciones.

relaciones, permitiendo la creación de una red de datos enlazados. Existen varias serializaciones disponibles para escribir RDF entre las que destacan N3 (*Notation3*) y *Turtle*, alternativas no-XML específicas para representar datos en la Web Semántica. Por otro lado también está RDF/XML y JSON. *Turtle* (*Terse RDF Triple Language*) es un subconjunto simplificado de N3 muy utilizado para representar los datos en forma de tripletas. Cada uno de los elementos de una triplete se puede representar como una IRI (*Internationalized Resource Identifier*)<sup>8</sup>, o en el caso del objeto también puede ser un literal.

- **RDFS** es una extensión del vocabulario u ontología básico RDF. Proporciona una forma de definir clases y propiedades para recursos RDF permitiendo a los usuarios crear vocabularios personalizados para describir sus propios datos (W3C, 2023). Con RDFS, se pueden definir relaciones entre clases y propiedades, así como jerarquías de subclases y subpropiedades.
- **OWL** es un lenguaje formal que se utiliza para modelar ontologías, es decir, una representación estandarizada que se utiliza para describir un conjunto de datos y las relaciones entre ellos (W3C, 2012a). OWL permite definir etiquetas de clases, propiedades y restricciones sobre datos

<sup>8</sup><https://datatracker.ietf.org/doc/html/rfc3987>

RDF siendo más expresivo que RDFS ya que permite más construcciones.

- **SPARQL** es el lenguaje de consulta que se utiliza para recuperar información de grafos RDF en la Web Semántica (W3C, 2012b). SPARQL permite realizar consultas complejas sobre datos RDF utilizando una sintaxis similar a SQL. La consulta básica SPARQL esta formada por: *SELECT*, donde se especifican las variables (se identifican por un símbolo de interrogación de cierre) que se deben recuperar en la consulta, *DISTINCT* (opcional) para que los resultados sean únicos y no muestre duplicados, y *WHERE* que define los patrones triples (sujeto-predicado-objeto) que deben coincidir en los datos RDF para que se devuelvan los resultados. Asimismo, también se pueden poner prefijos (*PREFIX*) antes de la clausula *SELECT* para abreviar los elementos y simplificar las consultas (las IRI de los prefijos utilizados en este trabajo se encuentran en el Anexo A) y especificaciones adicionales como filtros (*FILTER*), agrupaciones (*GROUP BY*), orden (*ORDER BY*)... También se pueden crear grafos RDF cambiando *SELECT* (para realizar consultas) por *CONSTRUCT*, con una sintaxis similar a la anterior. Además, existen varias extensiones del lenguaje de consulta SPARQL, como SPARQL-Generate que permite generar datos RDF de forma automatizada o GeoSPARQL que posibilita trabajar con datos geospaciales en RDF.

En la Figura 3.4 se presenta un ejemplo de consulta SPARQL que cuenta el número de árboles que hay en cada parcela del IFN3 y ordena el resultado por el identificador de la parcela de menor a mayor, donde se pueden observar algunos de los elementos definidos anteriormente.

```
PREFIX ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/>

SELECT DISTINCT ?plot COUNT(DISTINCT ?tree)

WHERE {
  ?tree a ifn:Tree ;
        ifn:isInPlot ?plot
}

ORDER BY ASC(?plot)
```

Figura 3.4: Ejemplo consulta SPARQL para contar el número de árboles de las parcelas del IFN3.

Durante la primera década de los 2 000 se produjo un importante desarrollo de los lenguajes y tecnologías de la Web Semántica, como RDF, OWL y SPARQL. Sin embargo, aunque estos lenguajes permitían definir y representar información estructurada, había una falta de datos semánticos

disponibles en la Web (Pascal, 2021). Esto generó un gran escepticismo ya que las expectativas iniciales eran muy altas y se esperaba una adopción más amplia de la Web Semántica por parte de la industria y los usuarios. Sin embargo, a finales de los 2 000, surgieron nuevas iniciativas que cambiaron esta situación. Una de ellas fue *Linked Open Data* (LOD) impulsada por Tim Berners-Lee, el creador de la WWW, que propuso una serie de principios de datos enlazados ampliamente aceptados, estableciendo una serie de prácticas y estándares para la publicación de datos en la Web de forma enlazada y accesible (Pascal, 2021).

Para almacenar estos datos existen diferentes *triplestores* (también conocidos como almacenes de tripletas o *RDF stores*), sistemas de gestión de bases de datos diseñados para almacenar, gestionar y recuperar datos semánticos en formato RDF. Un *triplestore* permite almacenar grandes volúmenes de tripletas RDF (sujeto-predicado-objeto), proporcionando una infraestructura para consultas eficientes y búsqueda de información. Además, los *triplestores* ofrecen una serie de herramientas para el análisis y la gestión de los datos RDF, lo que los convierte en una pieza clave para la implementación de aplicaciones y servicios de la Web Semántica y los LOD.

En conjunto, la Web Semántica y los LOD ofrecen una visión de la Web en la que los datos están disponibles de manera más estructurada, coherente y fácil de usar, lo que permite una mejor interoperabilidad entre las bases de datos. Los 4 principios de los datos enlazados planteados por Tim Berners-Lee, se enumeran a continuación (Berners-Lee, 2006):

1. Utilizar IRI como identificadores únicos de los recursos en la Web.
2. Las IRI deben ser accesibles a través de HTTP (*Hypertext Transfer Protocol*) para que las personas puedan buscar y encontrar información sobre ellas.
3. Las IRI deben proporcionar información útil mediante el uso de estándares Web abiertos como RDF o SPARQL.
4. Las IRI deben estar enlazadas con otras IRI para permitir que las personas puedan acceder a más información y navegar a través de los datos enlazados.

Teniendo en cuenta estos principios, Tim Berners-Lee propuso un esquema de 5 niveles o estrellas para incentivar la publicación de datos abiertos, interconectados y de calidad en la Web (Figura 3.5) (Berners-Lee, 2006). A continuación se describen los 5 niveles cuyas características son aditivas desde el primer nivel hasta el último:

1. Datos disponible en la Web en cualquier formato.
2. Datos accesibles en un formato estructurado (este es el caso de la base de datos del IFN publicada por el MITECO).
3. Datos estructurados en un formato no propietario (este es el caso de la base de datos geospacial de los recintos municipales del IGN).
4. Datos que siguen los estándares de la W3C (vocabularios u ontologías comunes).
5. Datos enlazados con otros datos que permiten una mayor contextualización (este es el caso del *dataset Cross-Forest*).

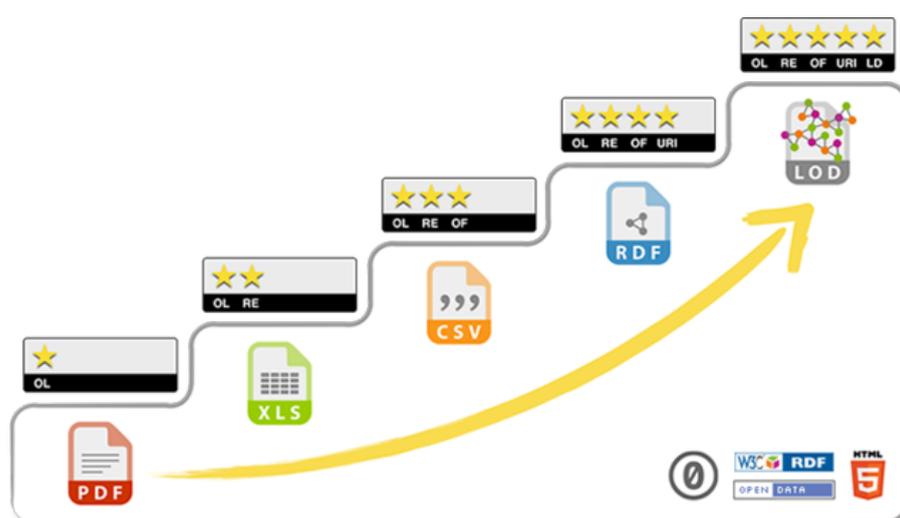


Figura 3.5: Esquema de 5 estrellas propuesto por Tim Berners-Lee para evaluar el nivel de disponibilidad e interconexión de los datos en la Web.

Las tecnologías de la Web Semántica y los principios LOD son aplicables a una amplia variedad de dominios, como se ejemplifica en la Figura 3.6. En el ámbito espacial, se han propuesto varias iniciativas para la representación de datos geospaciales, entre las cuales destaca GeoSPARQL y la ontología W3C Geo (Group, 2003). A diferencia de la ontología W3C Geo, que proporciona una representación básica y limitada de datos geospaciales, GeoSPARQL sigue los estándares y estructuras de los datos espaciales definidos por el OGC<sup>9</sup> (Figura 3.7) y permite realizar consultas y operaciones espaciales complejas en conjuntos de datos RDF (OGC, 2012a; Perry et al., 2012).

GeoSPARQL es una ontología pequeña y una extensión de SPARQL que soporta una amplia variedad de tipos de datos geospaciales, incluyendo pun-

<sup>9</sup><https://www.ogc.org/standard/geosparql/>

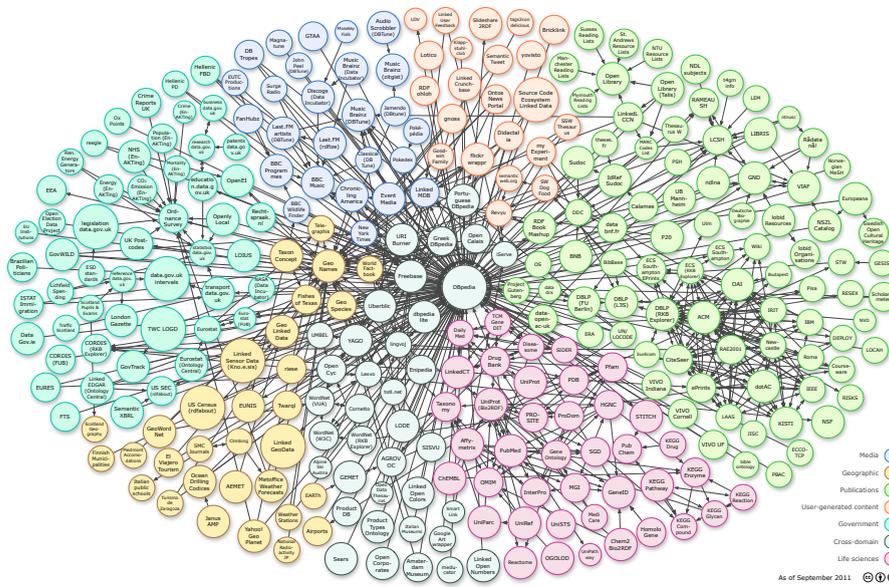


Figura 3.6: Grafo de ejemplo de los diferentes dominios que se pueden encontrar publicados en formato LOD en la Web. Fuente: <http://cas.lod-cloud.net/#diagram>.

tos, líneas y polígonos, así como operaciones geoespaciales comunes, como la interSección, la unión y la distancia entre geometrías. Los formatos espaciales para representar las geometrías de los objetos espaciales que utiliza GeoSPARQL son WKT y GML, tratados en la Sección 3.1.1. GeoSPARQL tiene varias aplicaciones prácticas ya que puede ser utilizado para integrar información geoespacial de diferentes fuentes y realizar análisis espaciales avanzados. Por este motivo, es el lenguaje de consulta que se utilizará en este trabajo para llevar a cabo la asignación de las geometrías de los municipales a las parcelas del IFN3.

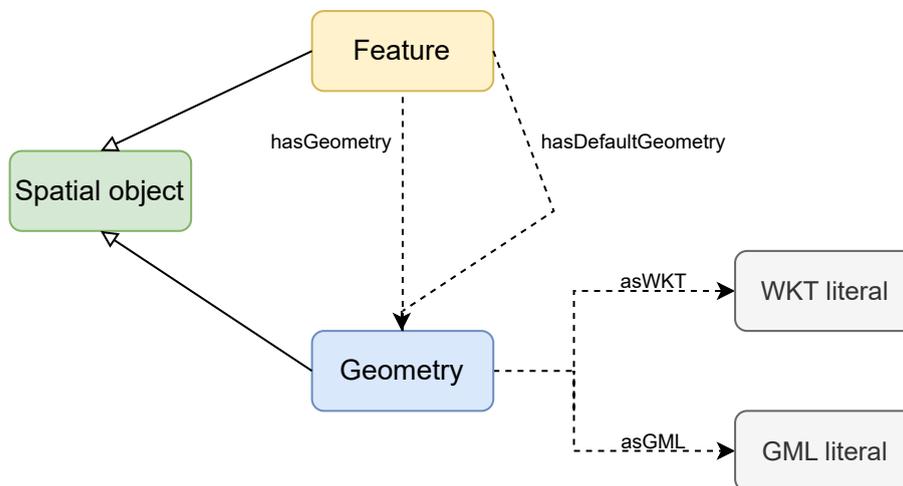


Figura 3.7: Diagrama que describe las clases que constituyen GeoSPARQL y sus propiedades.

Los almacenes de tripletas que dan soporte a GeoSPARQL permiten almacenar y consultar datos geospaciales de manera eficiente. Sin embargo, existen diferentes almacenes de tripletas que ofrecen soporte a GeoSPARQL, y cada uno tiene sus propias características. De acuerdo con el OGC (OGC, 2020) los productos que dan soporte a GeoSPARQL ordenados en orden alfabético son los siguientes: *Apache Jena Fuseki*, *Eclipse RDF4J*, *Ontop-spatial*, *Ontotext GraphDB*, *Openlink Virtuoso*, *Oracle Spatial and Graph*, *Parliament*, *Strabon* y *Stardog*.

En este trabajo se utilizarán y compararán dos almacenes de tripletas con soporte a GeoSPARQL como son Virtuoso y Fuseki:

- **Virtuoso**<sup>10</sup> es el almacén de tripletas utilizado en el *dataset Cross-forest* en su versión libre (versión 07.20.3230), aunque también tiene una versión propietaria. Virtuoso es una base de datos híbrida ya que combina un modelo de datos orientado a grafos y una base de datos relacional. Esto significa que se pueden almacenar y consultar tanto datos estructurados como no estructurados. Virtuoso es una herramienta madura y estable, con una larga historia y amplio uso en proyectos de datos enlazados. Este almacén de tripletas es especialmente adecuado para proyectos que requieren rapidez en las consultas, alta escalabilidad y capacidades de replicación y sincronización de datos.
- **Fuseki**<sup>11</sup> es un almacén de tripletas de código abierto desarrollada por Apache Jena en el que se puede integrar la implementación de GeoSPARQL<sup>12</sup>. Fuseki está enfocado exclusivamente en el modelo de datos orientado a grafos, es una herramienta más reciente que Virtuoso basada en Java y se ha ganado una gran popularidad en la comunidad de datos enlazados debido a su facilidad de uso y su enfoque en la creación de aplicaciones Web. Según el artículo de Jovanovik et al. (2021), Fuseki es el *triplestore* que implementa mejor el estándar de GeoSPARQL de los comparados.

Tanto Virtuoso como Fuseki tienen soporte para GeoSPARQL, como se ha mencionado anteriormente, sin embargo la implementación se ha llevado a cabo de manera diferente en ambas plataformas. El grado de soporte que estos dos *triplestores* dan a GeoSPARQL es importante para este TFM, ya que esto determina si es mejor emplear uno u otro para el almacén de bases de datos geospaciales y el uso de determinadas funciones espaciales. La implementación de GeoSPARQL en Virtuoso se aleja más de los estándares

---

<sup>10</sup><https://virtuoso.openlinksw.com/>

<sup>11</sup><https://jena.apache.org/documentation/fuseki2/>

<sup>12</sup><https://jena.apache.org/documentation/geosparql/>

propuestos por la OGC. Virtuoso solo soporta el formato WKT para el almacén de datos geospaciales, mientras que Fuseki utiliza los formatos WKT y GML.

### 3.2. Descripción de los datos fuente

En este trabajo se pretende integrar dos bases de datos espaciales: el *dataset Cross-Forest* que contiene las parcelas del IFN3 y sus coordenadas en formato RDF (LOD) y la base de datos del IGN que contiene los recintos municipales españoles actuales en formato *Shapefile*. En las siguientes secciones se explica con mayor detalle ambas bases de datos.

#### 3.2.1. Parcelas del IFN3

El IFN3 se encuentra publicado para su descarga gratuita en 100 archivos de bases de datos *Microsoft Access* (.mdb) agrupados por provincias en la página del MITECO<sup>13</sup>, donde también se encuentra un archivo pdf (Documentador Base de Datos de Campo<sup>14</sup> y Documentador de Base de Datos SIG<sup>15</sup>) que describe las Tablas que componen el IFN3 y sus atributos.

En el proyecto para la creación de la base de datos semántica *Cross-Forest* el grupo de investigación en el que se desarrolla este TFM convirtió la base de datos del IFN3 del MITECO a LOD definiendo un conjunto de ontologías o vocabularios (posiciones, medidas, especies...) para integrar y crear las relaciones entre los datos que a su vez se enlazaron con otras ontologías (NCBI, DBpedia, Wikidata, Cross-Nature). Además, se relacionó con otras bases de datos como el Mapa Forestal Español a escala 1:50.000 (MFE50) en el que se basa la distribución de las parcelas del IFN3. En el punto SPARQL *Cross-Forest*<sup>16</sup> se pueden consultar todas las parcelas con sus características y árboles asociados y descargar los datos de todas las parcelas de España en diferentes formatos (JSON, CSV, Turtle, N-Triple, XML, RDF/XML...) lo que facilita mucho su uso para diversos propósitos.

Las parcelas del IFN3 que se van a utilizar en este trabajo fueron primeramente definidas por el equipo del IFN del MITECO en gabinete, partiendo como referencia de, 1) la malla UTM (*Universal Transversal Mercator*) de 1km x 1km, 2) los usos del suelo definidos en el MFE50 y 3) las parcelas

<sup>13</sup><https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/informacion-disponible/ifn3.aspx>

<sup>14</sup>[https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/documentador\\_bdcampo\\_ifn3\\_tcm30-282240.pdf](https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/documentador_bdcampo_ifn3_tcm30-282240.pdf)

<sup>15</sup>[https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/documentador\\_bdsig\\_ifn3\\_tcm30-293905.pdf](https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/documentador_bdsig_ifn3_tcm30-293905.pdf)

<sup>16</sup><https://crossforest.gsic.uva.es/sparql>

que ya se habían apeado en la segunda edición del IFN (IFN2) entre 1986 y 2006. En gabinete se describieron 96 660 parcelas totales (que se encuentran en la Tabla `PCDatosMap`) identificadas con el código de la provincia en la que se encuentra la parcela, el número de estadillo de cada una y su clase. Estas pueden ser A (parcelas que ya se habían muestreado en el IFN2) o N (parcelas nuevas que se apearon por primera vez en el IFN3). Todas estas parcelas se asociaron a unas coordenadas del centro de la parcela en el sistema de referencia de coordenadas UTM ED50 dentro de la malla de 1km x 1km, lo que permite su posterior localización en campo (Figura 3.8).

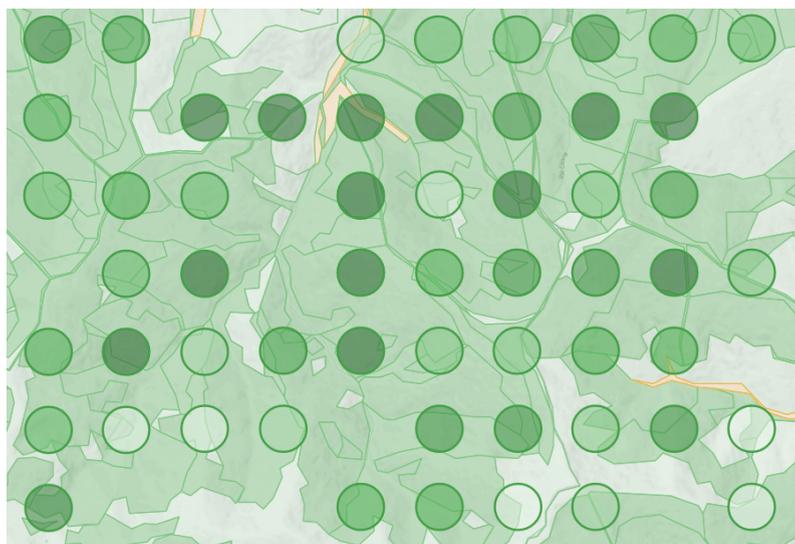


Figura 3.8: Distribución de las parcelas del IFN3 siguiendo una malla de 1km x 1km en las zonas donde el uso del suelo es arbolado según el MFE50. Fuente: <https://forestexplorer.gsic.uva.es/explorer/>.

Una vez realizado el trabajo de campo por el equipo del MITECO y Tragsatec, en el IFN3 se encuentran descritas 99 046 parcelas (en la Tabla `PCParcelas`) identificadas mediante el código de provincia, estadillo y clase definido en gabinete, al que se le añade una subclase en campo para identificar cambios en las parcelas de la clase A (parcelas inicialmente apeadas en el IFN2). Por este motivo hay parcelas en la Tabla `PCDatosMap` que en la Tabla `PCParcelas` se asocian a una o más parcelas. Por ejemplo, la parcela 01-0317-A en `PCDatosMap` se divide en 01-0317-A-3E y 01-0317-A-3C en `PCParcelas`, donde la subclase (3E, 3C) determina las condiciones y cambios de la parcela en campo.

Tras identificar discordancias entre las Tablas, Tragsatec, la empresa encargada por el MITECO de la realización del IFN, nos aclaró que además de las parcelas descritas en gabinete, en campo se apearon 545 parcelas de refuerzo (clase R), denominadas parcelas satélite de una parcela planetaria o

principal y se desestimaron otras por diversos motivos. La Tabla 3.1 muestra de forma resumida las parcelas que contiene finalmente el IFN3 y sus características.

Tabla 3.1: Discrepancias entre las parcelas de la Tabla PCDatosMap y PCParcelas en el IFN3.

	<b>PCDatosMap (gabinete)</b>	<b>PCParcelas (campo)</b>
<b>Identificador parcelas</b>	Prov - Esta - Cla	Prov - Esta - Cla - Subcla
<b>Nº de parcelas totales</b>	96 660	99 046
<b>Parcelas que están en ambas Tablas</b>	92 569	98 497
<b>Parcelas que no están en ambas Tablas</b>	4 091	549
	1 128 son tipo A (IFN2)	545 son tipo R (refuerzo)
	2 959 son tipo N (IFN3)	
<b>Desglose parcelas que no coinciden</b>	33-0999-N	33-9999-N-N
	17-0458-N	17-0458-A-4
	29-2293-N	29-2293-A-4
	29-2293-N	29-2321-A-4
<b>Aclaraciones Tragsatec</b>	4 087 parcelas posiblemente desestimadas en campo. 17-0458-N es 17-0458-A. 29-2293-N y 29-2293-N se clasificaron como A4 en campo por su condición de inaccesibles.	33-9999-N-N es 33-0999-N

En el *dataset Cross-Forest* las 99 046 parcelas del IFN3 (identificadas por provincia-estadillo-clase-subclase) se encuentran asociadas a las coordenadas UTM ED50 (*European Datum 1950*), coordenadas originales del IFN3 y también han sido transformadas a las coordenadas en el sistema geodésico de referencia de coordenadas geográficas mundial o *World Geodetic System 1984* (WGS84).

### 3.2.2. Recintos municipales españoles

Las delimitaciones territoriales de los municipios son fundamentales para la organización administrativa y política de España. A partir de los límites municipales se delimitan el resto de unidades administrativas por agregación (provincias, comunidades autónomas y el propio estado) (BOE, 1986). Con la llegada de la cartografía digital, en 2009 el IGN comenzó a digitalizar todas las líneas límite que antes se encontraban definidas en en las Actas de deslinde municipales según los títulos jurídicos inscritos en el Registro Central de Cartografía (RCC)<sup>17</sup>.

<sup>17</sup>Real Decreto 1545/2007, de 23 de noviembre, por el que se regula el Sistema Cartográfico Nacional [https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2007-20556](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2007-20556)

Para definir estos límites se utilizan los códigos asignados por el Instituto Nacional de Estadística (INE)<sup>18</sup>, que son una serie de 5 números que forman un código único que identifica cada municipio. Los primeros dos dígitos pertenecen al código de la provincia y los 3 siguientes al código del municipio dentro de la misma. Además, el INE añade un sexto dígito de control a cada código de municipio que es asignado mediante una regla de cálculo para detectar errores. Estos códigos INE se relacionan con la denominación del municipio correspondiente según Figuran inscritos en el Registro de Entidades Locales (REL)<sup>19</sup>. Todos los años se actualiza la información de los municipios y se publican las modificaciones en su denominación y/o en su código INE. En el Anexo B se muestra una Tabla con las modificaciones de los códigos INE de los municipios desde la edición de 2007 hasta 2023 que afectan a las parcelas del IFN3 y que se tendrán en cuenta a la hora de realizar la asignación de las parcelas a los municipios.

En la actualidad se encuentran publicados varios archivos de los recintos y límites municipales de España en formato *Shapefile* en el centro de descargas del IGN<sup>20</sup>. En este trabajo se han utilizado dos ficheros, uno relativo a las unidades administrativas de los recintos municipales de la Península y Baleares y otro que contiene los recintos municipales de las Islas Canarias (Figura 3.9). Los sistemas de referencia en los que están definidos son ETRS89 (*European Terrestrial Reference System 1989*) para la Península y Baleares y REGCAN95 (Red Geodésica Nacional por Técnicas Espaciales Canarias 1995) en las Islas Canarias, que es casi idéntico a WGS84 (*World Geodetic System 1984*). Las diferencias entre ETRS89 y WGS84 son despreciables para la mayoría de SIG, aunque ETRS89 es un sistema de referencia de coordenadas definido para Europa y por tanto se considera más preciso en esta zona, mientras que WGS84 es global.

Estos ficheros contienen asociados a la geometría de los municipios, además de otros atributos que no hemos utilizado en este trabajo, el nombre oficial de los municipios en la actualidad y el código nacional (CCAAPPMMMM, donde CC es el código del país, AA, es el código de la comunidad autónoma, PP es el código de la provincia, MMMMM es el código del municipio) del que hemos extraído los últimos 5 dígitos relativos al código del municipio.

---

<sup>18</sup>[https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica\\_C&cid=1254736177031&menu=ultiDatos&idp=1254734710990](https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177031&menu=ultiDatos&idp=1254734710990)

<sup>19</sup><https://ssWeb.seap.minhap.es/REL/>

<sup>20</sup><https://centrodedescargas.cnig.es/CentroDescargas/index.jsp>

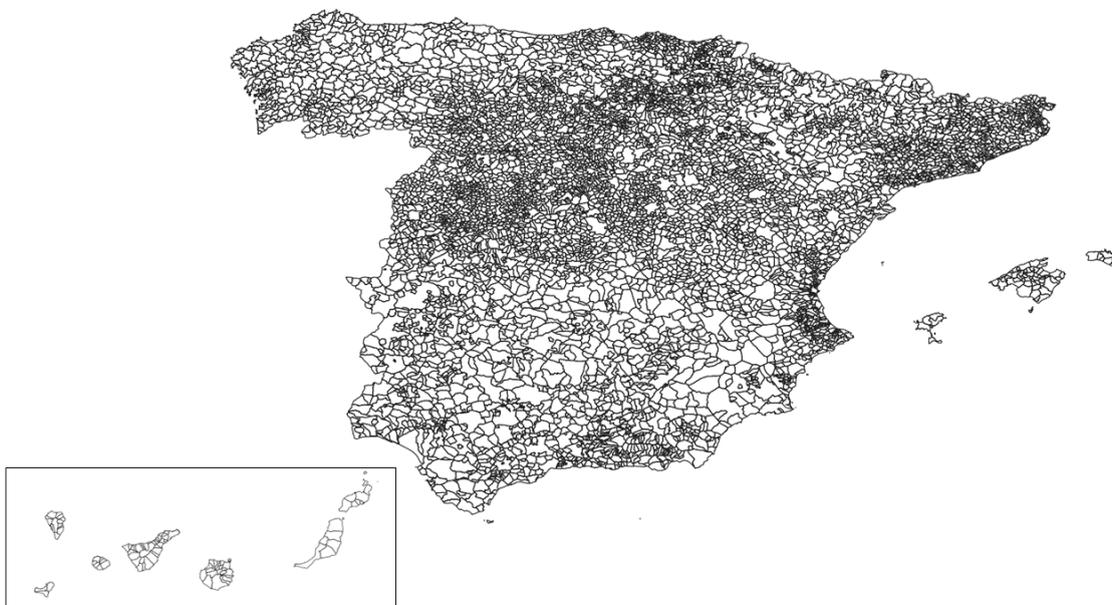


Figura 3.9: Recintos municipales de España de la base de datos del IGN de 2023.

En los archivos *shapefile* de los recintos municipales publicados por el IGN hay un total de 8 217 geometrías (polígonos y multipolígonos) que cubren todo el territorio español. Esas 8 217 geometrías comprenden:

- 8 131 municipios oficiales a fecha de este trabajo con código INE definido e inscritos en el registro de Entidades Locales.
- 81 territorios cuya jurisdicción es compartida por varios términos municipales. De forma genérica se llaman comunidades jurisdiccionales. En la base de datos del IGN presentan un código INE ficticio de 5 dígitos que empieza por 53.
- 5 territorios que jurisdiccionalmente no pertenecen a ningún municipio sino que dependen directamente del Estado. Estos son las posesiones españolas en África: Isla del Perejil, Islas Alhucemas, Islas Chafarinas y Peñón de Vélez de la Gomera. Presentan un código INE ficticio de 5 dígitos que empieza por 54. Estos territorios no se tendrán en cuenta en este trabajo ya que no existen parcelas del IFN en los mismos.

En el *dataset Cross-Forest* se encuentran definidos en RDF los 8 131 municipios oficiales y sus propiedades: provincia al que pertenece cada municipio, nombre, código INE de municipio, dígito de control... (trabajo llevado a cabo en el proyecto Erasmus+ *VirtualForests*<sup>21</sup>) pero no los límites

<sup>21</sup><https://virtualforests.eu/>

de los términos municipales (geometrías), ni las comunidades jurisdiccionales o los territorios que dependen directamente del estado. En la Figura 3.10 se presenta la consulta SPARQL que muestra las características asociadas al municipio de Palencia de código INE 34120 en el *dataset Cross-Forest*. En el resultado de la consulta, que se muestra en la Tabla 3.2, se puede ver que el municipio carece de geometría asociada, necesaria para realizar este trabajo.

```
SELECT DISTINCT *
WHERE {
  municipality:34120 ?p ?o.
}
```

Figura 3.10: Consulta SPARQL al *dataset Cross-Forest* de las propiedades y objetos asociados al municipio de Palencia (las IRI de los prefijos utilizados se encuentran en el Anexo A)

Tabla 3.2: Propiedades y valores asociadas al municipio de Palencia en el *dataset Cross-Forest*.

p	o
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#Municipio">http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#Municipio</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://www.w3.org/2002/07/owl#NamedIndividual">http://www.w3.org/2002/07/owl#NamedIndividual</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/Municipality">https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/Municipality</a>
<a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a>	"Palencia"@es
<a href="http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#provincia">http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#provincia</a>	<a href="https://datos.iepnb.es/recurso/sector-publico/territorio/Provincia/34">https://datos.iepnb.es/recurso/sector-publico/territorio/Provincia/34</a>
<a href="https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasCheckDigit">https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasCheckDigit</a>	2
<a href="https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasMunicipalityCode">https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasMunicipalityCode</a>	"120"

En el caso de las comunidades jurisdiccionales también será necesario crear las tripletas que describan las características de las mismas ya que no se encuentran incluidas actualmente en el *dataset Cross-Forest*. Para ello se utilizarán los atributos asociados a las geometrías: provincia, nombre de la entidad, código INE de municipio ficticio y dígito de control. Este último se generará a partir de un *script* desarrollado mediante el software R (Anexo C) que aplica las reglas de transformación que sigue el INE para generar los dígitos de control de los municipios según lo descrito por Garoia Rubio et al. (1982).

### 3.3. Procesado geoespacial

En esta Sección se van a tratar, por un lado, la conversión a RDF de las geometrías de los recintos municipales españoles, y por otro lado, la integración y preparación de los datos en los entornos de pruebas.

### 3.3.1. Conversión de las geometrías de los municipios a RDF

Con el propósito de convertir las geometrías de los límites municipales de *Shapefile*, formato en el que se encuentran actualmente en la base de datos del IGN, a *Turtle*, formato para serializar grafos RDF, se llevará a cabo el siguiente procedimiento, que se esquematiza en la Figura 3.11. Cabe mencionar, que las propiedades asociadas a los municipios (toponimia, código INE...) ya se encuentran en el *dataset Cross-Forest* como se ha explicado en el Capítulo 1 y la Sección 3.2.2.

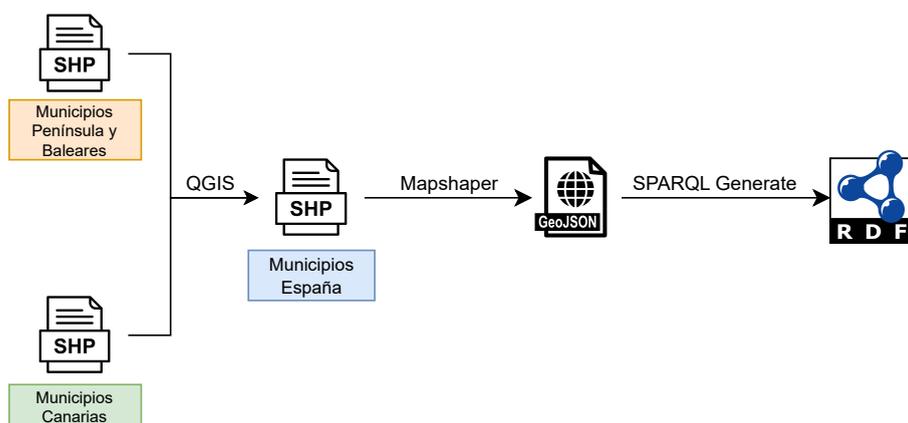


Figura 3.11: Diagrama de flujo del proceso de conversión de los recintos municipales españoles de *Shapefile* a RDF

1. **Unión o *merge* de los dos archivos *Shapefile* de municipios** utilizando QGIS<sup>22</sup>: por un lado, el que contiene los límites municipales de la Península y las Islas Baleares, y por otro, lado el que contiene lo límites municipales de las Islas Canarias para obtener un *Shapefile* con sistema de referencia de coordenadas WGS84. Aunque los dos archivos se encuentren en sistemas de referencia de coordenadas diferentes (ETRS89 y REGCAN95) estos son compatibles por lo que no es necesario realizar ninguna transformación de sistemas de referencia.
2. **Creación de una nueva columna (**IDmuni**)** en la Tabla de atributos asociada a las geometrías del *Shapefile* de los recintos municipales mediante la extracción de los 5 últimos dígitos del código nacional de los municipios. Para ello se utilizará la función `substr` en la calculadora de campos de la Tabla de atributos en QGIS.
3. **Conversión del *Shapefile* de los recintos municipales a GeoJSON** mediante *Mapshaper*<sup>23</sup>. El formato GeoJSON facilita la conversión de

<sup>22</sup><https://qgis.org/es/site/>

<sup>23</sup><https://mapshaper.org/>

las geometrías a RDF, ya que SPARQL-Generate puede leer este formato pero no puede leer *Shapefile*.

4. **Conversión de los GeoJSON de los municipios a RDF** utilizando un *script* de shell<sup>24</sup> desarrollado durante el proyecto *Cross-Forest* para ejecutar SPARQL-Generate y generar un archivo en formato *Turtle* (extensión *.ttl*) con las geometrías de los municipios asociadas a su código INE en RDF.

### 3.3.2. Integración de datos

En esta Sección se tratará la metodología para realizar la asignación de las parcelas del IFN3 a los municipios en los que se encuentran utilizando dos *triplestores* (Virtuoso y Fuseki) y un SIG (QGIS). Todo ello se llevará a cabo en el Sistema de Referencia de Coordenadas WGS84, en el que se encuentran tanto los recintos municipales como las posiciones de las parcelas del IFN3.

#### 3.3.2.1. Integración mediante SPARQL

Para la integración de los recintos municipales en RDF con las parcelas del IFN3 primero se descargarán estos datos en formato *Turtle* del punto SPARQL del *dataset Cross-Forest*. La descarga se llevará a cabo mediante SPARQL CONSTRUCT, una cláusula de consulta en SPARQL que se utiliza para construir grafos RDF, como ya se ha detallado en la Sección 3.1.2. La integración de los datos de los recintos municipales en RDF con las parcelas del IFN3 se realizará primero en dos entornos de pruebas, configurados con dos almacenes de tripletas que dan soporte a GeoSPARQL, Virtuoso y Fuseki. De esta manera se podrá comparar ambas implementaciones, por un lado la de Virtuoso, el almacén de tripletas utilizado en el *dataset Cross-forest*. Por otro lado la de Fuseki, uno de los almacenes de tripletas de libre uso que según Jovanovik et al. (2021) presenta la mejor implementación de GeoSPARQL al compararlo con otros almacenes de tripletas.

Para la configuración del entorno de pruebas con el almacén de tripletas Virtuoso se utilizará la imagen de Docker<sup>25</sup> *OpenLink Virtuoso Open Source Edition 7.2*<sup>26</sup> que simplifica su instalación y facilita su despliegue. En el caso de Fuseki, es necesario descargar la última versión de su página de descargas (Versión 4.7.0)<sup>27</sup>, y agregar la implementación de GeoSPARQL para Fuseki descargandola del repositorio de maven para la versión 4.7.0<sup>28</sup>,

<sup>24</sup><https://github.com/Cross-Forest/scripts/blob/master/sparql-generate.sh>

<sup>25</sup><https://www.docker.com/>

<sup>26</sup><https://hub.docker.com/r/openlink/virtuoso-opensource-7>

<sup>27</sup><https://jena.apache.org/download/>

<sup>28</sup><https://mvnrepository.com/artifact/org.apache.jena/jena-geosparql>

ya que Fuseki con GeoSPARQL integrado no se encuentra en Docker, como sería lo deseable.

Una vez creados los entornos de pruebas, se cargarán 1) los archivos en formato *Turtle* de las parcelas con sus coordenadas (latitud y longitud), 2) las propiedades de los municipios y 3) las geometrías de los recintos municipales transformados a *Turtle* (proceso descrito en la Sección 3.3.1). Una vez que tanto los datos de las posiciones de las parcelas del IFN3 como los municipios españoles estén en RDF en los entornos de pruebas es posible realizar la asignación de las parcelas a los municipios mediante una consulta GeoSPARQL.

Para realizar las asignaciones se creará una consulta adaptada a las funciones de GeoSPARQL implementadas en los diferentes almacenes de tripletas que se van a utilizar. Por una lado, Virtuoso utiliza la función espacial `st_within`<sup>29</sup> para determinar si una geometría se encuentra dentro de otra geometría. La función toma dos argumentos geométricos siguiendo la sintaxis básica: `st_within (geometría1, geometría2)` y devuelve un valor booleano que indica si la primera geometría está completamente dentro de la segunda geometría. Las geometrías pueden ser cualquier tipo de objeto geométrico soportado por GeoSPARQL, como puntos, líneas o polígonos. Virtuoso, usa una extensión propietaria de GeoSPARQL (prefijo `bif`) lo que quiere decir que no se puede usar la misma consulta en otros almacenes de tripletas, como Fuseki, si no que requiere adaptaciones de las funciones `bif`.

Por otro lado, la consulta GeoSPARQL en Fuseki difiere un poco de la que se realiza en Virtuoso. Para asignar las parcelas a sus municipios se utilizará la función `spatialF:convertLatLon`. `spatialF`, es una extensión propietaria de Fuseki para convertir las coordenadas de latitud y longitud en un punto en el espacio. También se empleará la función `geof:sfWithin`, función de GeoSPARQL definida por la OGC, para determinar si ese punto se encuentra dentro del polígono del municipio, de manera similar a la función `st_within` que emplea Virtuoso.

### 3.3.2.2. Integración mediante un SIG

Es muy habitual la utilización de SIG en el tratamiento y gestión de datos geospaciales. Por este motivo se pretende realizar la asignación de las parcelas a los recintos municipales utilizando un SIG de libre uso como QGIS y compararla con las asignaciones realizadas con SPARQL.

<sup>29</sup>[https://docs.openlinksw.com/virtuoso/fn\\_st\\_within/](https://docs.openlinksw.com/virtuoso/fn_st_within/)

Para realizar la asignación de las parcelas a los municipios en QGIS ya disponemos de un archivo con los recintos municipales de la Península y las Islas en el Sistema de Coordenadas WGS84 en formato *Shapefile*, resultado de los procesos realizados en la Sección 3.3.1. Por lo que será necesario tener otro *Shapefile* con las posiciones de todas las parcelas del IFN3. Este *Shapefile* se generará a partir de las parcelas y sus coordenadas asociadas transformadas a WGS84 en el *dataset Cross-Forest*. Para la descarga de los datos de las parcelas se llevará a cabo una consulta al punto SPARQL de *Cross-Forest* especificando que los resultados generados se descarguen en formato CSV. Una vez descargadas en CSV los identificadores de las parcelas con sus coordenadas asociadas (latitud y longitud) se importarán en QGIS como *Shapefile*.

La asignación de las parcelas a sus municipios se llevará a cabo en QGIS mediante la herramienta que permite unir atributos por localización (*Join attributes by location*) de capas vectoriales. Esta herramienta posibilita la utilización de la función *Within* que tiene en cuenta la ubicación espacial de dos capas vectoriales y determina qué geometría esta completamente dentro de otra geometría, de manera similar a lo que hace la función `geof:sfWithin` de GeoSPARQL. Previamente a la realización de la asignación, se creará en QGIS un índice espacial tanto del *Shapefile* de los recintos municipales como de las parcelas, ya que este archivo no se crea por defecto dentro del *Shapefile* y mejorara significativamente la velocidad y la eficiencia del procesamiento de datos espaciales.

Finalmente, se transformará el *Shapefile* resultante con la asignación de las parcelas a municipios a GeoJSON y posteriormente a RDF mediante un *script* de SPARQL-Generate. Esto permitirá realizar las comparaciones de los resultados en RDF de las asignaciones llevadas a cabo por los tres procedimientos empleados.

### **3.4. Comparación y validación de los resultados**

Una vez obtenidos los resultados de la asignación de las parcelas con el código INE del municipio al que pertenecen utilizando la implementación de GeoSPARQL de Virtuoso y Fuseki y la asignación realizada mediante QGIS, se llevará a cabo una validación cruzada entre las tres asignaciones utilizando funciones para la gestión de datos en la versión 4.2.2 del software R (R Core Team, 2017). Para la comparación se exportarán los resultados del código de las parcelas asociado al código INE del municipio asignado

en formato CSV de las tres asignaciones. Posteriormente, se cargarán en RStudio y se unirán las Tablas por el código de la parcela con la función `merge` para luego realizar un `subset` de las parcelas cuyo código INE coincida.

Además, también se compararán los resultados obtenidos en este trabajo con las asignaciones de los códigos INE a sus parcelas correspondientes llevadas a cabo en la base de datos del IFN3 por el equipo del MITECO. Para ello se tendrá en cuenta las modificaciones de los códigos INE que han sufrido algunos municipios desde la publicación del inventario (2007) hasta la actualidad según lo publicado en la página Web del INE de las relaciones de municipios y sus códigos<sup>30</sup> (Anexo B).

### 3.5. Aplicación de los resultados

Con la información de la asignación de las parcelas a los municipios a los que pertenecen se podrán extraer datos y estadísticas básicas por municipio. Para poner en valor e ilustrar los resultados obtenidos y la potencialidad de los LOD, se asociarán las parcelas asignadas a sus municipios con variables a nivel de parcela calculadas en el IFN3, como son el número de pies/ha (N) y el área basimétrica (G) de cada especie.

Estas variables relativas son ampliamente utilizadas en el ámbito forestal para el estudio y comparación de parcelas. Por lo que a modo de ejemplo en este trabajo se realizará un mapa mediante QGIS de especies dominantes según el área basimétrica, expresada en  $m^2/ha$ , y el número de pies por hectárea de cada municipio español. Para ello, se integrarán los datos de estas variables forestales, que se encuentran asociadas a las parcelas del IFN3 en el *dataset Cross-Forest*, en el entorno de pruebas donde se habrá cargado la asignación de las parcelas al códigos INE del municipio al que pertenecen una vez validada. Con esto, se desarrollarán consultas SPARQL para calcular la media de N y G de cada especie por municipio y se empleará el valor máximo de la media para representar en dos mapas la especie dominante de cada municipio de acuerdo a estas variables.

Además, se evaluará el caso concreto de las especies dominantes de los municipios que se encuentran dentro de la influencia la Red de Parques Nacionales Española, compuesta por 16 espacios naturales protegidos para la conservación de su flora y fauna, al considerarse especialmente relevantes.

---

<sup>30</sup>[https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica\\_C&cid=1254736177031&menu=ultiDatos&idp=1254734710990](https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177031&menu=ultiDatos&idp=1254734710990)

La gestión de este tipo de espacios naturales tiene gran importancia en España, ya que es uno de los países de la Unión Europea que contribuye más a su biodiversidad (Ayanz, 2014). Además, cuentan con planes específicos para su gestión, puestos en marcha por los organismos autonómicos que en colaboración con las entidades locales, asociaciones, instituciones y organizaciones relacionadas llevan a cabo una gestión adecuada para la protección de estas áreas (MITECO, s.f.-a).

Para delimitar los municipios que se encuentran dentro de los Parques Nacionales se usarán los *Shapefile* del área de influencia socioeconómica (AIS) publicado por el MITECO<sup>31</sup>. Las AIS se definen como el territorio constituido por los términos municipales que aportan terreno a los Parques Nacionales, así como, excepcionalmente, por otros municipios directamente relacionados, en los que las administraciones públicas llevarán a cabo políticas activas para su gestión y desarrollo (BOE, 2014).

### **3.6. Publicación de los resultados como LOD**

Una vez obtenidas tanto las geometrías de los municipios españoles, como la asignación validada de las parcelas del IFN3 al código INE del municipio al que pertenecen en formato *Turtle*, estos dos archivos se cargarán en el *dataset Cross-forest* integrándose así con el resto de información con la que ya cuenta. De esta forma se podrán realizar consultas y descargar los datos en diferentes formatos para una gran variedad de usos. Además, los archivos utilizados y generados, así como los *scripts* desarrollados en este trabajo, se publicarán de manera abierta en un repositorio GitHub creado para este propósito.

### **3.7. Diagrama de flujo resumen de todo el proceso**

En la Figura 3.12 se presenta un diagrama a modo resumen de los procesos llevados a cabo para cumplir con los objetivos planteados.

---

<sup>31</sup><https://www.miteco.gob.es/es/parques-nacionales-oapn/red-parques-nacionales/sig/ais.aspx>

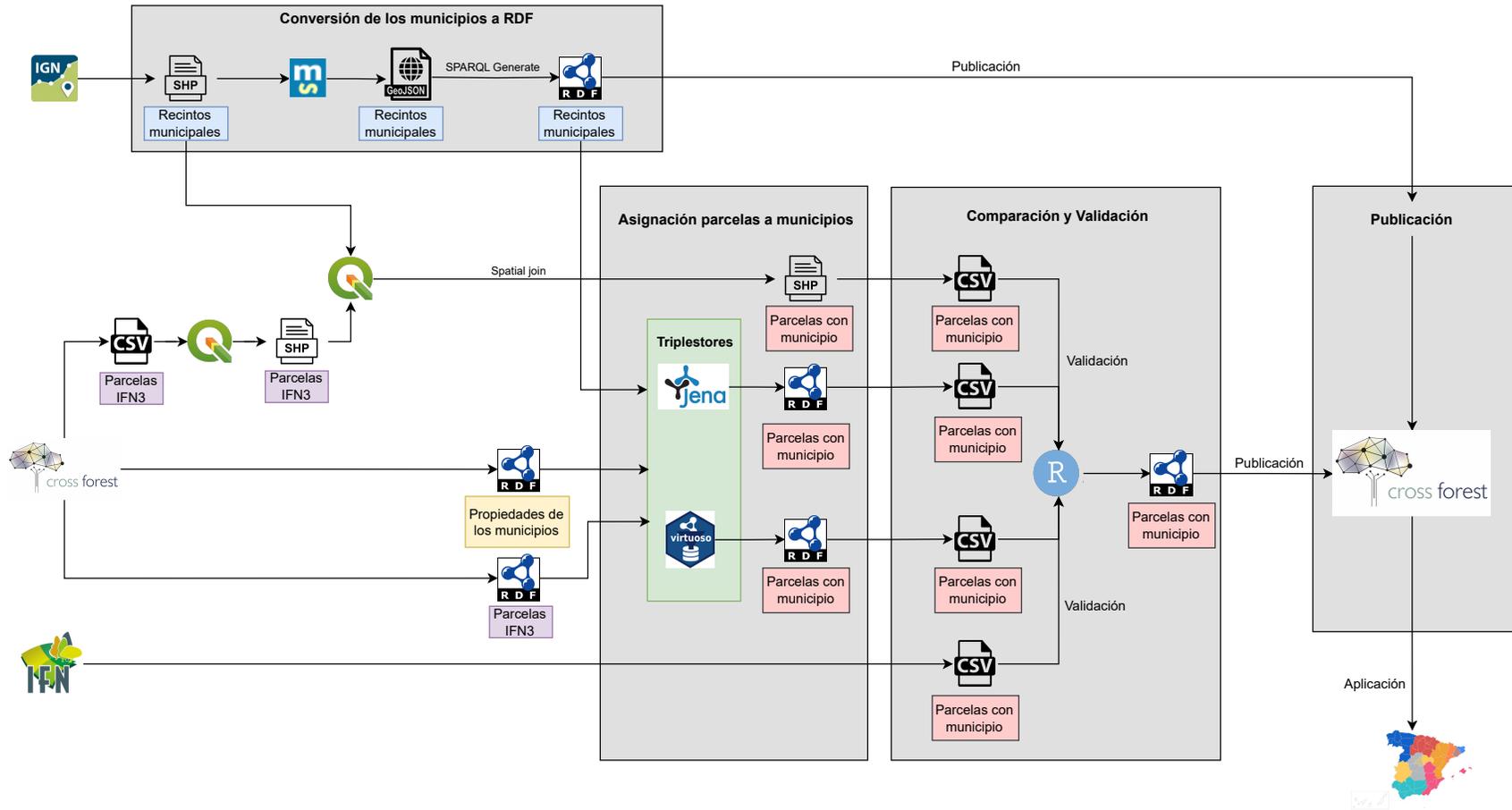


Figura 3.12: Diagrama de flujo donde se detalla todo el proceso llevado a cabo en este trabajo.

## Capítulo 4

# RESULTADOS

En este Capítulo se van a tratar los resultados obtenidos, desde la conversión de las geometrías de los municipios a RDF, la asignación de las parcelas a sus respectivos términos municipales y su validación, hasta la publicación en abierto de los datos generados. Además, se presentarán mapas que muestren la utilidad de la integración de la base de datos espacial de los recintos municipales del Instituto Geográfico Nacional con la base de datos semántica *Cross-Forest* en el ámbito forestal.

### 4.1. Conversión de los recintos municipales a RDF

Como resultado de la transformación de las geometrías de los municipios españoles de formato *Shapefile* a GeoJSON mediante *Mapshaper* (Sección 3.3.1), se ha generado un archivo GeoJSON de polígonos y multipolígonos representados por sus coordenadas en WGS84 (latitud y longitud) y asociados a una serie de propiedades que se corresponden con los atributos de cada *feature* en el archivo *Shapefile*. En la Figura 4.1 se presenta el código GeoJSON generado para el municipio de Palencia, donde solo se muestran algunas coordenadas de los vértices del polígono que forman los límites del mismo a modo de ejemplo. En la Figura 4.1 también se puede ver la propiedad `IDmuni` extraída como se explica en la Sección 3.3.1 para identificar el municipio por el código INE de 5 dígitos.

Una vez convertidas las geometrías con sus propiedades a GeoJSON se ha creado un *script* SPARQL-Generate (Anexo D) que utiliza diversas ontologías y prefijos RDF para generar una lista de las geometrías de los municipios asociadas a su código INE en formato RDF (*Turtle*). Para automatizar las ejecuciones de SPARQL-Generate se ha utilizado un *script* de shell, desarrollado durante el proyecto *Cross-Forest*<sup>1</sup>. En la Figura 4.2 se puede observar

---

<sup>1</sup><https://github.com/Cross-Forest/scripts/blob/master/sparql-generate.sh>

```

{
  "type": "Feature", // define este objeto como una "entidad" geográfica
  "geometry": { // descripción la geometría de la entidad
    "type": "Polygon", // tipo de geometría (en este caso, un polígono)
    "coordinates": [ // describe las coordenadas (lat, lon) del polígono
      [ [-4.6532803540000005,41.949459208], [-4.653179586,41.948361883000004],
        [-4.653578359,41.946987174], [-4.653606339,41.946203217000004], ... ]
    ]
  },
  "properties": { //contiene información de la Tabla atributos del Shapefile
    "INSPIREID": "ES.IGN.BDDAE.34073434120",
    "COUNTRY": "ES",
    "NATLEV": "https://inspire.ec.europa.eu/codelist/
AdministrativeHierarchyLevel/4thOrder",
    "NATLEVNAME": "Municipio",
    "NATCODE": "34073434120",
    "NAMEUNIT": "Palencia",
    "CODNUT1": "ES4",
    "CODNUT2": "ES41",
    "CODNUT3": "ES414",
    "IDmuni": "34120"
  }
}

```

Figura 4.1: Ejemplo de la geometría del municipio de Palencia en formato GeoJSON.

un ejemplo del formato *Turtle* generado. El municipio de Palencia (código INE 34120) se identifica en el *dataset Cross-Forest* por *municipality:34120*, que ahora tiene un polígono asociado, *polygon:34120-432* (identificador formado por el código INE del municipio y el código del sistema de referencia de coordenadas WGS84<sup>2</sup> en el que están las coordenadas del polígono). Este polígono correspondiente a la superficie del municipio de Palencia, está a su vez asociado, con los vértices que forman su geometría en formato WKT.

```

@prefix ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/> .
@prefix epsg: <http://epsg.w3id.org/ontology/> .
@prefix wkt: <http://www.opengis.net/ont/geosparql#> .
@prefix spo: <http://crossforest.eu/position/ontology/> .
@prefix polygon: <https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/
mfe/polygon/> .
@prefix crs: <http://crossforest.eu/epsg/data/crs/> .
@prefix municipality: <https://datos.iepnb.es/recurso/sector-publico/medio-
ambiente/ifn/municipality>

municipality:34120 spo:hasPolygon polygon:34120-4326.
polygon:34120-4326 epsg:hasCoordinateReferenceSystem crs:4326 ;
    wkt:asWKT "POLYGON((-4.6532803540000005,41.949459208,,-4.653179586,
    41.948361883000004,,-4.653578359,41.946987174,,-4.653606339,
    41.946203217000004,,-...))"^^wkt:wktLiteral .

```

Figura 4.2: Ejemplo de formato *Turtle* del municipio de Palencia asociado a su geometría.

<sup>2</sup><https://epsg.io/4326>

## 4.2. Asignación de las parcelas del IFN3 a los recintos municipales

La configuración mediante línea de comandos de los entornos de pruebas con los almacenes de tripletas Virtuoso y Fuseki se muestran en el Anexo E y Anexo F, respectivamente. Para la descarga de los datos en formato *Turtle* del *dataset Cross-Forest* se desarrollaron las siguientes consultas SPARQL CONSTRUCT, con el fin de descargar las parcelas y sus coordenadas asociadas en WGS84 (Figura 4.3) y las propiedades de los municipios (Figura 4.4) para posteriormente cargar estos datos junto con las geometrías de los municipios generadas previamente en ambos almacenes de tripletas, según lo descrito en el Anexo E y Anexo F.

```

CONSTRUCT {
  ?plot a ifn:Plot;
    spo:hasPosition ?pos.
  ?pos ?p ?o.
}

WHERE {
  ?plot a ifn:Plot.
  ?plot spo:hasPosition ?pos .
  ?pos ?p ?o.
  FILTER EXISTS {?pos spo:hasCoordinateReferenceSystem crs:4326. }
}

```

Figura 4.3: Consulta SPARQL CONSTRUCT para descargar las parcelas del IFN3 y sus coordenadas asociadas del *dataset Cross-Forest* en formato *Turtle* (las IRI de los prefijos utilizados se encuentran en el Anexo A).

```

CONSTRUCT {?m ?p ?o}

WHERE {
  ?m a muni:Municipio.
  ?m ?p ?o.
}

```

Figura 4.4: Consulta SPARQL CONSTRUCT para descargar las características asociadas a los municipios en formato *Turtle* (las IRI de los prefijos utilizados se encuentran en el Anexo A).

Como resultado de cargar también el archivo *Turtle* con las geometrías de los municipios al entorno de pruebas junto con los datos de las propiedades de los municipios, los municipios cuentan ahora con una geometría asociada con la propiedad `spo:hasPolygon` a la IRI del identificador del polígono. La IRI del polígono a su vez esta asociada a las coordenadas en formato WKT de los vértices de la geometría del municipio como se mostraba en la Figura 4.2. Esto se puede ver en la Tabla 4.1 que indica el resultado de la consulta

anteriormente descrita en la Figura 3.10 donde se piden las propiedades del municipio de Palencia.

Tabla 4.1: Ejemplo de integración del polígono de los límites del municipio de Palencia con sus propiedades.

p	o
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#Municipio">http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#Municipio</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://www.w3.org/2002/07/owl#NamedIndividual">http://www.w3.org/2002/07/owl#NamedIndividual</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/Municipality">https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/Municipality</a>
<a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a>	"Palencia"@es
<a href="http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#provincia">http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#provincia</a>	<a href="https://datos.iepnb.es/recurso/sector-publico/territorio/Provincia/34">https://datos.iepnb.es/recurso/sector-publico/territorio/Provincia/34</a>
<a href="https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasCheckDigit">https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasCheckDigit</a>	2
<a href="https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasMunicipalityCode">https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/hasMunicipalityCode</a>	"120"
<a href="http://crossforest.eu/position/ontology/hasPolygon">http://crossforest.eu/position/ontology/hasPolygon</a>	<a href="https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/polygon/34120-4326">https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/polygon/34120-4326</a>

#### 4.2.1. Asignación mediante SPARQL

A continuación, se muestran los resultados producto de la realización de consultas de datos geospaciales RDF. Para ello se han empleado las funciones de GeoSPARQL en los almacenes de tripletas Virtuoso y Fuseki en la asignación de las geometrías de los municipios a las parcelas del IFN3, que posteriormente de compararan con las realizadas con QGIS.

##### 4.2.1.1. Virtuoso

Para la asignación de las parcelas del IFN3 a los municipios con Virtuoso se realizó la consulta que se muestra en la Figura 4.5.

```

SELECT ?plot ?muni
WHERE {
  ?plot a ifn:Plot;
        spo:hasPosition ?pos.
  ?pos spo:hasCoordinateReferenceSystem crs:4326 ;
        axis:106 ?lat ;
        axis:107 ?lng .
  ?muni spo:hasPolygon ?poly .
  ?poly wkt:asWKT ?wkt_muni .
  FILTER (bif:st_within(bif:st_point(xsd:double(?lng), xsd:double(?lat)),
        ?wkt_muni))
}

```

Figura 4.5: Consulta SPARQL para la asignación de la posición de las parcelas a sus recintos municipales utilizando Virtuoso (las IRI de los prefijos utilizados se encuentran en el Anexo A).

Sin embargo, los resultados obtenidos con esta consulta no fueron correctos en la mayoría de los casos, a pesar de que todas las parcelas fueron asignadas a al menos un municipio. Más del 70 % de las parcelas se asociaron a más de un municipio dando como resultado falsos positivos:

- 26 591 (26,85 %) parcelas asociadas a 1 municipio.
- 46 668 (47,12 %) parcelas asociadas a 2 municipios.
- 20 869 (21,07 %) parcelas asociadas a 3 municipios.
- 4 284 (4,33 %) parcelas asociadas a 4 municipios.
- 560 (0,57 %) parcelas asociadas a 5 municipios.
- 71 (0,06 %) parcelas asociadas a 6 municipios.
- 2 parcelas asociadas a 7 municipios.

En la Tabla 4.2 se presenta el caso concreto de una de las parcelas (22-0139-A-1) a la que Virtuoso asignó a 7 municipios. En la Figura 4.6 se muestra un mapa donde se localiza la parcela en el municipio de código INE 22131. No obstante Virtuoso la ha asignado además de a ese municipio, a los seis municipios de los alrededores (22006, 22068, 22901, 22028, 22032, 22130). A consecuencia de detectar este problema hemos abierto un *issue* en Virtuoso (<https://github.com/openlink/virtuoso-opensource/issues/1098>), que a fecha de la entrega de este trabajo no tiene respuesta que solucione esta cuestión.

Tabla 4.2: Municipios asignados por Virtuoso a la parcela 22-0139-A-1 utilizando la consulta de la Figura 4.5 (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A).

<b>plot</b>	<b>muni</b>
plot: 22-0139-A-1	municipality: 22006
plot: 22-0139-A-1	municipality:22068
plot:22-0139-A-1	municipality:22131
plot:22-0139-A-1	municipality:22901
plot:22-0139-A-1	municipality:22028
plot:22-0139-A-1	municipality:22032
plot:22-0139-A-1	municipality:22130

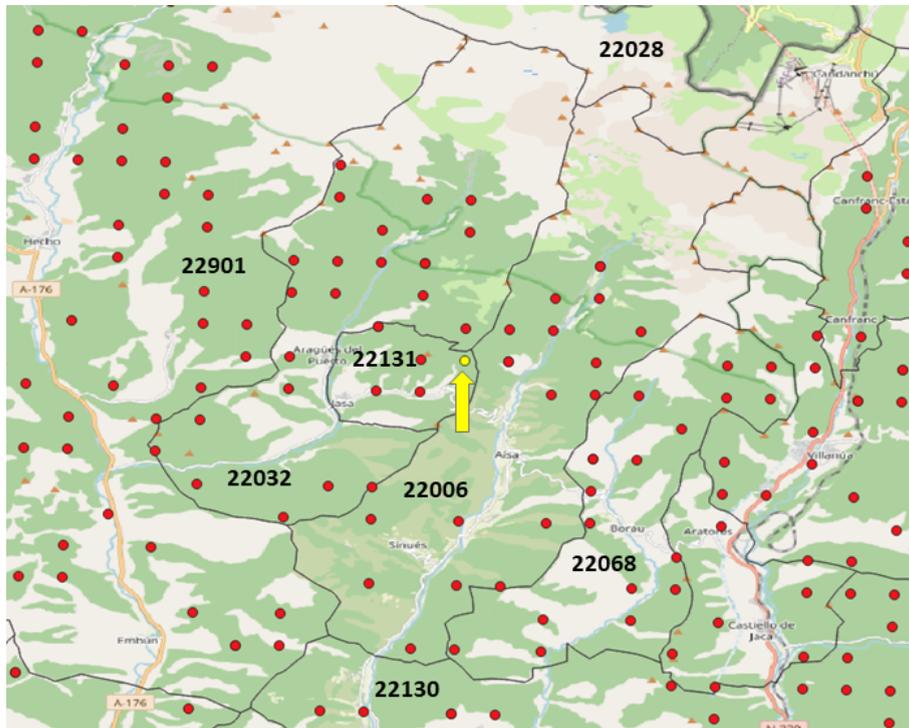


Figura 4.6: Mapa de localización de la parcela 22-0139-A-1 (en amarillo) dentro de su municipio y los municipios colindantes a los que Virtuoso la asigna.

#### 4.2.1.2. Fuseki

La consulta GeoSPARQL realizada para la asignación de las parcelas del IFN3 mediante Fuseki se muestra en la Figura 4.7:

```

SELECT ?plot ?muni
WHERE {
  ?plot a ifn:Plot;
    spo:hasPosition ?spo.
  ?spo spo:hasCoordinateReferenceSystem crs:4326>;
    axis:106 ?lat;
    axis:107 ?lng .
  ?muni spo:hasPolygon ?poly .
  ?poly wkt:asWKT ?wkt_muni.
  BIND(spatialF:convertLatLon(?lat, ?lng) as ?point)
  FILTER(geof:sfWithin(?point, ?wkt_muni))
}

```

Figura 4.7: Consulta SPARQL para la asignación de la posición de las parcelas a sus recintos municipales utilizando Fuseki (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A).

Los resultados obtenidos en esta asignación fueron los siguientes:

- 99 026 (99,98 %) parcelas asignadas a un único municipio.
- 19 (0,02 %) parcelas no asignadas a ningún municipio.

### 4.2.2. Asignación mediante un SIG

Por otro lado, para la unión de las parcelas a sus municipios por localización mediante QGIS se desarrolló un *script* de shell (Anexo J). Este *script* descarga las parcelas con sus coordenadas en WGS84 del *dataset Cross-Forest* en CSV, convierte de forma automática el CSV a *Shapefile*, crea el índice espacial para el *Shapefile* de las parcelas y el de los recintos municipales y realiza la unión espacial de ambos archivos, generando otro *Shapefile* de tipo punto con la asignación. La Tabla de atributos del archivo resultante muestra el IRI de la parcela según las especificaciones descritas en el *dataset Cross-Forest* asociado a las coordenadas donde se localiza el centro de la parcela y las propiedades del recinto municipal en el que se localiza dicha parcela publicadas por el IGN, donde se incluye el código INE (IDmuni) extraído previamente. La asignación dio como resultado final:

- 99 026 (99,98 %) parcelas asignadas a un único municipio.
- 19 (0,02 %) parcelas no asignadas a ningún municipio.

En la Tabla 4.3 se muestra la asignación realizada para la parcela 22-0139-A-1:

Tabla 4.3: Asignación de la parcela 22-0139-A-1 al municipio de código INE 22131 y sus propiedades asociadas mediante QGIS.

<b>plot</b>	<a href="https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/plot/22-0139-A-1">https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/plot/22-0139-A-1</a>
<b>lng</b>	-0.632501
<b>lat</b>	42.705148999999999
<b>INSPIRED</b>	ES.IGN.BDDAE.3402222131
<b>COUNTRY</b>	COUNTRY: ES
<b>NATLEV</b>	<a href="https://inspire.ec.europa.eu/codelist/AdministrativeHierarchyLevel/4thOrder">https://inspire.ec.europa.eu/codelist/AdministrativeHierarchyLevel/4thOrder</a>
<b>NATLEVNAME</b>	Municipio
<b>NATCODE</b>	3402222131
<b>NAMEUNIT</b>	Jasa
<b>CODNUT1</b>	ES2
<b>CODNUT2</b>	ES24
<b>CODNUT3</b>	ES241
<b>IDmuni</b>	22131

### 4.2.3. Resultados de las asignaciones en RDF

Una vez realizadas las consultas SPARQL para obtener las asignaciones de las parcelas a los municipios que las contienen, se llevó a cabo una consulta SPARQL CONSTRUCT para Virtuoso (Anexo H) y Fuseki (Anexo I) mediante línea de comandos para crear grafos RDF de forma automática con los resultados obtenidos. En el caso de QGIS, el *Shapefile* resultante

de la asignación se exportó a GeoJSON y después de GeoJSON a RDF mediante un *script* SPARQL-generate desarrollado para este propósito (Anexo G).

Un ejemplo del resultado obtenido en formato *Turtle* mediante Fuseki para la parcela 22-0139-A-1 se puede ver en la Figura 4.8, donde se observa la tripleta formada por el identificador de dicha parcela (sujeto), la relación creada para la asignación de parcelas a municipios en el *dataset Cross-Forest* `ifn:isInMunicipality` (predicado) y el identificador del municipio de código INE 22131 (objeto). Esta asignación coincide con la representación visual en la Figura 4.6. La generación de estas tripletas permiten publicar los datos directamente en el *dataset Cross-Forest* y así obtener el municipio en el que esta la parcela como una propiedad de la misma sin necesidad de hacer otra consulta utilizando GeoSPARQL. De esta manera, al consultar las parcelas del IFN3 también se obtendrá, a parte de otras características como las especies que contiene, la posición etc., el municipio al que pertenece que esta asociado a su vez con sus características (provincia, nombre, geometría...).

```
@prefix ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/> .
@prefix municipality: <https://datos.iepnb.es/recurso/sector-publico/medio-
ambiente/ifn/municipality/> .
@prefix plot: <https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn
/plot/>

plot:22-0139-A-1 ifn:isInMunicipality municipality:22131 .
```

Figura 4.8: Ejemplo parcela 22-0139-A-1 asignada al código INE de un municipio en formato *Turtle*.

### 4.3. Comparación y validación de los resultados obtenidos

Tras las asignaciones de las parcelas a los municipios mediante los tres procedimientos (consultas GeoSPARQL en Virtuoso y Fuseki, y unión por localización en QGIS), en la Tabla 4.4 se comparan los resultados obtenidos. Los tres procedimientos se han llevado a cabo con un equipo con la siguientes características: Procesador Intel(R) Core(TM) i7-1165G7 2.80GHz, memoria RAM instalada 16,0 GB, disco de almacenamiento SSD 512 GB y sistema operativo Windows 11 Pro (x64).

En la Tabla 4.4 se puede observar que los resultados generados en Fuseki y QGIS son idénticos, aunque los tiempos empleados en la asignación por Fuseki son mucho mayores. Mientras que Virtuoso asigna las parcelas a los

municipios en relativamente poco tiempo, similar al empleado por QGIS, pero asigna más del 70 % de las parcelas a municipios incorrectos, lo que hace que sea necesario desestimar este resultado.

Tabla 4.4: Comparación de los resultados de la asignación de las parcelas del IFN3 a sus municipios con Virtuoso Fuseki y QGIS.

	<b>Virtuoso</b>	<b>Fuseki</b>	<b>QGIS</b>
<b>Tiempo de ejecución</b>	0:02:98	1:56:46	0:02:09
<b>Nº de parcelas asignadas a municipios</b>	99 045 (100,00 %)	99 026 (99,98 %)	99 026 (99,98 %)
<b>Nº de asignaciones correctas</b>	26 591 (26,85 %)	99 026 (99,98 %)	99 026 (99,98 %)
<b>Nº de asignaciones incorrectas</b>	72 454 (73,15 %)	0	0
<b>Nº de parcelas no asignadas</b>	0	19 (0,02 %)	19 (0,02 %)

Por otro lado, se han detectado 19 parcelas que han sido asignadas por Virtuoso a los municipios más cercanos pero que no han sido asignados a ningún municipio por Fuseki o QGIS. Esto es debido a que la coordenada central de estas 19 parcelas se encuentra fuera del límite de cualquier municipio español. En la Tabla 4.5 se muestran estas parcelas, sus coordenadas, su localización real, el municipio más cercano y la distancia al mismo.

Además, en la Tabla 4.6 se compararon las asignaciones correctas realizadas con Fuseki o QGIS con las anotaciones del código INE asociado a las parcelas en el IFN3. Esta comparación se realizó tanto con las anotaciones con respecto a los códigos INE de los municipios realizadas antes del apeo de las parcelas en campo (PCDatosMap) y una vez se han trasladado los equipos al lugar donde se encuentran las parcelas (PCParcelas). En la validación de la asignaciones de las parcelas a los códigos de los municipios en el IFN3 se han tenido en cuenta los cambios que se han producido en los códigos de los municipios desde 2007 hasta 2023, los cuales se muestran en el Anexo B. Este análisis comparativo de los resultados se ha llevado a cabo mediante el software R (R Core Team, 2017) (Anexo K), donde se han homogeneizado los datos y se han puesto en común (Tabla 4.6).

Tabla 4.5: Parcelas que se encuentran fuera de los límites de los municipios españoles.

<b>Código Parcela (Prov-Est- Cla-Subcl)</b>	<b>Longitud (WGS84)</b>	<b>Latitud (WGS84)</b>	<b>Localización</b>	<b>Municipio más cercano (código INE)</b>	<b>Distancia al municipio más cercano (m)</b>
07-0254-N-N	1.529893	39.117113	Mar Mediterráneo	Sant Joan de Labritja (07050)	4
07-0277-N-N	1.599989	39.081934	Mar Mediterráneo	Sant Joan de Labritja (07050)	2
07-0429-N-N	1.510364	38.945664	Mar Mediterráneo	Santa Eulària des Riu (07054)	20
07-0510-N-N	3.197933	39.928199	Mar Mediterráneo	Pollença (07042)	1
07-0517-N-N	3.162747	39.928199	Mar Mediterráneo	Pollença (07042)	6
07-0815-N-N	3.454003	39.729206	Mar Mediterráneo	Capdepera (07014)	80
07-1157-N-N	2.533785	39.485881	Mar Mediterráneo	Calvià (07011)	49
07-1257-N-N	2.975705	39.351654	Mar Mediterráneo	Campos (07013)	26
07-1325-N-N	3.103292	39.297546	Mar Mediterráneo	Santanyi	9
07-1331-N-N	3.091683	39.288545	Mar Mediterráneo	Santanyi (07057)	61
10-3125-A-3C	-7.273028	39.311689	Portugal	Valencia de Alcántara (10203)	4
10-3125-A-3E	-7.273028	39.311689	Portugal	Valencia de Alcántara (10203)	4
15-0051-A-3C	-7.860201	43.669363	Mar Cantábrico	Ortigueira (15061)	8
15-0051-A-3E	-7.860201	43.669363	Mar Cantábrico	Ortigueira (15061)	8
17-0136-A-1	2.658758	42.377957	Francia	Albanyà (17003)	176
17-3004-N-N	1.976707	42.490971	Francia	Llívia (17094)	46
31-0169-N-N	-1.424025	43.268070	Francia	Baztan (31050)	9
31-0597-N-N	-1.306336	43.068289	Francia	Luzaide/Valcarlos (31248)	12
33-0099-N-N	-6.178044	43.568172	Mar Cantábrico	Cudillero (33021)	1

Tabla 4.6: Comparación de la asignación de las parcelas a sus municipios en en IFN3.

	<b>PCDatosMap (gabinete)</b>	<b>PCParcelas (campo)</b>
<b>Parcelas con código INE</b>	80 286 (81 %)	46 443 (47 %)
<b>Parcelas sin código INE</b>	18 759 (19 %)	52 602 (53 %)
<b>Código INE igual en las dos Tablas</b>	45 670 (98 %)	
<b>Código INE distinto en las dos Tablas</b>	773 (2 %)	
<b>Parcelas bien asignadas (comparación con Fuseki/ QGIS)</b>	44 910 (56 %)	41 074 (88 %)
<b>Parcelas mal asignadas (comparación con Fuseki/QGIS)</b>	35 376 (44 %)	5 369 (12 %)

Los resultados muestran en la Tabla 4.6 que casi la mitad de las parcelas asignadas a sus municipios antes de ir a campo (PCDatosMap) están mal asignadas y que el 88 % de las parcelas asignadas a un código municipal durante los trabajos de campo (PCParcelas) son correctas, las cuales coinciden en un 98 % en las dos Tablas.

#### 4.4. Aplicación de los resultados

En las siguientes secciones se han desarrollado consultas SPARQL y mapas para mostrar la aplicabilidad de los resultados obtenidos en las secciones anteriores. Las métricas que se muestran en esta Sección se han podido llevar a cabo gracias de la integración en LOD de la base de datos de los recintos municipales españoles y el *dataset Cross-Forest* que contiene la información recogida en el IFN3.

##### 4.4.1. Estadísticas a partir de los resultados obtenidos

A partir de la asignación de las parcelas a los municipios se pueden obtener varias estadísticas realizando consultas SPARQL al conjunto de datos. Como por ejemplo, para contar el número de parcelas que tiene cada municipio se realizó la consulta que se muestra en la Figura 4.9 obteniendo que el municipio con mayor cantidad de parcelas es Cuenca (código INE: 16078) con 370 parcelas.

```

SELECT ?muni COUNT (DISTINCT ?plot) as ?nplot

WHERE{
  ?plot a ifn:Plot ;
  ifn:isInMunicipality ?muni.
}

GROUP BY ?muni
ORDER BY DESC (?nplot)

```

Figura 4.9: Consulta SPARQL para obtener el número de parcelas del IFN3 en cada municipio (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A).

También se realizó la consulta que se muestra en la Figura 4.10 para obtener el número medio de parcelas por municipio. El resultado de dicha consulta mostró que los municipios que cuentan con alguna parcela tienen de media 16 parcelas del IFN3, teniendo en cuenta que de los 8 217 municipios y comunidades jurisdiccionales que forman España, 1 828 no tienen ninguna parcela del inventario (resultado de la consulta de la Figura 4.11).

```

SELECT AVG(?nplot) as ?mean
{
  SELECT ?muni COUNT (DISTINCT ?plot) as ?nplot

  WHERE{
    ?plot a ifn:Plot ;
    ifn:isInMunicipality ?muni.
  }

  GROUP BY ?muni
}

```

Figura 4.10: Consulta SPARQL para obtener la media de parcelas del IFN3 por municipio en España (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A).

```

SELECT COUNT(DISTINCT ?municipio) as ?Count

WHERE {
  ?municipio a ifn:Municipality.
  MINUS
  {?plot ifn:isInMunicipality ?municipio.}
}

```

Figura 4.11: Consulta SPARQL para saber cuántos municipios españoles no tienen parcelas del IFN3 (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A).

#### 4.4.2. Mapas de especies dominantes por municipio

Se llevaron a cabo dos consultas SPARQL para descargar los datos de las variables asociadas a las parcelas por especie y municipio del entorno de producción (punto SPARQL del *dataset Cross-Forest*<sup>3</sup>) en formato CSV. Para obtener, por un lado, la media del área basimétrica (G) (Figura 4.12) y por otro, la media del número de pies/ha (N) (Figura 4.13) a partir de las parcelas comprendidas dentro de cada municipio.

<sup>3</sup><https://crossforest.gsic.uva.es/sparql>

```

SELECT ?muni ?species AVG(?G) as ?meanG

WHERE{
  ?plot a ifn:Plot ;
    ifn:containsSpeciesPlot ?infoSpeciesPlot ;
    ifn:isInMunicipality ?muni.
  ?infoSpeciesPlot ifn:hasBasalAreaInM2byHA ?G;
    ifn:hasSpecies ?species.
}

GROUP BY ?muni ?species

```

Figura 4.12: Consulta SPARQL para obtener el área basimétrica media (m<sup>2</sup>/ha) por especie y municipio (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A).

```

SELECT ?muni ?species AVG(?N) as ?meanN

WHERE{
  ?plot a ifn:Plot ;
    ifn:containsSpeciesPlot ?infoSpeciesPlot ;
    ifn:isInMunicipality ?muni.
  ?infoSpeciesPlot ifn:hasNumberOfTreesInUnitsByHA ?N ;
    ifn:hasSpecies ?species.
}

GROUP BY ?muni ?species

```

Figura 4.13: Consulta SPARQL para obtener el número medio de pies por hectárea por especie y por municipio (las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A).

Finalmente, utilizando el paquete `dplyr` (Wickham et al., 2023) para la manipulación de datos en R se calculó el valor más alto del área basimétrica (m<sup>2</sup>/ha) y del número de pies por hectárea para cada municipio español (Anexo L) y se transformaron los códigos de las especies definidos en el IFN3 por su correspondiente nombre científico. Todo esto permitió representar la información en los mapas que se muestran en la Figura 4.14, que muestra las especies dominantes por municipio según el área basimétrica media, y en la Figura 4.15, que presenta las especies dominantes por municipio según la media del número de pies/ha.

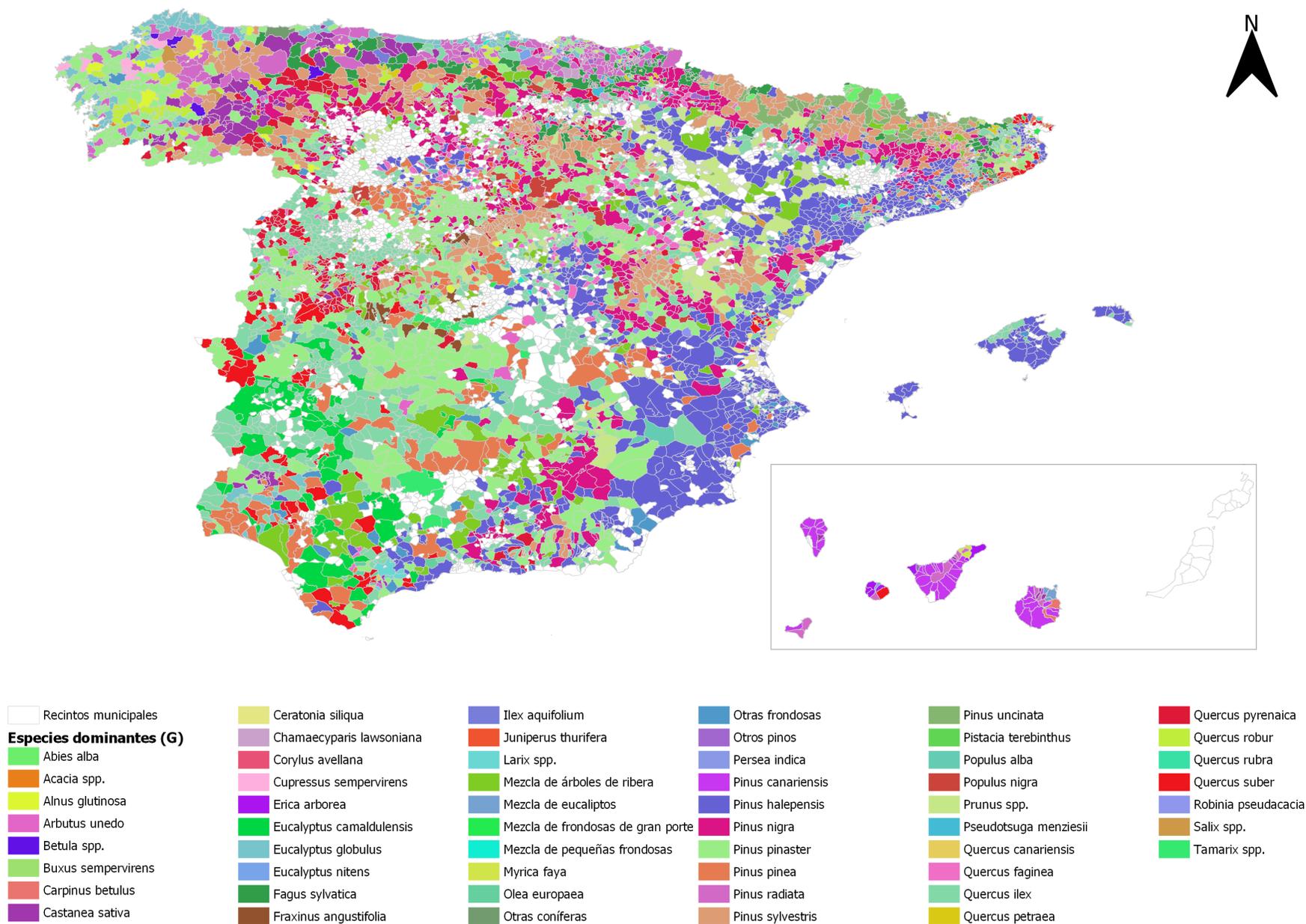


Figura 4.14: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ).

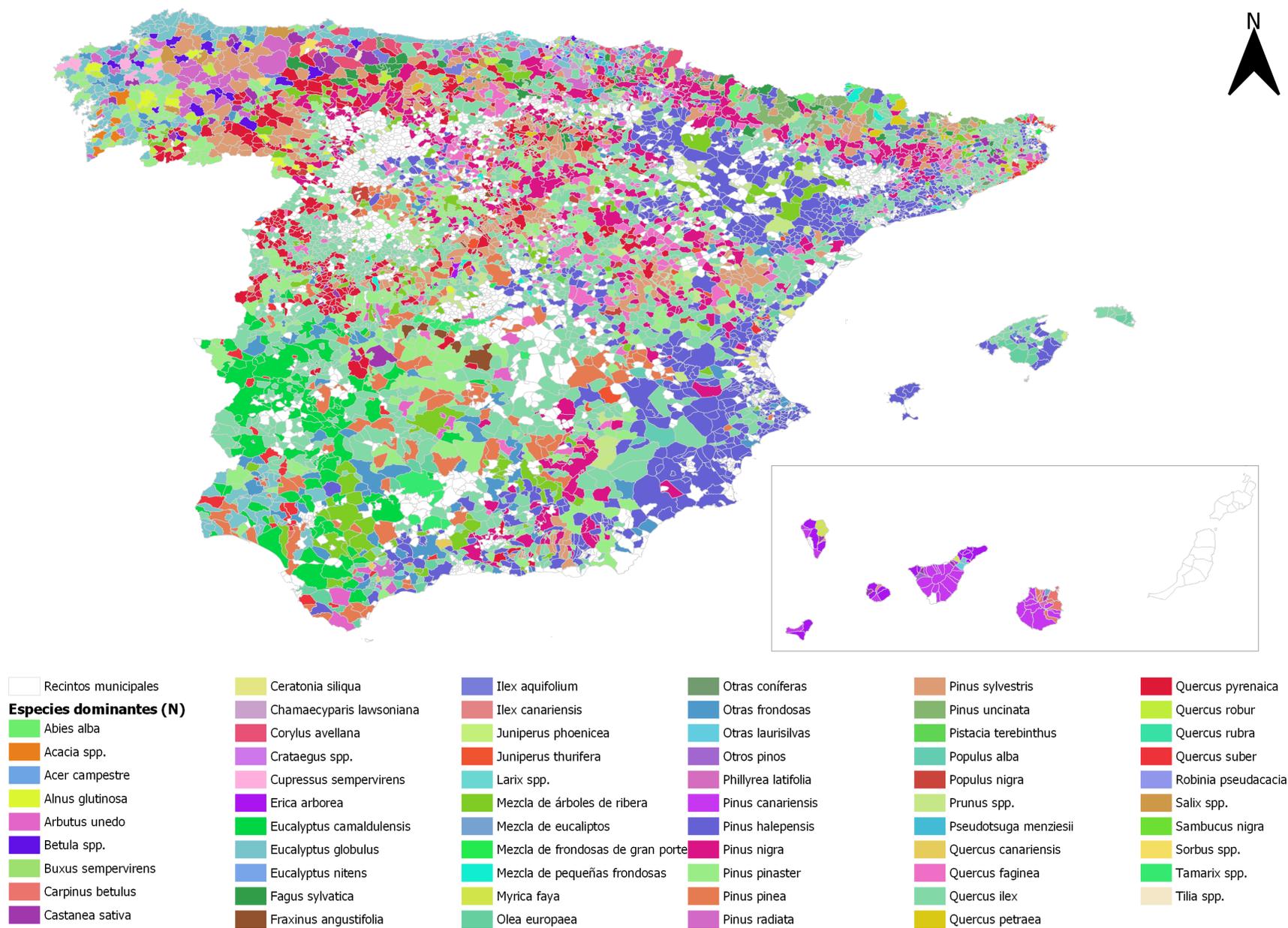


Figura 4.15: Mapa de especies dominantes por municipio según el número de pies por hectárea.

#### 4.4.2.1. Parques Nacionales

En la Tabla 4.7 se muestran los datos obtenidos de las especies dominantes por municipio, según la media del área basimétrica y el número de pies/ha del área de influencia socioeconómica de la Red de Parques Nacionales de España. Además, también se ilustra el número de parcelas del IFN3 que se encuentran dentro de esta zona. Para conocer esto último se ha desarrollado la consulta SPARQL que se muestra en la Figura 4.16, para el caso concreto de Picos de Europa. Donde se buscan las parcelas que se encuentran en los municipios que comprenden el parque y su área de influencia.

Asimismo, en el Anexo M se ilustran mediante mapas las especies dominantes de los municipios de los 15 Parques Nacionales Españoles que cuentan con parcelas del IFN3 (Timanfaya no tiene parcelas con estos datos) en base a su área basimétrica (G) y al número de pies/ha (N). Se han generado dos mapas para 14 de los 15 Parques Nacionales, uno que representa la dominancia de especies según la media de los m<sup>2</sup>/ha de cada especie por municipio y otro según la media de densidad de pies de cada especie por municipio. Sin embargo, el Parque Nacional de Tablas de Daimiel obtuvo las mismas especies dominantes tanto por área basimétrica como por pies/ha por lo que se representó en un mismo mapa.

```

SELECT COUNT(DISTINCT ?plot) as ?Count

WHERE {
  VALUES ?municipio {municipality:24106 municipality:24116 municipality:33003
    municipality:33008 municipality:33012 municipality:33043
    municipality:33046 municipality:33047 municipality:39015
    municipality:39022}
  ?plot ifn:isInMunicipality ?municipio.
}

```

Figura 4.16: Consulta SPARQL para obtener el número de parcelas del IFN3 que hay el Picos de Europa. Las IRI de los prefijos utilizados en la consulta se encuentran en el Anexo A.

Tabla 4.7: Tabla resumen de las especies dominantes por municipio dentro del área de influencia socioeconómica de la Red de Parques Nacionales española teniendo en cuenta, los datos del IFN3 de número de pies/ha y el área basimétrica.

Parque Nacional	Municipios	Nº Parcelas del IFN3	Especies arbóreas dominantes por municipio (m <sup>2</sup> /ha)	Especies arbóreas dominantes por municipio (Nº pies/ha)
<b>Aigüestortes y Lago de San Mauricio</b>	25024, 25025, 25043, 25082, 25086, 25309, 25227, 25243, 25245, 25903	394	<i>Abies alba</i> <i>Pinus sylvestris</i> <i>Pinus uncinata</i>	<i>Abies alba</i>
				<i>Betula spp.</i> <i>Mezcla de pequeñas frondosas</i> <i>Pinus halepensis</i> <i>Pinus sylvestris</i> <i>Pinus uncinata</i> <i>Quercus ilex</i>
<b>Archipiélado de Cabrera</b>	07040, 07059	31	<i>Pinus halepensis</i>	<i>Olea europaea</i> <i>Pinus halepensis</i>
<b>Cabañeros</b>	13006, 13049, 13060, 13072, 45113, 45075	241	<i>Pinus pinaster</i>	<i>Pinus pinaster</i> <i>Pinus pinea</i>
<b>Caldera de Taburiente</b>	38007, 38008, 38016, 38027, 38030, 38033, 38037, 38029, 38047	597	<i>Castanea sativa</i> <i>Pinus canariensis</i>	<i>Castanea sativa</i>
				<i>Erica arborea</i> <i>Myrica faya</i> <i>Pinus canariensis</i>
<b>Doñana</b>	21005, 21040, 41012, 41079	314	<i>Mezcla de árboles de ribera</i> <i>Pinus pinea</i>	<i>Eucalyptus camaldulensis</i>
				<i>Eucalyptus globulus</i> <i>Mezcla de árboles de ribera</i> <i>Pinus pinea</i>
<b>Garajonay</b>	38002, 38003, 30021, 38036, 38040, 38050	225	<i>Erica arborea</i> <i>Persea indica</i> <i>Pinus radiata</i> <i>Quercus suber</i>	<i>Castanea sativa</i>
				<i>Erica arborea</i> <i>Ilex canariensis</i> <i>Pinus canariensis</i>
<b>Islas Atlánticas</b>	15073, 36004, 36057, 36060	85	<i>Acacia spp.</i> <i>Eucalyptus globulus</i> <i>Pinus pinaster</i>	<i>Acacia spp.</i>
				<i>Eucalyptus globulus</i>
<b>Monfragüe</b>	10056, 10057, 10058, 10070, 10097, 10103, 10116, 10123, 10160, 10173, 10175, 10176, 10182, 10190, 10197	292	<i>Eucalyptus camaldulensis</i> <i>Pinus pinaster</i> <i>Pinus pinea</i> <i>Quercus ilex</i> <i>Quercus suber</i>	<i>Eucalyptus camaldulensis</i>
				<i>Otras frondosas</i> <i>Pinus pinaster</i> <i>Pinus pinea</i> <i>Quercus ilex</i>

Parque Nacional	Municipios	Nº Parcelas del IFN3	Especies arbóreas dominantes por municipio (Área basimétrica)	Especies arbóreas dominantes por municipio (Nº pies/ha)
<b>Ordesa y Monte Perdido</b>	22057, 22069, 22107, 22189, 22227, 22230	223	<i>Pinus sylvestris</i> <i>Pinus uncinata</i>	<i>Abies alba</i> <i>Fagus sylvatica</i> <i>Pinus nigra</i> <i>Pinus uncinata</i>
<b>Picos de Europa</b>	24106, 24116, 33003, 33008, 33012, 33043, 33046, 33047, 39015, 39022	337	<i>Castanea sativa</i> <i>Fagus sylvatica</i> <i>Quercus ilex</i> <i>Pinus radiata</i>	<i>Corylus avellana</i> <i>Fagus sylvatica</i> <i>Pinus sulvestris</i> <i>Quercus ilex</i> <i>Quercus pyrenaica</i>
<b>Sierra de Guadarrama</b>	28003, 28018, 28023, 28034, 28038, 28068, 28076, 28082, 28085, 28087, 28093, 28097, 28112, 28120, 28144, 40007, 40026, 40059, 40076, 40093, 40113, 40139, 40152, 40155, 40181, 40188, 40190, 40194, 40199, 40206, 40207, 40901, 40904, 53057	744	<i>Fraxinus angustifolia</i> <i>Pinus pinaster</i> <i>Pinus sylvestris</i> <i>Prunus spp.</i> <i>Quercus ilex</i> <i>Quercus pyrenaica</i>	<i>Fraxinus angustifolia</i> <i>Juniperus thurifera</i> <i>Pinus pinaster</i> <i>Pinus sylvestris</i> <i>Quercus ilex</i> <i>Quercus pyrenaica</i>
<b>Sierra de las Nieves</b>	29013, 29023, 29031, 29040, 29060, 29061, 29073, 29076, 29077, 29084, 29090, 29100, 29904	523	<i>Larix spp.</i> <i>Otras frondosas</i> <i>Pinus halepensis</i> <i>Pinus pinaster</i> <i>Quercus suber</i>	<i>Larix spp.</i> <i>Olea europaea</i> <i>Otras frondosas</i> <i>Pinus halepensis</i> <i>Pinus pinaster</i> <i>Pinus pinea</i> <i>Pinus radiata</i> <i>Quercus ilex</i>

Parque Nacional	Municipios	Nº Parcelas del IFN3	Especies arbóreas dominantes por municipio (Área basimétrica)	Especies arbóreas dominantes por municipio (Nº pies/ha)
<b>Sierra Nevada</b>	04001, 04002, 04005, 04014, 04020, 04023, 04028, 04029, 04030, 04038, 04043, 04045, 04046, 04055, 04057, 04065, 04067, 04071, 04073, 04077, 04080, 04901, 18005, 18010, 18018, 18030, 18032, 18033, 18035, 18040, 18042, 18068, 18069, 18071, 18074, 18094, 18097, 18108, 18112, 18114, 18116, 18117, 18119, 18121, 18123, 18134, 18143, 18151, 18163, 18176, 18189, 18182, 18183, 18193, 18901, 18903, 18904, 18907	630	<i>Mezcla de árboles de ribera</i> <i>Otras frondosas</i> <i>Pinus halepensis</i> <i>Pinus nigra</i> <i>Pinus pinaster</i> <i>Pinus sylvestris</i> <i>Quercus ilex</i> <i>Quercus pyrenaica</i>	<i>Mezcla de árboles de ribera</i> <i>Otras frondosas</i> <i>Pinus halepensis</i> <i>Pinus nigra</i> <i>Pinus pinaster</i> <i>Pinus sylvestris</i> <i>Quercus ilex</i> <i>Quercus pyrenaica</i>
<b>Tablas de Daimiel</b>	13039, 13083, 13096	11	<i>Arbutus unedo</i> <i>Mezcla de árboles de ribera</i> <i>Quercus ilex</i>	<i>Arbutus unedo</i> <i>Mezcla de árboles de ribera</i> <i>Quercus ilex</i>
<b>Teide</b>	38001, 38005, 38012, 38015, 38017, 38018, 38019, 38020, 38022, 38026, 38031, 38034, 38040, 38052	882	<i>Pinus canariensis</i> <i>Pinus radiata</i>	<i>Otros pinos</i> <i>Pinus canariensis</i>
<b>Timanfaya</b>	35029, 35034	0	-	-

#### 4.5. Publicación de los resultados en el *dataset Cross-Forest*

Para que los resultados obtenidos en este trabajo se encuentren a disposición pública, se ha creado un repositorio en GitHub<sup>4</sup> donde se encuentran los siguientes *scripts* y datos generados para su posterior publicación en el *dataset Cross-Forest*:

- *Script* SPARQL-Generate para convertir los archivos de GeoJSON a RDF (*Turtle*) de las geometrías de los recintos municipales y sus atributos asociados (Anexo D) y de la asignación de las parcelas a los municipios llevada a cabo por QGIS (Anexo G).
- Archivo en formato *Turtle* de las geometrías de los recintos municipales y las Comunidades jurisdiccionales asociadas a su código INE<sup>5</sup>.
- *Scripts* para la asignación de forma automática de las parcelas a sus recintos municipales mediante Virtuoso (Anexo H), Fuseki (Anexo I) y QGIS (Anexo J)
- Archivo en formato *Turtle* de las parcelas del IFN3 asociadas a los códigos de los municipios dónde se encuentran<sup>6</sup>.
- Archivos *Shapefile* de los polígonos de los municipios con información de las especies dominantes por municipio, según el área basimétrica (m<sup>2</sup>/ha)<sup>7</sup>.
- Archivos *Shapefile* de los polígonos de los municipios con información de las especies dominantes por municipio, según el número de pies/ha<sup>8</sup>.

---

<sup>4</sup>[https://github.com/nacresle/GeoSPARQL\\_TFM](https://github.com/nacresle/GeoSPARQL_TFM)

<sup>5</sup>[https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Datos/Recintos\\_municipales/recintos\\_TODO.ttl](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Datos/Recintos_municipales/recintos_TODO.ttl)

<sup>6</sup>[https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Datos/Asignacion\\_parcela\\_muni/Asig\\_plot\\_muni.ttl](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Datos/Asignacion_parcela_muni/Asig_plot_muni.ttl)

<sup>7</sup>[https://github.com/nacresle/GeoSPARQL\\_TFM/tree/main/Mapas/Espa%C3%B1a/dominante\\_G](https://github.com/nacresle/GeoSPARQL_TFM/tree/main/Mapas/Espa%C3%B1a/dominante_G)

<sup>8</sup>[https://github.com/nacresle/GeoSPARQL\\_TFM/tree/main/Mapas/Espa%C3%B1a/dominante\\_N](https://github.com/nacresle/GeoSPARQL_TFM/tree/main/Mapas/Espa%C3%B1a/dominante_N)

## Capítulo 5

# DISCUSIÓN

Los resultados de este TFM muestran cómo la integración de datos siguiendo los principios de los datos abiertos enlazados (LOD) permiten realzar su valor. Esto posibilita la realización de estadísticas y la creación de mapas partiendo como base de la organización territorial de municipio de forma más sencilla y eficiente. Muchos estudios forestales emplean el municipio como la entidad local en la que se basan sus investigaciones. Sin embargo, al no tener los datos integrados de los límites municipales con las bases de datos forestales, utilizan Sistemas de Información Geográfica (SIG) para relacionar de forma local los datos concretos que usan en sus estudios. Los formatos en los que se encuentra el Inventario Forestal Nacional (IFN) publicado por el MITECO y la cantidad de archivos que contiene divididos por provincias dificulta mucho hacer estos análisis a nivel de municipio, lo que se facilita enormemente si se dispone de la información integrada.

En el caso de la red de Parques Nacionales cobra especial importancia el conocimiento de las zonas forestales en las áreas de influencia socio-económica que comprenden los municipios que tienen parte o toda su superficie catalogada como Parque Nacional (Bonache López et al., 2016). Un ejemplo de la utilización de los límites municipales para el estudio de la biodiversidad en el Parque Nacional de Picos de Europa es el llevado a cabo por Guadilla-Sáez et al. (2019). En este estudio se compararon datos de bosques templados caducifolios dentro y fuera del Parque Nacional para evaluar la efectividad de las áreas protegidas utilizando datos de las parcelas del tercer inventario nacional, teniendo en cuenta también el efecto antropológico ejercido por las poblaciones de los municipios que comprenden el parque y sus alrededores. En dicha investigación utilizaron un SIG para el tratamiento de los datos de las parcelas concretas evaluadas y su relación a los términos municipales donde se localizan. Sin embargo, los resultados de este TFM proporcionan las parcelas del inventario, y su información asociada, interconectada con los datos de los términos municipales para toda

España, facilitando su posterior análisis tanto a pequeña como a gran escala.

Asimismo, otras investigaciones como la de Gil et al. (2011) utilizaron un SIG para evaluar la cantidad de carbono almacenado en los municipios de Castilla y León teniendo en cuenta la biomasa aérea a lo largo del tiempo empleando la información del IFN. Además, en este estudio se evalúa la cantidad de biomasa aérea residual de los bosques por regiones que podría ser utilizada como fuente de bionergía; y la potencialidad de ciertos municipios de Castilla y León para el establecimiento de industrias que utilicen la biomasa como materia prima, lo que favorecerían la economía local. En este caso, también se relacionan mediante un SIG los datos de las parcelas de la segunda y tercera edición de los inventarios nacionales de una región concreta con los municipios en los que se encuentran.

Por otro lado, Sánchez-García et al. (2015) pretenden crear una base de datos geospacial mediante un SIG que sirva como herramienta de apoyo para investigar el potencial de existencias de biomasa forestal en el caso de estudio de la provincia de Asturias. Su objetivo es ayudar a la administración pública en la toma de decisiones y la formulación de nuevas estrategias energéticas al proporcionar información sobre la oferta y demanda de la biomasa a distintos niveles, entre ellos el municipal. Para la oferta de biomasa por municipio utilizan como base los datos de las parcelas de la tercera edición del IFN en la que, como se ha comprobado en este trabajo, la información relativa a los términos municipales no es fiable y está parcialmente disponible. Por ese motivo, necesitan realizar la integración de los datos forestales con los límites municipales en un SIG para realizar mapas y estadísticas de oferta de biomasa por municipio y así relacionarla con la demanda en Asturias. Los datos proporcionados en este TFM facilitarían la ampliación de esta base de datos geospacial a nivel nacional, permitiendo analizar la biomasa forestal para los municipios de toda España.

Esta información es muy útil para la planificación y el manejo de montes públicos mediante la realización de planes de ordenación o de prevención contra incendios. En el trabajo de Martínez et al. (2009) se analiza la causalidad de los incendios forestales por municipio o Marey-Pérez y Rodríguez-Vicente (2009) evalúa los cambios temporales y espaciales del uso de la tierra y la situación socioeconómica ligada a la forestación a nivel de municipio en la provincia de Lugo. Además, los resultados de este TFM facilitan la gestión de los servicios ecosistémicos de las áreas forestales a nivel de municipio donde haya parcelas del IFN, pudiendo ser un impulso importante en la economía del mismo, ya que debido a la despoblación rural

y la falta de recursos económicos, muchos montes no son gestionados y los aprovechamientos se han abandonado. La gestión sostenible de los montes municipales puede tener un potencial de ganancia económica que no está reconocido por los ayuntamientos y que podría ser utilizado en su beneficio.

Actualmente, no existen muchos estudios en el ámbito de bases de datos forestales que utilicen las tecnologías de la Web Semántica para la integración y publicación de los datos en abierto. Sin embargo, estas tecnologías se están introduciendo cada vez más en la forma en la que tenemos de gestionar, manejar y almacenar los datos abiertos en la Web. En la Unión Europea entre los años 2011 y 2016 se llevó a cabo el proyecto TREES4FUTURE<sup>1</sup> en el que utilizaron los principios de la Web Semántica con el objetivo de proporcionar a la comunidad forestal europea un fácil acceso a fuentes de información no integradas para el aprovechamiento a corto y largo plazo de los recursos forestales tanto por parte de la comunidad investigadora como por parte de agentes con intereses socioeconómicos (Lokers et al., 2015).

En este trabajo, la integración del *dataset Cross-Forest* con los recintos municipales utilizando LOD ha sido posible gracias a la implementación de GeoSPARQL en almacenes de tripletas. A pesar de que este tipo de integraciones se suelen realizar con un SIG en el ámbito forestal, cuando surge la necesidad de utilizar datos geospaciales pertenecientes a diferentes fuentes y a gran escala, la falta de interoperabilidad de estos sistemas dificulta la homogeneización e integración de bases de datos y su publicación en abierto. No obstante, hay varios estudios que evalúan la adecuación de la implementación de GeoSPARQL en diferentes almacenes de tripletas mostrando que el soporte de GeoSPARQL varía considerablemente entre las distintas implementaciones actualmente disponibles.

Jovanovik et al. (2021) realizaron una comparación de los diferentes almacenes de tripletas con soporte a GeoSPARQL (GeoSPARQL Fuseki, GraphDB, Virtuoso, TryplyDB, RDF4J y Stardog) para medir de forma objetiva el alcance de los mismos y verificar el cumplimiento del estándar GeoSPARQL. Para ello se basaron en los requisitos definidos por la OGC para el estándar GeoSPARQL y realizaron una serie de consultas SPARQL a una base de datos de prueba con una única o un conjunto de respuestas esperadas para comprobar el cumplimiento de dichos requisitos por los distintos almacenes de tripletas. Uno de los requisitos testeados fue que las implementaciones debían permitir la propiedad `geo:sfWithin`, la cual usamos en este trabajo para realizar la integración de las parcelas del IFN3

---

<sup>1</sup><https://cordis.europa.eu/project/id/284181>

con los recintos municipales españoles.

En el estudio presentado anteriormente, compararon los almacenes de tripletas empleados en este TFM con otros almacenes de tripletas, utilizaron Virtuoso en su versión 7.2 de libre uso, la versión que hemos usado en la asignación de parcelas a municipios, y GeoSPARQL Fuseki versión 3.17.0, una versión anterior a la utilizada. El almacén de tripletas con mayor número de respuestas a las consultas SPARQL correctas fue Fuseki, con un 82,75 % de respuestas correctas, seguido de GraphDB con un 69,75 % y Virtuoso con un 63,46 %. Por este motivo, decidimos comparar la asignación realizada por Virtuoso, el almacén de tripletas que contenía el punto SPARQL del *dataset Cross-Forest* con Fuseki, obteniendo en este TFM que Fuseki realizaba mejor la consulta espacial `geo:sfWithin` mientras que Virtuoso daba falsos positivos.

Jovanovik et al. (2021) explican en su trabajo que Fuseki es el único sistema con soporte completo de los formatos GML y WKT y la única implementación que contiene todas las extensiones de GeoSPARQL definidas por la OGC. Sin embargo, Fuseki también produjo resultados incorrectos en funciones cubiertas por la extensión de reescritura de consultas y algunas de la extensión de geometría. Además, Fuseki no puede manejar literales WKT y GML vacíos, al igual que los otros almacenes de tripletas comparados. Estos errores de Fuseki, no se advirtieron en este trabajo, ya que únicamente se emplearon geometrías en formato WKT sin literales vacíos y la función `geo:sfWithin` que no pareció dar errores, aunque el tiempo empleado para la obtención de los resultados correctos fue mucho mayor que con Virtuoso o QGIS.

Por otro lado, Virtuoso ofrece soporte para literales WKT, pero no para literales GML, lo que, como hemos explicado anteriormente, no nos afecta en este trabajo. Sin embargo, los errores en las consultas testeadas son asociados en gran parte a que Virtuoso transforma los literales WKT a un tipo de literal interno específico de Virtuoso dando discrepancias con las respuestas esperadas. En este TFM se han detectado errores concretamente en la función `geo:sfWithin` que Virtuoso implementa como `bif:st_within` al no asignar correctamente una geometría (polígono) que contiene a otra (punto), aunque la asignación ha sido mucho más rápida que en el caso de Fuseki. Además, otros usuarios de este almacén de tripletas en su versión de libre uso también han reportado errores en esta y otras funciones de GeoSPARQL que implementa Virtuoso. Concretamente, en la actualidad hay 26 *issues* abiertos relativos a la implementación de GeoSPARQL

en Virtuoso<sup>2</sup>. No obstante, no se han encontrado *benchmarks* donde se compruebe específicamente el funcionamiento de la implementación de la función `geo:sfWithin` en Virtuoso ni se reporte, como en el caso de este TFM, su incorrecta implementación.

En este trabajo se realizó, además, la comparación de la asignación de parcelas a municipios utilizando almacenes de tripletas y un SIG de uso libre como QGIS. En los estudios realizados de manejo de datos geospaciales está muy extendida la utilización de herramientas SIG sobre todo en el ámbito científico, como se ha mencionado anteriormente, y no tanto la utilización de almacenes de tripletas y LOD. Sin embargo, hemos podido comprobar que el uso de ambas herramientas puede ser complementario para el manejo de datos geospaciales. Por un lado la eficiencia de un SIG en tiempo y bondad de los resultados es mejor que la de los almacenes de tripletas utilizados (Virtuoso y Fuseki). Sin embargo, es necesario convertir los datos en formato LOD a otro formato compatible con los SIG y luego transformarlos otra vez a LOD para poder publicarlos de forma abierta y enlazada con las distintas bases de datos. Mientras que utilizando Fuseki, los resultados obtenidos en formato LOD son satisfactorios aunque la obtención de los mismos es más lenta. No obstante para este caso concreto, donde la asignación de las parcelas a sus términos municipales es un proceso que una vez comprobado que es correcto solo se realizaría una vez para obtener el grafo con los datos enlazados de parcelas y municipios. Por lo tanto, los resultados obtenidos con Fuseki permiten llevar a cabo el proceso en menos pasos aunque la generación de dicho resultado sea más lenta. En el caso de este TFM el tiempo empleado no es un factor especialmente relevante.

---

<sup>2</sup><https://github.com/openlink/virtuoso-opensource/issues?q=is%3Aissue+is%3Aopen+geosparql>



## Capítulo 6

# CONCLUSIONES

El Inventario Forestal Nacional es una base de datos compuesta por una gran cantidad de información actualizada periódicamente de los ecosistemas forestales españoles y que, actualmente, en muchos casos no se está aprovechando todo su potencial debido a la dificultad en el manejo e integración de la información publicada. Además, con el paso de las ediciones del mismo se van añadiendo nuevas variables de interés creando una base de información abierta cada vez más completa para el estudio, gestión y conservación de estos ecosistemas. Sin embargo, como hemos visto anteriormente, la dificultad de manejo de los formatos en los que se encuentra disponible esta base de datos hace que la mayoría de los estudios que la emplean sean en zonas muy concretas, o que la información no sea muy utilizada en el apoyo para la gestión sostenible de estas zonas.

Asimismo, se ha comprobado en este trabajo que la información relativa a los términos municipales en los que se encuentran las parcelas del inventario están parcialmente disponibles y son poco fiables. Esta situación hace más complicada la localización de las parcelas y el conocimiento por parte de los agentes regionales de cuáles se encuentran dentro de sus límites municipales y, por tanto, pueden tener en cuenta como potencial fuente de información para diferentes propósitos. La integración de los resultados generados en este TFM con los datos del inventario y otras bases de datos forestales en formato LOD en el *dataset Cross-Forest* facilita enormemente la utilización y acceso a esta información por parte, no solo del personal investigador, sino también del personal gestor de las administraciones provinciales y municipales.

Por otra parte, se han cumplido los objetivos propuestos al inicio de esta trabajo. Se han convertido los recintos municipales publicados por el Instituto Geográfico Nacional a datos abiertos enlazados (LOD) lo que ha hecho posible asignar de forma automática las parcelas del IFN3 a los municipios utilizando GeoSPARQL. Además, se ha verificado la asignación

con diferentes almacenes de tripletas que implementan GeoSPARQL, comprobando que Virtuoso no ha realizado una correcta implementación de la función `geo:sfWithin` y que Fuseki llevó a cabo de forma adecuada la asignación de las 99 026 parcelas a sus municipios pero empleando más tiempo (casi dos horas). También se ha utilizado QGIS, Sistema de Información Geográfica de libre uso popular en el manejo de datos geospaciales, obteniendo buenos resultados en poco tiempo pero mediante un proceso más largo para lograr obtener las asignaciones en formato LOD. Al comparar los resultados obtenidos con la información disponible de los códigos municipales asociados a las parcelas en el IFN3 hemos verificado que existen algunos errores y discrepancias.

Como consecuencia de este trabajo, además de los objetivos propuestos en un principio se han revisado los datos de las parcelas del IFN3 comprobando que existían algunos errores de anotación entre las Tablas publicadas por el MITECO. Del mismo modo, se han localizado 19 parcelas no asignadas a ningún municipio por QGIS o Fuseki debido a que las posiciones de las mismas las localizan fuera de los límites geográficos de España. Por otro lado, con toda la información verificada e integrada en formato LOD ha sido posible realizar estadísticas por municipio y mapas de las especies dominantes, centrándonos en el caso de los Parques Nacionales por ser zonas de especial importancia para su conservación y gestión. Asimismo, todos los datos y *script* generados se han publicado en GitHub siguiendo los principios FAIR (*Findable, Accessible, Interoperable, and Reusable*) y se pretende publicar de forma integrada en el entorno de producción del *dataset Cross-Forest*.

Finalmente, este trabajo se ha presentado en el congreso internacional *Conference Artificial Intelligence and Ecosystems Management* y enviado al *XXVI IUFRO WORLD CONGRESS*, en el que esta pendiente de aprobación. Además, abre un gran número de posibilidades para trabajos futuros basados en la entidad territorial de municipio, utilizando las funciones de GeoSPARQL para el manejo de datos espaciales y la potencialidad de los LOD para interconectar bases de datos. Una de las operaciones que soporta GeoSPARQL y que no se ha incluido en este trabajo es el cálculo de distancias entre dos geometrías que se podría emplear para conocer la distancia entre las parcelas del IFN3 y el límite de los municipios más cercanos. Además, la integración de los datos generados en el *dataset Cross-Forest* posibilita la utilización de los datos del Mapa Forestal Nacional que se encuentra también incluido en esta base de datos y que contiene la superficie del uso del suelo de España dividido en teselas. Con esta información sería viable calcular la su-

perficie de los tipos de ecosistemas forestales que contiene cada municipio, información muy relevante y que facilitaría en gran medida su gestión.



## Capítulo 7

# AGRADECIMIENTOS

Esta investigación se ha desarrollado dentro del marco del proyecto LOD.For.Trees. Particularmente me gustaría dar las gracias a mis tutores, cuya orientación y disposición han sido fundamentales para integrar el mundo forestal y los Datos Abiertos Enlazados. A José Miguel Giménez, quien no ha dudado en resolver y ayudar con cualquier consulta que me ha ido surgiendo. A Elena Camacho Arranz, por su disponibilidad a la hora de proporcionarnos los datos y las aclaraciones necesarias con respecto a las delimitaciones municipales españolas. Y por último, a Tragsatec por solventar las dudas con respecto a los datos del inventario forestal nacional.



# Referencias

- Alberdi, I., Syoval, V., Condes, S., Cañellas, I., & Vallejo, R. (2016). El Inventario Forestal Nacional español, una herramienta para el conocimiento, la gestión y la conservación de los ecosistemas forestales arbolados. *Ecosistemas*, 25(3), 88-97. <https://doi.org/10.7818/ECOS.2016.25-3.10>
- Alberdi, I., Vallejo, R., G., Á.-G., Condes, S., González-Ferreiro Eduardo, Guerrero, S., Hernández, L., Martínez-Jauregui Maria, Montes, F., Oliveira, N., Pasalodos-Tato Maria, Robla, E., D., R.-G., Sánchez-González Mariola, Syoval, V., Miguel Alfonso, S., Sixto, H., & Cañellas, I. (2017). The multi-objective Spanish National Forest Inventory. *Forest Systems*, 26(2), e04S. <https://doi.org/10.5424/fs/2017262-10577>
- Ayanz, A. S. M. (2014). Gestión y conservación en la Red de Parques Nacionales de España. *Ambienta: La revista del Ministerio de Medio Ambiente*, (106), 16-25.
- Berners-Lee, T. (2006). Linked Data - Design issues [Accessed: Feb. 20, 2023]. <https://www.w3.org/DesignIssues/LinkedData.html>
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data: The Story so Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5, 1-22. <https://doi.org/10.4018/978-1-60960-593-3.ch008>
- BOE. (1985). Ley 7/1985, de 2 de abril, Reguladora de las Bases del Régimen Local. <https://www.boe.es/buscar/doc.php?id=BOE-A-1985-5392>
- BOE. (1986). Real Decreto 1690/1986, de 11 de julio, por el que se aprueba el Reglamento de Población y Demarcación Territorial de las Entidades Locales. <https://www.boe.es/buscar/act.php?id=BOE-A-1986-21944>
- BOE. (2003). Ley 43/2003, de 21 de noviembre, de Montes. <https://www.boe.es/buscar/act.php?id=BOE-A-2003-21339>
- BOE. (2007). Directiva 2007/2/CE del Parlamento Europeo y del Consejo, de 14 de marzo de 2007, por la que se establece una infraestructura de información espacial en la Comunidad Europea (Inspire). <https://www.boe.es/buscar/doc.php?id=DOUE-L-2007-80587>
- BOE. (2014). Ley 30/2014, de 3 de diciembre, de Parques Nacionales. <https://www.boe.es/buscar/act.php?id=BOE-A-2014-12588>
- Bonache López, J., de Mingo-Sancho García, G., Serrada Hierro, J., Amengual Ramis, P., Perales Rodríguez, J., Martínez Ruiz, R., Rodado Hernández, S., & Albornos Llorente, E. (2016). El seguimiento y la evaluación a largo plazo en la Red española de Parques Nacionales: *Ecosistemas*, 25(1), 31-48. <https://doi.org/10.7818/ECOS.2016.25-1.05>
- Breunig, M., Bradley, P. E., Jahn, M., Kuper, P., Mazroob, N., Rösch, N., Al-Doori, M., & Stefanakis, M., Emmanuel y Jadidi. (2020). Geospatial Data Management Research: Progress and Future Directions. *ISPRS International Journal of Geo-Information*, 9(2). <https://doi.org/10.3390/ijgi9020095>
- Brodeur, J., Bédard, Y., & Edwards, B., Geoffrey y Moulin. (2003). Revisiting the Concept of Geospatial Data Interoperability within the Scope of Human Communication Processes.

- Transactions in GIS*, 7(2), 243-265. <https://doi.org/https://doi.org/10.1111/1467-9671.00143>
- Broszofske, K. D., Froese, R. E., & Falkowski, A., Michael J. y Banskota. (2013). A Review of Methods for Mapping and Prediction of Inventory Attributes for Operational Forest Management. *Forest Science*, 60(4), 733-756. <https://doi.org/10.5849/forsci.12-134>
- Butler, H., & Daly, M. (2016). The GeoJSON format. *RFC 7946*. <https://www.rfc-editor.org/rfc/rfc7946>
- de la Fuente, G. C. (2020). *Guía práctica para la publicación de Datos Espaciales*. <https://datos.gob.es/es/documentacion/guia-practica-para-la-publicacion-de-datos-espaciales>
- ESRI. (s.f.). Geoprocessing considerations for shapefile output. [%5Curl%7Bhttps://desktop.arcgis.com/es/arcmap/latest/manage-data/shapefiles/geoprocessing-considerations-for-shapefile-output.htm%7D](https://desktop.arcgis.com/es/arcmap/latest/manage-data/shapefiles/geoprocessing-considerations-for-shapefile-output.htm%7D)
- Fierro-García, B., Roldán-Zamarrón, A., Lerner-Cuzzi, M., Baiget-Llompарт, R., Aguirre-Briones, F., Gonzalvo-Morales, V., Vega-Gorgojo, G., Gimenez García, J., Ordoñez-Alonso, C., Bravo-Oviedo, F., Fonseca, A., Gomes, A., Fanego-Lobo, A., & Lorenzana-Campillo, J. (2022). CROSS-FOREST, un proyecto de datos forestales abiertos de España y Portugal. *Sociedad Española de Ciencias Forestales*.
- Garoia Rubio, E., Rodríguez-Ponga Salamanca, P., & Villan Criado, I. (1982). Métodos biyectivos de detección de errores: ventajas del método ABC. *Estadística española*, (94), 67-101.
- Gil, M. V., Blanco, D., Carballo, M. T., & Calvo, L. F. (2011). Carbon stock estimates for forests in the Castilla y León region, Spain. A GIS based method for evaluating spatial distribution of residual biomass for bio-energy. *Biomass and Bioenergy*, 35(1), 243-252. <https://doi.org/https://doi.org/10.1016/j.biombioe.2010.08.004>
- Group, W. S. W. I. (2003). *Basic Geo (WGS84 lat/long) Vocabulary* [Accessed: May. 01, 2023]. <https://www.w3.org/2003/01/geo/>
- Guadilla-Sáez, S., Pardo-de-Santayana, M., Victoria, R.-G., & Svenning, J.-C. (2019). Biodiversity conservation effectiveness provided by a protection status in temperate forest commons of north Spain. *Forest Ecology and Management*, 433, 656-666. <https://doi.org/https://doi.org/10.1016/j.foreco.2018.11.040>
- Guerra-Hernández, J., & Pascual, A. (2021). Using GEDI lidar data and airborne laser scanning to assess height growth dynamics in fast-growing species: a showcase in Spain. *Forest Ecosystems*. <https://doi.org/10.1186/s40663-021-00291-2>
- Jovanovik, M., Homburg, T., & Spasić, M. (2021). A GeoSPARQL Compliance Benchmark. *ISPRS International Journal of Geo-Information*, 10(7). <https://doi.org/10.3390/ijgi10070487>
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. SAGE Publications Ltd. <https://doi.org/https://doi.org/10.4135>
- Kügeler, S., Antje y Jirka. (2022). *Geospatial Trends 2022: Opportunities for data.europa.eu from emerging trends in the geospatial community*. <https://doi.org/10.2830/077789>
- Lillesand, T. M., Kiefer, R. W., & Chipman, J. (2008). *Remote sensing and image interpretation* (6th). John Wiley & Sons.
- Lokers, R., van Randen, Y., Knapen, R., Gaubitzer, S., Zudin, S., & Janssen, S. (2015). Improving Access to Big Data in Agriculture and Forestry Using Semantic Technologies. En E. Garoufallou, R. J. Hartley & P. Gaitanou (Eds.), *Metadata and Semantics Research* (pp. 369-380). Springer International Publishing.
- Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2005). *Geographic Information Systems and Science*. John Wiley & Sons.
- Lorenzo-Sáez, E., Oliver-Villanueva, J.-V., Lemus-Zúñiga, L.-G., Urchueguía, J. F., & Lerma-Arce, V. (2022). Development of sectorial and territorial information system to monitor

- GHG emissions as local and regional climate governance tool: Case study in Valencia (Spain). *Urban Climate*, 42, 101125. <https://doi.org/https://doi.org/10.1016/j.uclim.2022.101125>
- Marey-Pérez, M., & Rodríguez-Vicente, V. (2009). Forest transition in Northern Spain: Local responses on large-scale programmes of field-afforestation [Formalisation of Ly Rights in the South]. *Ly Use Policy*, 26(1), 139-156. <https://doi.org/https://doi.org/10.1016/j.lyusepol.2008.02.004>
- Martínez, J., Vega-García, C., & Chuvieco, E. (2009). Human-caused wildfire risk rating for prevention planning in Spain. *Journal of Environmental Management*, 90(2), 1241-1252. <https://doi.org/https://doi.org/10.1016/j.jenvman.2008.07.005>
- Meddens, A. J. H., Steen-Adams, M. M., yrew T Hudak, Mauro, F., & y Jacob Strunk, P. M. B. (2022). Specifying geospatial data product characteristics for forest and fuel management applications. *Environmental Research Letters*, 17(4), 045025. <https://doi.org/10.1088/1748-9326/ac5ee0>
- MITECO. (s.f.-a). *Historia de la Red de Parques Nacionales*. Consultado el 31 de mayo de 2023, desde <https://www.miteco.gob.es/es/parques-nacionales-oapn/red-parques-nacionales/historia.aspx>
- MITECO. (s.f.-b). *Historia del Inventario Forestal Nacional de España*. Consultado el 20 de febrero de 2023, desde [https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/index\\_historia\\_inventario\\_forestal\\_espana.aspx](https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/index_historia_inventario_forestal_espana.aspx)
- OGC. (2012a). GeoSPARQL Specification. <https://docs.opengeospatial.org/is/12-052r2/12-052r2.html%7D>
- OGC. (2012b). *OGC GeoSPARQL: A geographic query language for RDF data, version 1.0*.
- OGC. (2020). *OGC Benefits of Representing Spatial Data Using Semantic and Graph Technologies*. Consultado el 20 de febrero de 2023, desde <http://docs.ogc.org/wp/19-078r1/19-078r1.html>
- OGC. (2021). OGC GeoPackage. <http://www.geopackage.org/%7D>
- Ortuño-Pérez, S. F., & Martín-Fernández, A. J. (2006). Forest externalities, demography and rural development in inly Spain. *Forest Policy and Economics*, 8(2), 109-122. <https://doi.org/https://doi.org/10.1016/j.forpol.2004.05.007>
- Otsu, K., Díaz, L., & Schade, S. (2011). A Workflow for Analyzing Interoperability of Geospatial Data Sets at Different Administrative Levels - A Case Study in Forest Fire Data.
- Pascal, H. (2021). A Review of the Semantic Web Field. *Commun. ACM*, 64(2), 78-83. <https://doi.org/10.1145/3397512>
- Perry, M., Herring, J. R., & Kolas, D. (2012). GeoSPARQL: Enabling a geospatial semantic web. *Semantic Web*, 3(4), 355-370.
- R Core Team. (2017). *R: A Language and Environment for Statistical Computing*. <https://www.R-project.org/>
- Saeedeh, C., Eskyari y Emilio. (2015). Fire danger assessment in Iran based on geospatial information. *International Journal of Applied Earth Observation and Geoinformation*, 42, 57-64. <https://doi.org/https://doi.org/10.1016/j.jag.2015.05.006>
- Sánchez-García, S., Canga, E., Tolosana, E., & Majada, J. (2015). A spatial analysis of woodfuel based on WISDOM GIS methodology: Multiscale approach in Northern Spain. *Applied Energy*, 144, 193-203. <https://doi.org/https://doi.org/10.1016/j.apenergy.2015.01.099>
- Sayar, A., Pierce, M., & Fox, G. (2014). OGC Compatible Geographical Information Services.
- Sieber, R. (2019). Open Data and Geospatial [Print version]. En T. Davies, S. Walker, M. Rubinstein & F. Perini (Eds.), *The State of Open Data: Histories and Horizons*. African Minds; International Development Research Centre. <https://doi.org/10.5281/zenodo.2677845>

- Taylor, L., Janowicz, K., Cox, S., Portele, C., & Purves, R. S. (2017). Spatial Data on the Web Best Practices [[Online; accessed 6-April-2023]]. [%5Curl%7Bhttps://www.w3.org/TR/sdw-bp/%7D](https://www.w3.org/TR/sdw-bp/)
- Vega-Gorgojo, G., Giménez-García, J., Ordóñez, C., & Bravo, F. (2022). Pioneering Easy-to-Use Forestry Data with Forest Explorer. *Semantic Web*, 13, 147-162. <https://doi.org/10.3233/SW-210430>
- W3C. (2012a). *OWL 2 Web Ontology Language*. <https://www.w3.org/TR/owl2-overview/>
- W3C. (2012b). *OWL 2 Web Ontology Language*. <https://www.w3.org/TR/owl2-overview/>
- W3C. (2014). *RDF 1.1 Concepts and Abstract Syntax*. <https://www.w3.org/TR/rdf11-concepts/>
- W3C. (2017). *Spatial Data on the Web Best Practices*. <https://www.w3.org/TR/sdw-bp/>
- W3C. (2023). *RDF 1.2 Schema*. <https://www.w3.org/TR/rdf-schema/>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation* [<https://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>].
- Zhang, C., Li, W., & Zhao, T. (2008). Geospatial data sharing based on Geospatial Semantic Web technologies. *Journal of Spatial Science*, 52. <https://doi.org/10.1080/14498596.2007.9635121>

# **Anexos**



## Anexo A

# Lista de los prefijos utilizados en este trabajo con sus IRI correspondientes

Tabla A.1: Prefijos y sus IRI correspondientes utilizados en este trabajo por orden de aparición

Prefijos	IRI
ifn	<a href="https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/">&lt;https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/&gt;</a>
municipality	<a href="https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/municipality/">&lt;https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/municipality/&gt;</a>
spo	<a href="http://crossforest.eu/position/ontology/">&lt;http://crossforest.eu/position/ontology/&gt;</a>
crs	<a href="http://crossforest.eu/epsg/data/crs/">&lt;http://crossforest.eu/epsg/data/crs/&gt;</a>
muni	<a href="http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#">&lt;http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio#&gt;</a>
axis	<a href="http://epsg.w3id.org/ontology/axis/">&lt;http://epsg.w3id.org/ontology/axis/&gt;</a>
wkt	<a href="http://www.opengis.net/ont/geosparql#">&lt;http://www.opengis.net/ont/geosparql#&gt;</a>
polygon	<a href="https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/polygon/">&lt;https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/polygon/&gt;</a>
plot	<a href="https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/plot/">&lt;https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/plot/&gt;</a>
spatialF	<a href="http://jena.apache.org/function/spatial#">&lt;http://jena.apache.org/function/spatial#&gt;</a>
geof	<a href="http://www.opengis.net/def/function/geosparql/">&lt;http://www.opengis.net/def/function/geosparql/&gt;</a>



## Anexo B

# Cambios en los códigos INE de los municipios de las ediciones de 2007 y 2023

Tabla B.1: Tabla que muestra los cambios en los códigos INE de los municipios a partir de 2007, cuando se publicó el IFN3, hasta 2023, año de finalización de este TFM

<b>Provincia</b>	<b>INE 2007</b>	<b>INE 2023</b>
04	029	904
06	015	903
10	128	902
10	076	903
10	180	904
10	180	905
11	021	903
14	030	901
14	060	902
15	026, 063	902
18	105/6	065
18	020	077
18	020	106
18	158	914
18	105	915
18	140	916
21	017	902
29	015	902
29	084	903
29	084	904
36	011, 012	902
38	013	901
41	095	904
43	148	907



## Anexo C

# *Script* en R para generar el dígito de control de los municipios

Este *script* en R se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Check\\_digit.R](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Check_digit.R)

```
1 #cargar los municipios
2 muni <- read.csv("Muni_ES.csv")
3
4 #Pegar un 0 delante de los identificadores de municipio menores de 5 caracteres
5 muni$IDmuni<-ifelse((nchar(muni$IDmuni))<5,
6                     paste0("0",muni$IDmuni),
7                     muni$IDmuni)
8
9 #comprobar que todos los identificadores de municipio tienen 5 caracteres
10 unique(nchar(muni$IDmuni))
11
12 #Dividir los digitos del identificador del municipio
13 muni$uno <- substr(muni$IDmuni, start = 1, stop = 1)
14 muni$dos <- substr(muni$IDmuni, start = 2, stop = 2)
15 muni$tres <- substr(muni$IDmuni, start = 3, stop = 3)
16 muni$cuatro <- substr(muni$IDmuni, start = 4, stop = 4)
17 muni$cinco <- substr(muni$IDmuni, start = 5, stop = 5)
18
19
20 #Transformar los digitos segun Tabla INE
21 muni$check1 <- ifelse (muni$uno == 0, 0,
22                       ifelse(muni$uno == 1, 2,
23                               ifelse(muni$uno == 2, 4,
24                                       ifelse(muni$uno == 3, 6,
25                                               ifelse(muni$uno == 4, 8,
26                                                       ifelse(muni$uno == 5, 1,
27                                                             ifelse(muni$uno == 6, 3,
28                                                                     ifelse(muni$uno == 7, 5,
29                                                                           ifelse(muni$uno ==8, 7,
30                                                                                   ifelse(muni$uno == 9, 9, NA))))))))))
31
32 muni$check2 <- ifelse (muni$dos == 0, 0,
33                       ifelse(muni$dos == 1, 3,
34                               ifelse(muni$dos == 2, 8,
35                                       ifelse(muni$dos == 3, 2,
36                                               ifelse(muni$dos == 4, 7,
37                                                       ifelse(muni$dos == 5, 4,
38                                                                                   ifelse(muni$dos == 6, 1,
```

```

39         ifelse(muni$dos == 7, 5,
40         ifelse(muni$dos ==8, 9,
41         ifelse(muni$dos == 9, 6, NA)))))))))
42
43 muni$check3 <- ifelse (muni$tres == 0, 0,
44         ifelse(muni$tres == 1, 1,
45         ifelse(muni$tres == 2, 2,
46         ifelse(muni$tres == 3, 3,
47         ifelse(muni$tres == 4, 4,
48         ifelse(muni$tres == 5, 5,
49         ifelse(muni$tres == 6, 6,
50         ifelse(muni$tres == 7, 7,
51         ifelse(muni$tres ==8, 8,
52         ifelse(muni$tres == 9, 9, NA)))))))))
53
54 muni$check4 <- ifelse (muni$cuatro == 0, 0,
55         ifelse(muni$cuatro == 1, 2,
56         ifelse(muni$cuatro == 2, 4,
57         ifelse(muni$cuatro == 3, 6,
58         ifelse(muni$cuatro == 4, 8,
59         ifelse(muni$cuatro == 5, 1,
60         ifelse(muni$cuatro == 6, 3,
61         ifelse(muni$cuatro == 7, 5,
62         ifelse(muni$cuatro ==8, 7,
63         ifelse(muni$cuatro == 9, 9, NA)))))))))
64
65 muni$check5 <- ifelse (muni$cinco == 0, 0,
66         ifelse(muni$cinco == 1, 3,
67         ifelse(muni$cinco == 2, 8,
68         ifelse(muni$cinco == 3, 2,
69         ifelse(muni$cinco == 4, 7,
70         ifelse(muni$cinco == 5, 4,
71         ifelse(muni$cinco == 6, 1,
72         ifelse(muni$cinco == 7, 5,
73         ifelse(muni$cinco ==8, 9,
74         ifelse(muni$cinco == 9, 6, NA)))))))))
75 #sumar los 5 digitos transformados
76 muni$sum <- rowSums(muni[, 16:20])
77
78 #Dividir la suma anterior entre 10 excepto si la suma da 0
79 muni$div <- ifelse(muni$sum ==0, 0, muni$sum/10)
80
81 #Extraer el dividendo de la disvision anterior
82 muni$res <- substr(muni$div, start = 3, stop = 3)
83
84 #Al substraer los numeros pasan a ser caracteres por lo que los volvemos a
    pasar a numeros
85 muni$res <- as.numeric(muni$res)
86
87 #Los que son NA los transformamos en 0
88 muni$res[is.na(muni$res)] <- 0
89
90 #los que son 0 se mantienen como 0 y a los demas se les resta 10
91 muni$dc <- ifelse(muni$res == 0, 0, (10 - muni$res))
92
93 Check_digit_T <- muni[, c(1:10, 24)]
94
95 #Exportar csv con los digitos de control de cada municipio
96 write.csv(Check_digit_T, "Check_digit_Todos.csv")

```

## Anexo D

# Consulta SPARQL-Generate para convertir las geometrías de los municipios asociadas a su código INE de GeoJSON a RDF

Este *script* de SPARQL-Generate se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Conversion\\_municipios\\_RDF/Municipios.rqg](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Conversion_municipios_RDF/Municipios.rqg)

```
1 BASE <https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/>
2 PREFIX iter: <http://w3id.org/sparql-generate/iter/>
3 PREFIX fun: <http://w3id.org/sparql-generate/fn/>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 PREFIX ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/>
6 PREFIX spo: <http://crossforest.eu/position/ontology/>
7 PREFIX epsg: <http://epsg.w3id.org/ontology/>
8 PREFIX crs: <http://crossforest.eu/epsg/data/crs/>
9 PREFIX wkt: <http://www.opengis.net/ont/geosparql#>
10 PREFIX polygon: <https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/
    mfe/polygon/>
11 PREFIX municipality: <https://datos.iepnb.es/recurso/sector-publico/medio-
    ambiente/ifn/municipality>
12
13 GENERATE {
14
15     <municipality/{?IDMUNI}> a ifn:Municipality ;
16         spo:hasPolygon <polygon/{?IDMUNI}-4326> ;
17
18     <polygon/{?IDMUNI}-4326> a spo:Polygon ;
19         epsg:hasCoordinateReferenceSystem crs:4326 ;
20         wkt:asWKT ?geometricCoordinates ;
21
22 }
23 SOURCE <urn:source> AS ?source
24 ITERATOR iter:GeoJSON(?source) AS ?geometricCoordinates ?properties
25 WHERE {
26     BIND(fun:JSONPath(?properties, "$.IDmuni") AS ?idmuni)
27     BIND("{?idmuni}" AS ?IDMUNI)
28     BIND(xsd:integer(?idmuni) AS ?IDMUNI_INT)
29 }
```



## Anexo E

# Despliegue de un contenedor Docker de Virtuoso

Este código se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Asignacion\\_parcela\\_municipio/Configurar\\_Virtuoso.txt](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Asignacion_parcela_municipio/Configurar_Virtuoso.txt)

```
1 #Creacion de una instancia de Virtuoso en WINDOWS
2 docker run #comando principal para ejecutar un contenedor de Docker.
3 --name my_virtddb #establece el nombre del contenedor como "my_virtddb"
4 -d #el contenedor debe ejecutarse en modo demonio (segundo plano)
5 --env DBA_PASSWORD=mysecret #establece la variable de entorno DBA_PASSWORD con
   el valor mysecret, que se utiliza para establecer la contraseña del usuario
   administrador (DBA) de Virtuoso.
6 --publish 1111:1111 #Este puerto se utiliza para conectarse a la consola de
   administración Web de Virtuoso.
7 --publish 8890:8890 #Este puerto se utiliza para conectarse a la interfaz de
   consulta SPARQL.
8 --volume C:\Users\Usuario\Desktop\MASTER\_0_TFM\Datos_Virtuoso:/database #mapea
   la carpeta local Datos_Virtuoso al directorio /database dentro del
   contenedor. Este directorio contiene los datos de la base de datos.
9 --volume C:\Users\Usuario\Desktop\MASTER\_0_TFM\Datos:/datos #mapea la carpeta
   local Datos al directorio /datos dentro del contenedor. Este directorio
   contiene los datos que se van a cargar en la base de datos.
10 openlink/virtuoso-opensource-7:latest #especifica la imagen de Docker que se va
   a utilizar para crear el contenedor
11
12 #Correr el contenedor
13 docker start my_virtddb
14
15 #Acceder al contenedor
16 docker exec -i my_virtddb isql 1111 dba mysecret
17
18 #Cargar todos los archivos Turtle de la carpeta datos en el grafo especificado
19 ld_dir ('/datos', '*.ttl', 'http://crossforest.eu');
20 rdf_loader_run();
21
22 #Listar los datos de la base de datos
23 select * from DB.DBA.load_list;
```



## Anexo F

# Configuración de una instancia en Fuseki con GeoSPARQL

Este código se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Asignacion\\_parcela\\_municipio/Configurar\\_FUSEKI\\_GEOSPARQL.txt](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Asignacion_parcela_municipio/Configurar_FUSEKI_GEOSPARQL.txt)

```
1 java -jar jena-fuseki-geosparql-4.7.0.jar
2
3 -rf Datos\*.ttl #indica que se deben cargar en memoria los archivos con
4   extension ".ttl" (Turtle) que se encuentren en la carpeta "Datos"
5
6 --index_expiry 5000000000000,500000000000,500000000000 #establece los tiempos
7   de expiracion de los indices utilizados por el programa
8
9 -p 3034 #indica el puerto en el que se va a escuchar las solicitudes HTTP
10
11 -t ./geosparql-tdb2 #indica la carpeta en la que se va a guardar la base de
12   datos TDB (triple store)
13
14 -t2 #indica que se deben usar dos bases de datos TDB, una para datos y otra
15   para indices espaciales
16
17 --update #permite realizar operaciones de actualizacion en la base de datos
18
19 --spatial_index ./geosparql-tdb2/spatial.index #indica la ubicacion de los
20   indices espaciales que se van a utilizar.
21
22 -d Municipios #Nombre de la base de datos espacial, en este caso Municipios
```



## Anexo G

# Consulta SPARQL-Generate para convertir la asignación realizada mediante QGIS de las parcelas a los municipios de GeoJSON a RDF

Este *script* de SPARQL-Generate se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Asignacion\\_parcela\\_municipio/Asig\\_muni\\_plot.rqg](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Asignacion_parcela_municipio/Asig_muni_plot.rqg)

```
1 BASE <https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/>
2 PREFIX iter: <http://w3id.org/sparql-generate/iter/>
3 PREFIX fun: <http://w3id.org/sparql-generate/fn/>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 PREFIX ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/>
6 PREFIX municipality: <https://datos.iepnb.es/recurso/sector-publico/medio-
  ambiente/ifn/municipality>
7 PREFIX plot: <https://datos.iepnb.es/recurso/sector-publico/medio-ambiente/ifn/
  plot>
8
9
10 GENERATE {
11     <plot/{?IDPLOT}> ifn:isInMunicipality <municipality/{?IDMUNI}> ;
12     .
13 }
14
15 }
16
17 SOURCE <urn:source> AS ?source
18 ITERATOR iter:GeoJSON(?source) AS ?geometry ?properties
19 WHERE {
20     BIND(fun:JSONPath(?properties, "$.plot") AS ?idPlot)
21     BIND(STRAFTER(?idPlot, "plot/") AS ?IDPLOT)
22     BIND(xsd:integer(?idPlot) AS ?IDPLOT_INT)
23     BIND(fun:JSONPath(?properties, "$.IDmuni") AS ?idmuni)
24     BIND("{?idmuni}" AS ?IDMUNI)
25     BIND(xsd:integer(?idmuni) AS ?IDMUNI_INT)
26 }
```



## Anexo H

# Asignación parcelas-municipios con Virtuoso

Este *Script* de shell se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Asignacion\\_parcela\\_municipio/asig\\_plot\\_muni\\_VIRTUOSO.sh](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Asignacion_parcela_municipio/asig_plot_muni_VIRTUOSO.sh)

```
1 curl -X POST http://localhost:8890/sparql --data-urlencode #solicitud POST a
  un servidor SPARQL (localhost)
2
3 "query=
4 PREFIX ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/>
5 PREFIX spo: <http://crossforest.eu/position/ontology/>
6 PREFIX crs: <http://epsg.w3id.org/data/crs/>
7 PREFIX axis: <http://epsg.w3id.org/ontology/axis/>
8 PREFIX wkt: <http://www.opengis.net/ont/geosparql#>
9
10 CONSTRUCT {?plot ifn:isInMunicipality ?muni}
11 WHERE {
12     ?plot a ifn:Plot;
13         spo:hasPosition ?pos.
14     ?pos spo:hasCoordinateReferenceSystem crs:4326 ;
15         axis:106 ?lat ;
16         axis:107 ?lng .
17     ?muni spo:hasPolygon ?poly .
18     ?poly wkt:asWKT ?wkt_muni .
19     FILTER (bif:st_within(bif:st_point(xsd:double(?lng), xsd:double(?lat)), ?
  wkt_muni))
20 }"
21
22 --data-urlencode "default-graph-uri=http://crossforest.eu" #IRI grafo con la
  base de datos
23
24 -H 'Accept:application/x-turtle' #Resultados generados en formato Turtle
25
26 > muni_plot_Virtuoso.ttl #archivo resultante
```



## Anexo I

# Asignación parcelas-municipios con Fuseki

Este *Script* de shell se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Asignacion\\_parcela\\_municipio/asig\\_plot\\_muni\\_FUSEKI.sh](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Asignacion_parcela_municipio/asig_plot_muni_FUSEKI.sh)

```
1 curl -X POST --data #solicitud POST a un servidor SPARQL
2
3 "query=
4 PREFIX spatialF: <http://jena.apache.org/function/spatial#>
5 PREFIX wkt: <http://www.opengis.net/ont/geosparql#>
6 PREFIX spo: <http://crossforest.eu/position/ontology/>
7 PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
8 PREFIX ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/>
9 PREFIX crs: <http://epsg.w3id.org/data/crs/>
10 PREFIX axis: <http://epsg.w3id.org/ontology/axis/>
11
12 CONSTRUCT {?plot ifn:isInMunicipality ?muni}
13 WHERE {
14     ?plot a ifn:Plot; spo:hasPosition ?spo.
15     ?spo spo:hasCoordinateReferenceSystem crs:4326;
16         axis:106 ?lat;
17         axis:107 ?lng .
18     ?muni spo:hasPolygon ?poly .
19     ?poly wkt:asWKT ?wkt_muni.
20     BIND(spatialF:convertLatLon(?lat, ?lng) as ?point)
21     FILTER(geof:sfWithin(?point, ?wkt_muni)).
22 }"
23
24 -H 'Accept:application/x-turtle' #Resultados generados en formato Turtle
25
26 http://localhost:3034/Municipios/query #Servidor SPARQL de Fuseki configurado
    anteriormente
27
28 > Asig_plot_N.ttl #archivo resultante
```



## Anexo J

# Asignación parcelas-municipios con QGIS

Este *Script* de shell se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Asignacion\\_parcela\\_municipio/asig\\_plot\\_muni\\_QGIS.sh](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Asignacion_parcela_municipio/asig_plot_muni_QGIS.sh)

```
1 #Descarga de los datos de las posiciones de las parcelas de punto SPARQL de
   pruebas
2 curl -k -G https://crossforest.gsic.uva.es/pruebas/sparql --data-urlencode
   query='
3 PREFIX ifn: <https://datos.iepnb.es/def/sector-publico/medio-ambiente/ifn/>
4 PREFIX spo: <http://crossforest.eu/position/ontology/>
5 SELECT ?plot ?lat ?lng
6 WHERE {
7   ?plot a ifn:Plot ;
8     spo:hasPosition ?pos .
9   ?pos spo:hasCoordinateReferenceSystem <http://epsg.w3id.org/data/crs/4326> ;
10     <http://epsg.w3id.org/ontology/axis/106> ?lat ;
11     <http://epsg.w3id.org/ontology/axis/107> ?lng .
12 }' -H "Accept: text/csv" > plots.csv
13
14 #Este comando utiliza la herramienta ogr2ogr para convertir el archivo CSV de
   las parcelas a un archivo shapefile
15 ogr2ogr -f "ESRI Shapefile" #La opcion -f se utiliza para especificar el
   formato de salida
16 -oo X_POSSIBLE_NAMES=lng #se utilizan para indicar que columna del archivo csv
   contiene las coordenadas X (longitud)
17 -oo Y_POSSIBLE_NAMES=lat #se utilizan para indicar que columna del archivo csv
   contiene las coordenadas Y (latitud)
18 -a_srs EPSG:4326 #especificar el sistema de coordenadas de referencia
19 plots_IFN3.shp plots.csv
20
21 #Este comando utiliza la herramienta ogrinfo para crear un indice espacial en
   el archivo shapefile plots_IFN3.shp y recinto_muni.shp.
22 ogrinfo -sql "CREATE SPATIAL INDEX ON plots_IFN3" plots_IFN3.shp
23 ogrinfo -sql "CREATE SPATIAL INDEX ON recintos_muni" recintos_muni.shp
24
25 #Utiliza la herramienta qgis_process para ejecutar la operacion "join
   attributes by location" en QGIS que da como resultado un archivo Shapefile.
26 qgis_process run native:joinattributesbylocation
27 --INPUT=plots_IFN3.shp
28 --JOIN=recintos_muni.shp
```

```
29 --PREDICATE=5 #es el tipo de relación espacial que se usará para unir las
    capas. En este caso, se usa el valor 5 que representa la relación "within"
    (dentro de).
30 --JOIN_FIELDS="" #se deja en blanco porque se van a incluir todos los campos.
31 --METHOD=0 #el valor 0 significa "tomar los atributos del objeto de la capa
    de unión".
32 --DISCARD_NONMATCHING=0 #se deja en 0 para que se incluyan en la capa de
    salida los elementos de la capa de entrada que no tienen una correspondencia
    en la capa de unión
33 --PREFIX="" #especifica un prefijo que se le agregará a los nombres de los
    campos de la capa de unión en la capa de salida.
34 --OUTPUT=Asig_plot_muni.shp #nombre de la capa de salida
35
36 #Transforma el shapefile de la asignación a GeoJSON
37 ogr2ogr -f GeoJSON Asig_plot_muni.json Asig_plot_muni.shp
38
39 #Utiliza el SPARQL-Generate para transformar el GeoJSON a RDF
40 /home/natalia/QGIS_automatiko/sparql-generate.sh -q /home/natalia/QGIS_
    automatiko/Asig_muni_plot.rqg -i "/home/natalia/QGIS_automatiko/Asig_plot_
    muni.json" -p -Xmx10000M -v
```

## Anexo K

# *Script* en R para la comparación de asignaciones de parcelas del IFN3 a municipios

Este *script* en R se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Compara\\_muni\\_inventario.R](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Compara_muni_inventario.R)

```
1 #####
2 ##### Librerías #####
3 #####
4
5 #install.packages("sqldf")
6 library(sqldf) #Consultas SQL
7
8 #install.packages("dplyr")
9 library(dplyr) #Manipulación de datos
10
11 #####
12 ##### Datos asignación Fuseki #####
13 #####
14
15 #Cargar datos
16 Asig_FUSEKI <- read.csv("Plot_muni_TODOS.csv")
17
18 #Substraer de las IRI de Parcela y Municipio los códigos de Provincia,
19   Estadillo, Clase, Subclase y Código INE
20 Asig_FUSEKI$Provincia <- substr(Asig_FUSEKI$plot, start = 71, stop = 72)
21 Asig_FUSEKI$Estadillo <- substr(Asig_FUSEKI$plot, start = 74, stop = 77)
22 Asig_FUSEKI$Clase_PCP <- substr(Asig_FUSEKI$plot, start = 79, stop = 79)
23 Asig_FUSEKI$Subclase <- substr(Asig_FUSEKI$plot, start = 81, stop = 83)
24 Asig_FUSEKI$INE <- substr(Asig_FUSEKI$muni, start = 79, stop = 85)
25 Asig_FUSEKI$INE3 <- substr(Asig_FUSEKI$INE, start = 3, stop = 5)
26
27 #####
28 ##### Datos asignación QGIS #####
29 #####
30 #Cargar datos
31 Asig_QGIS <- read.csv("Plot_muni_QGIS.csv")
32
33 #Substraer columnas que me interesan y cambiar nombre
34 Asig_QGIS$Provincia <- substr(Asig_QGIS$plot, start = 71, stop = 72)
```

```

35 Asig_QGIS$Estadillo <- substr(Asig_QGIS$plot, start = 74, stop = 77)
36 Asig_QGIS$Clase_PCPPar <- substr(Asig_QGIS$plot, start = 79, stop = 79)
37 Asig_QGIS$Subclase <- substr(Asig_QGIS$plot, start = 81, stop = 83)
38
39 Asig_QGIS$IDmuni<-ifelse((nchar(Asig_QGIS$IDmuni))<5,
40                          paste0("0", Asig_QGIS$IDmuni),
41                          Asig_QGIS$IDmuni)
42
43 Asig_QGIS <- Asig_QGIS[, c(14:17, 13)]
44
45 names(Asig_QGIS) <- c("Provincia", "Estadillo", "Clase_PCPPar", "Subclase",
46                      "INE_QGIS")
47
48 #####
49 ##### Comparar asignaciones FUSEKI y QGIS #####
50 #####
51
52 #Unir las dos Tablas
53 Fuseki_QGIS <- merge(Asig_FUSEKI, Asig_QGIS, by= c("Provincia", "Estadillo", "
54             Clase_PCPPar", "Subclase"))
55
56 #Comparar código INE asignado por FUSEKI y por QGIS
57 Comp_F_Q <- subset(Fuseki_QGIS, Fuseki_QGIS$INE == Fuseki_QGIS$INE_QGIS) #100%
58 coincidencias
59
60 #####
61 ##### Datos Tabla IFN3 PCDatosMap #####
62 #####
63
64 #Cargar datos
65 PCDatosMap <- read.csv("pcdatosmap.csv")
66
67 #Extraer solo columnas que me interesan
68 PCDatosMap <- PCDatosMap[, c(2:4, 7, 8, 9)]
69
70 #Homogeneizar datos Provincia, Estadillo y Código INE
71 PCDatosMap$Provincia<-ifelse((nchar(PCDatosMap$Provincia))<2,
72                               paste0("0", PCDatosMap$Provincia),
73                               PCDatosMap$Provincia)
74
75 PCDatosMap$Estadillo<-ifelse((nchar(PCDatosMap$Estadillo))<4,
76                               paste0("0", PCDatosMap$Estadillo),
77                               PCDatosMap$Estadillo)
78
79 PCDatosMap$Estadillo<-ifelse((nchar(PCDatosMap$Estadillo))<4,
80                               paste0("0", PCDatosMap$Estadillo),
81                               PCDatosMap$Estadillo)
82
83 PCDatosMap$INE3<-ifelse( nchar(PCDatosMap$INE)==5,
84                          substr(PCDatosMap$INE, start = 3, stop = 5),
85                          PCDatosMap$INE)
86
87 PCDatosMap$INE3<-ifelse ( nchar(PCDatosMap$INE) ==4,
88                          substr(PCDatosMap$INE, start = 2, stop = 4),
89                          PCDatosMap$INE3)
90
91 PCDatosMap$INE3<-ifelse ( nchar(PCDatosMap$INE)==2,
92                          paste0("0", PCDatosMap$INE) ,
93                          PCDatosMap$INE3)
94
95 PCDatosMap$INE3<-ifelse ( nchar(PCDatosMap$INE) ==1,
96                          paste0("00", PCDatosMap$INE),
97                          PCDatosMap$INE3)

```

```

93
94 #Número de Parcelas con código INE de municipio
95 sum(!is.na(PCDatosMap$INE3) & PCDatosMap$INE3 != "000") #77 212 tienen código
  INE
96 sum(is.na(PCDatosMap$INE3)) #19 184 NA's en código municipio
97
98 #Corrección de discrepancias
99 PCDatosMap[c(32966, 61044, 61070), 3] <- "A"
100
101 #####
102 ##### Datos Tabla IFN3 PCParcelas #####
103 #####
104
105 #Cargar datos
106 PCParcelas <- read.csv("pcparcelas.csv")
107
108 #Extraer columnas que me interesan
109 PCParcelas <- PCParcelas[,c(2:5, 6, 7, 16)]
110
111 #Cambiar nombre columnas par que coincidan con la otra Tabla
112 names(PCParcelas) <- c("Provincia", "Estadillo", "Clase", "Subclase", "CoorX",
  "CoorY", "INE")
113
114 #Homogeneizar códigos de Provincia, Estadillo y Código INE
115 PCParcelas$INE <- as.numeric(PCParcelas$INE)
116 PCParcelas$Provincia<-ifelse((nchar(PCParcelas$Provincia))<2,
  paste0("0", PCParcelas$Provincia),
  PCParcelas$Provincia)
117
118 PCParcelas$Estadillo<-ifelse((nchar(PCParcelas$Estadillo))<4,
  paste0("0", PCParcelas$Estadillo),
  PCParcelas$Estadillo)
119
120 PCParcelas$Estadillo<-ifelse((nchar(PCParcelas$Estadillo))<4,
  paste0("0", PCParcelas$Estadillo),
  PCParcelas$Estadillo)
121
122 PCParcelas$Estadillo<-ifelse((nchar(PCParcelas$Estadillo))<4,
  paste0("0", PCParcelas$Estadillo),
  PCParcelas$Estadillo)
123
124 PCParcelas$Estadillo<-ifelse((nchar(PCParcelas$Estadillo))<4,
  paste0("0", PCParcelas$Estadillo),
  PCParcelas$Estadillo)
125
126 PCParcelas$INE3<-ifelse ( nchar(PCParcelas$INE) == 4, substr(PCParcelas$INE,
  start = 2, stop = 4), PCParcelas$INE)
127
128 PCParcelas$INE3<-ifelse ( nchar(PCParcelas$INE)==2, paste0("0", PCParcelas$INE)
  , PCParcelas$INE3)
129
130 PCParcelas$INE3<-ifelse ( nchar(PCParcelas$INE) ==1, paste0("00", PCParcelas$
  INE), PCParcelas$INE3)
131
132 #Número de parcelas con código INE de municipio
133 sum(!is.na(PCParcelas$INE) & PCParcelas$INE3 != "000") #46 443 tienen código
  INE
134 sum((is.na(PCParcelas$INE) | PCParcelas$INE3 == "000")) #52 603 no tienen có
  digo INE
135
136 #Corrección de errores
137 PCParcelas$Estadillo <- ifelse(PCParcelas$Estadillo == 9999, 999, PCParcelas$
  Estadillo)
138
139 #####
140 ##### Unión de las dos Tablas del IFN3 y comparación Código INE #####
141 #####
142
143 #Unión ambas Tablas por las columnas de Provincia y Estadillo
144 Parcelas_IFN3 <- merge(PCDatosMap, PCParcelas, by= c("Provincia", "Estadillo"))

```

```

145 summary(Parcelas_IFN3)
146
147 #cuantas parcelas tienen dato de INE y que ese dato no sea 000
148 sum(!is.na(Parcelas_IFN3$INE3.x) & Parcelas_IFN3$INE3.x != "000") #80.386
    tienen código INE en PCDatosMap
149 sum(!is.na(Parcelas_IFN3$INE3.y) & Parcelas_IFN3$INE3.y != "000") #46.443
    tienen código INE en PCParcelas
150
151 #Cambiar el nombre de las columnas para entender la Tabla mejor
152 names(Parcelas_IFN3) <- c("Provincia", "Estadillo", "Clase_PcMap", "CoorX_
    PcMap", "CoorY_PcMap", "INE_PcMap", "INE3_PcMap", "Clase_PcPar",
153     "Subclase", "CoorX_PcPar", "CoorY_PcPar", "INE_PcPar",
    "INE3_PcPar")
154
155
156 #Parcelas con mismo código INE en PCDatosMap y PCParcelas
157 INE_igual <- subset(Parcelas_IFN3, ((Parcelas_IFN3$INE3_PcMap == Parcelas_IFN3$
    INE3_PcPar) & Parcelas_IFN3$INE3_PcPar != "000")) #45 670 iguales
158
159 #Parcelas con distinto código INE en PCDatosMap y PCParcelas
160 INE_dist <- subset(Parcelas_IFN3, Parcelas_IFN3$INE3_PcMap != Parcelas_IFN3$
    INE3_PcPar) #773 distintas
161
162 #Parcelas sin código INE en algunos de los dos
163 sum(is.na(Parcelas_IFN3$INE3_PcMap) | is.na(Parcelas_IFN3$INE3_PcPar)) #52 783
164
165 #####
166 ##### Comparar Código INE de las parcelas del IFN3 #####
167 ##### con la asignación realizada con FUSEKI #####
168 #####
169
170 #Unir Tabla Parcelas_IFN3 con Asig_FUSEKI
171 merge_Pcpar <- merge(Parcelas_IFN3, Asig_FUSEKI, by= c("Provincia", "Estadillo"
    , "Clase_PcPar", "Subclase"))
172
173 #Actualizar código INE PcMap
174 merge_Pcpar$INE3_PcMap <- ifelse((merge_Pcpar$INE3_PcMap ==128 &
175     merge_Pcpar$INE ==10902), 902,
176     ifelse((merge_Pcpar$INE3_PcMap ==180 &
177     merge_Pcpar$INE == 10904), 904,
178     ifelse((merge_Pcpar$INE3_PcMap ==180 &
179     merge_Pcpar$INE == 10905), 905,
180     ifelse((merge_Pcpar$INE3_PcMap == "026" |
181     merge_Pcpar$INE3_PcMap == "063" )
182     & merge_Pcpar$INE == 15902), 902,
183     ifelse((merge_Pcpar$INE3_PcMap == "011" |
184     merge_Pcpar$INE3_PcMap == "012" )
185     & merge_Pcpar$INE == 36902), 902,
186     ifelse((merge_Pcpar$INE3_PcMap == "013" &
187     merge_Pcpar$INE == 38901), 901,
188     ifelse((merge_Pcpar$INE3_PcMap == "015" &
189     merge_Pcpar$INE == "06902"), 902,
190     ifelse((merge_Pcpar$INE3_PcMap == "060" &
191     merge_Pcpar$INE == 48915), 915,
192     ifelse((merge_Pcpar$INE3_PcMap == 105 &
193     merge_Pcpar$INE == 18915), 915,
194     merge_Pcpar$INE3_PcMap)))))))))
195
196 #Actualizar Código INE PcPar
197 merge_Pcpar$INE3_PcPar <- ifelse((merge_Pcpar$INE3_PcPar ==128 &

```

```

198         merge_PCpar$INE ==10902), 902,
199         ifelse((merge_PCpar$INE3_PCPar ==180 &
200             merge_PCpar$INE == 10904), 904,
201         ifelse((merge_PCpar$INE3_PCPar ==180 &
202             merge_PCpar$INE == 10905), 905,
203         ifelse(((merge_PCpar$INE3_PCPar == "026" |
204             merge_PCpar$INE3_PCPar == "063" )
205             & merge_PCpar$INE == 15902), 902,
206         ifelse(((merge_PCpar$INE3_PCPar == "011" |
207             merge_PCpar$INE3_PCPar == "012" ) &
208             merge_PCpar$INE == 36902), 902,
209         ifelse((merge_PCpar$INE3_PCPar == "013" &
210             merge_PCpar$INE == 38901), 901,
211         ifelse((merge_PCpar$INE3_PCPar == "015" &
212             merge_PCpar$INE == "06902"), 902,
213         ifelse((merge_PCpar$INE3_PCPar == "060" &
214             merge_PCpar$INE == 48915), 915,
215         ifelse((merge_PCpar$INE3_PCPar == 105 &
216             merge_PCpar$INE == 18915), 915,
217             merge_PCpar$INE3_PCPar)))))))))
218
219 #Comparar Código INE con datos de PCDatosMap
220 Asig_PCMap <- subset(merge_PCpar,
221     merge_PCpar$INE3 == merge_PCpar$INE3_PCMap) #44 910
222     iguales
223 Asig_PCMap_d <- subset(merge_PCpar,
224     merge_PCpar$INE3 != merge_PCpar$INE3_PCMap) #35 376
225     distintos
226
227 #Comparar Código INE con datos de PCParcelas
228 Asig_PCPar <- subset(merge_PCpar,
229     merge_PCpar$INE3 == merge_PCpar$INE3_PCPar) #44 910
230     iguales
231 Asig_PCPar_d <- subset(merge_PCpar,
232     merge_PCpar$INE3 != merge_PCpar$INE3_PCPar) #5 369
233     distintos

```



## Anexo L

# ***Script* de R para asociar de la especie al valor máximo de área basimétrica y el Número de pies/ha por municipio**

Este *script* de R se encuentra disponible en: [https://github.com/nacresle/GeoSPARQL\\_TFM/blob/main/Scripts/Mapas\\_SP\\_dominante.R](https://github.com/nacresle/GeoSPARQL_TFM/blob/main/Scripts/Mapas_SP_dominante.R)

```
1 #Librerías utilizadas
2 library(dplyr)
3
4 #####
5 ##### Especies dominantes numero pies por ha #####
6 #####
7
8
9 #Cargar datos
10 pies_ha <- read.csv("N.csv")
11
12 #Crear columna del código INE extrayendolo del IRI del municipio
13 pies_ha$IDmuni <- substr(pies_ha$muni, start = 79, stop = 83)
14
15 #Crear columna del código de la especie extrayendolo del IRI de la especie
16 pies_ha$SP <- substr(pies_ha$especies, start = 69, stop = 70)
17
18 #Sustituir código de especie por nombre científico
19 pies_ha$Nombre_SP <- ifelse(pies_ha$SP == 7, "Acacia spp.",
20                             ifelse(pies_ha$SP == 8, "Phillyrea latifolia",
21                                     ifelse(pies_ha$SP == 15, "Crataegus spp.",
22                                             ifelse(pies_ha$SP == 18, "Chamaecyparis lawsoniana",
23                                                     ifelse(pies_ha$SP == 19, "Otras coníferas",
24                                                         ifelse(pies_ha$SP == 21, "Pinus sylvestris",
25                                                                 ifelse(pies_ha$SP == 22, "Pinus uncinata",
26                                                                     ifelse(pies_ha$SP == 23, "Pinus pinea",
27                                                                                             ifelse(pies_ha$SP == 24, "Pinus halepensis",
28                                                                                                 ifelse(pies_ha$SP == 25, "Pinus nigra",
29                                                                                                     ifelse(pies_ha$SP == 26, "Pinus pinaster",
30                                                                                                         ifelse(pies_ha$SP == 27, "Pinus canariensis",
31                                                                                                             ifelse(pies_ha$SP == 28, "Pinus radiata",
32                                                                                                                 ifelse(pies_ha$SP == 29, "Otros pinos",
33                                                                                                                     ifelse(pies_ha$SP == 31, "Abies alba",
34                                                                                                                         ifelse(pies_ha$SP == 32, "Larix spp.",
35                                                                                                                             ifelse(pies_ha$SP == 34, "Pseudotsuga menziesii",
```

```

36         ifelse(pies_ha$SP == 35, "Otros pinos",
37         ifelse(pies_ha$SP == 36, "Cupressus sempervirens",
38         ifelse(pies_ha$SP == 38, "Juniperus thurifera",
39         ifelse(pies_ha$SP == 39, "Juniperus phoenicea",
40         ifelse(pies_ha$SP == 41, "Quercus robur",
41         ifelse(pies_ha$SP == 42, "Quercus petraea",
42         ifelse(pies_ha$SP == 43, "Quercus pyrenaica",
43         ifelse(pies_ha$SP == 44, "Quercus faginea",
44         ifelse(pies_ha$SP == 45, "Quercus ilex",
45         ifelse(pies_ha$SP == 46, "Quercus suber",
46         ifelse(pies_ha$SP == 47, "Quercus canariensis",
47         ifelse(pies_ha$SP == 48, "Quercus rubra",
48         ifelse(pies_ha$SP == 50, "Mezcla de árboles de ribera",
49         ifelse(pies_ha$SP == 51, "Populus alba",
50         ifelse(pies_ha$SP == 53, "Tamarix spp.",
51         ifelse(pies_ha$SP == 54, "Alnus glutinosa",
52         ifelse(pies_ha$SP == 55, "Fraxinus angustifolia",
53         ifelse(pies_ha$SP == 57, "Salix spp.",
54         ifelse(pies_ha$SP == 58, "Populus nigra",
55         pies_ha$SP))))))))))))))))))))))))))))))))))))))
56
57 pies_ha$Nombre_SP <- ifelse(pies_ha$SP == 60, "Mezcla de eucaliptos",
58         ifelse(pies_ha$SP == 61, "Eucalyptus globulus",
59         ifelse(pies_ha$SP == 62, "Eucalyptus camaldulensis",
60         ifelse(pies_ha$SP == 64, "Eucalyptus nitens",
61         ifelse(pies_ha$SP == 65, "Ilex aquifolium",
62         ifelse(pies_ha$SP == 66, "Olea europaea",
63         ifelse(pies_ha$SP == 67, "Ceratonia siliqua",
64         ifelse(pies_ha$SP == 68, "Arbutus unedo",
65         ifelse(pies_ha$SP == 70, "Mezcla de frondosas de gran
66         porte",
67         ifelse(pies_ha$SP == 71, "Fagus sylvatica",
68         ifelse(pies_ha$SP == 72, "Castanea sativa",
69         ifelse(pies_ha$SP == 73, "Betula spp.",
70         ifelse(pies_ha$SP == 74, "Corylus avellana",
71         ifelse(pies_ha$SP == 76, "Acer campestre",
72         ifelse(pies_ha$SP == 77, "Tilia spp.",
73         ifelse(pies_ha$SP == 78, "Sorbus spp.",
74         ifelse(pies_ha$SP == 81, "Myrica faya",
75         ifelse(pies_ha$SP == 82, "Ilex canariensis",
76         ifelse(pies_ha$SP == 83, "Erica arborea",
77         ifelse(pies_ha$SP == 83, "Erica arborea",
78         ifelse(pies_ha$SP == 84, "Persea indica",
79         ifelse(pies_ha$SP == 89, "Otras laurisilvas",
80         ifelse(pies_ha$SP == 90, "Mezcla de pequeñas frondosas",
81         ifelse(pies_ha$SP == 91, "Buxus sempervirens",
82         ifelse(pies_ha$SP == 92, "Robinia pseudacacia",
83         ifelse(pies_ha$SP == 93, "Pistacia terebinthus",
84         ifelse(pies_ha$SP == 95, "Prunus spp.",
85         ifelse(pies_ha$SP == 97, "Sambucus nigra",
86         ifelse(pies_ha$SP == 98, "Carpinus betulus",
87         ifelse(pies_ha$SP == 99, "Otras frondosas",
88         pies_ha$Nombre_SP))))))))))))))))))))))))))))))))))
89 #Quedarnos con las columnas del Código del municipio, la media de N y el código
90     y
91 #el nombre científico de la especie
92 pies_ha2 <-pies_ha[, c(1, 4,5,3)]
93 #Crear otra Tabla con la especie dominante por municipio

```

```

94 pies_ha2_max<- pies_ha2 %>%
95   group_by(IDmuni) %>%
96   top_n(n = 1, wt = meanN)
97
98 #Guardar Tabla resultado en un csv
99 write.csv(pies_ha2, "pies_ha_max.csv")
100
101 #####
102 ##### Especies dominantes área basimétrica #####
103 #####
104
105 #Cargar datos
106 areaB <- read.csv("G.csv")
107
108 #Crear columna del código INE extrayendolo del IRI del municipio
109 areaB$IDmuni <- substr(areaB$muni, start = 79, stop = 83)
110
111 #Crear columna del código de la especie extrayendolo del IRI de la especie
112 areaB$SP <- substr(areaB$species, start = 69, stop = 70)
113
114 #Sustituir código de especie por nombre científico
115 areaB$Nombre_SP <- ifelse(areaB$SP == 7, "Acacia spp.",
116   ifelse(areaB$SP == 8, "Phillyrea latifolia",
117   ifelse(areaB$SP == 15, "Crataegus spp.",
118   ifelse(areaB$SP == 18, "Chamaecyparis lawsoniana",
119   ifelse(areaB$SP == 19, "Otras coníferas",
120   ifelse(areaB$SP == 21, "Pinus sylvestris",
121   ifelse(areaB$SP == 22, "Pinus uncinata",
122   ifelse(areaB$SP == 23, "Pinus pinea",
123   ifelse(areaB$SP == 24, "Pinus halepensis",
124   ifelse(areaB$SP == 25, "Pinus nigra",
125   ifelse(areaB$SP == 26, "Pinus pinaster",
126   ifelse(areaB$SP == 27, "Pinus canariensis",
127   ifelse(areaB$SP == 28, "Pinus radiata",
128   ifelse(areaB$SP == 29, "Otros pinos",
129   ifelse(areaB$SP == 31, "Abies alba",
130   ifelse(areaB$SP == 32, "Larix spp.",
131   ifelse(areaB$SP == 34, "Pseudotsuga menziesii",
132   ifelse(areaB$SP == 35, "Otros pinos",
133   ifelse(areaB$SP == 36, "Cupressus sempervirens",
134   ifelse(areaB$SP == 38, "Juniperus thurifera",
135   ifelse(areaB$SP == 39, "Juniperus phoenicea",
136   ifelse(areaB$SP == 41, "Quercus robur",
137   ifelse(areaB$SP == 42, "Quercus petraea",
138   ifelse(areaB$SP == 43, "Quercus pyrenaica",
139   ifelse(areaB$SP == 44, "Quercus faginea",
140   ifelse(areaB$SP == 45, "Quercus ilex",
141   ifelse(areaB$SP == 46, "Quercus suber",
142   ifelse(areaB$SP == 47, "Quercus canariensis",
143   ifelse(areaB$SP == 48, "Quercus rubra",
144   ifelse(areaB$SP == 50, "Mezcla de árboles de ribera",
145   ifelse(areaB$SP == 51, "Populus alba",
146   ifelse(areaB$SP == 53, "Tamarix spp.",
147   ifelse(areaB$SP == 54, "Alnus glutinosa",
148   ifelse(areaB$SP == 55, "Fraxinus angustifolia",
149   ifelse(areaB$SP == 57, "Salix spp.",
150   ifelse(areaB$SP == 58, "Populus nigra",
151   areaB$SP))))))))))))))))))))))))))))))))))))))
152
153 areaB2$Nombre_SP <- ifelse(areaB$SP == 60, "Mezcla de eucaliptos",

```

```

154         ifelse(areaB$SP == 61, "Eucalyptus globulus",
155         ifelse(areaB$SP == 62, "Eucalyptus camaldulensis",
156         ifelse(areaB$SP == 64, "Eucalyptus nitens",
157         ifelse(areaB$SP == 65, "Ilex aquifolium",
158         ifelse(areaB$SP == 66, "Olea europaea",
159         ifelse(areaB$SP == 67, "Ceratonia siliqua",
160         ifelse(areaB$SP == 68, "Arbutus unedo",
161         ifelse(areaB$SP == 70, "Mezcla de frondosas de gran porte"
162     ,
163         ifelse(areaB$SP == 71, "Fagus sylvatica",
164         ifelse(areaB$SP == 72, "Castanea sativa",
165         ifelse(areaB$SP == 73, "Betula spp.",
166         ifelse(areaB$SP == 74, "Corylus avellana",
167         ifelse(areaB$SP == 76, "Acer campestre",
168         ifelse(areaB$SP == 77, "Tilia spp.",
169         ifelse(areaB$SP == 78, "Sorbus spp.",
170         ifelse(areaB$SP == 81, "Myrica faya",
171         ifelse(areaB$SP == 82, "Ilex canariensis",
172         ifelse(areaB$SP == 83, "Erica arborea",
173         ifelse(areaB$SP == 84, "Persea indica",
174         ifelse(areaB$SP == 89, "Otras laurisilvas",
175         ifelse(areaB$SP == 90, "Mezcla de pequenas frondosas",
176         ifelse(areaB$SP == 91, "Buxus sempervirens",
177         ifelse(areaB$SP == 92, "Robinia pseudacacia",
178         ifelse(areaB$SP == 93, "Pistacia terebinthus",
179         ifelse(areaB$SP == 95, "Prunus spp.",
180         ifelse(areaB$SP == 97, "Sambucus nigra",
181         ifelse(areaB$SP == 98, "Carpinus betulus",
182         ifelse(areaB$SP == 99, "Otras frondosas",
183         areaB$Nombre_SP))))))))))))))))))))))))))))))
184
185 #Quedarnos con las columnas del Código del municipio la media de N y la especie
186     (Código
187     # y nombre científico)
188     areaB2 <-areaB[, c(1, 4,5,3)]
189
190 #Crear otra Tabla con la especie dominante por municipio
191     areaB2_max<- areaB2 %>%
192     group_by(IDmuni) %>%
193     top_n(n = 1, wt = meanG)
194
195 #Guardar Tabla resultado en un csv
196     write.csv(areaB2_max, "area_basimetrica_max.csv")

```

## Anexo M

# Mapas de especies dominantes de los Parques Nacionales Españoles

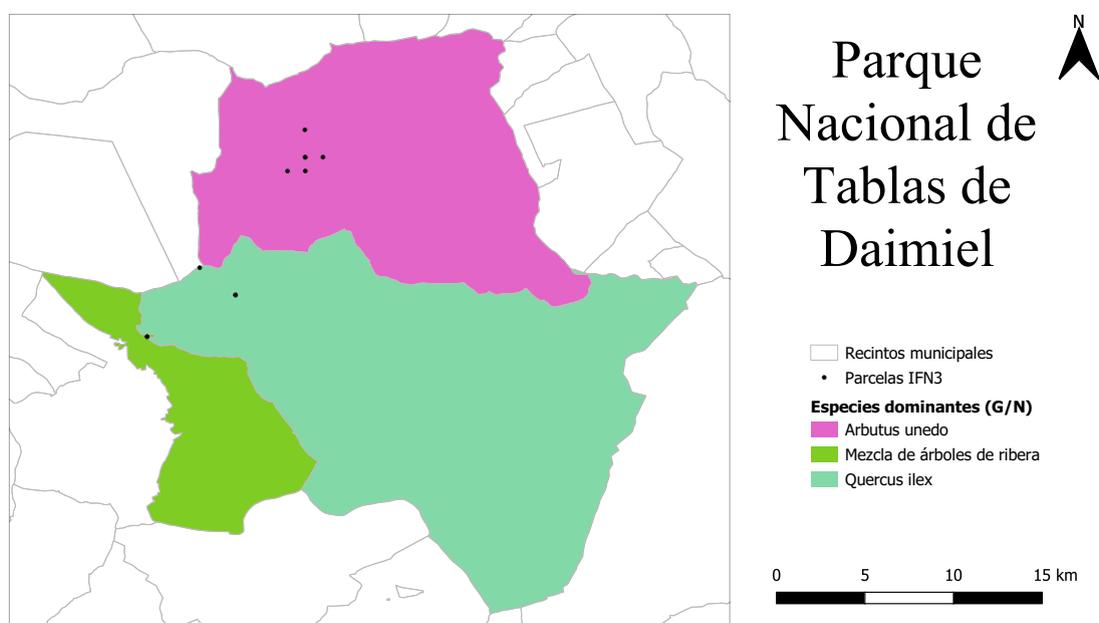


Figura M.1: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y el número de pies/ha y parcelas del IFN3 del Parque Nacional de Tablas de Daimiel

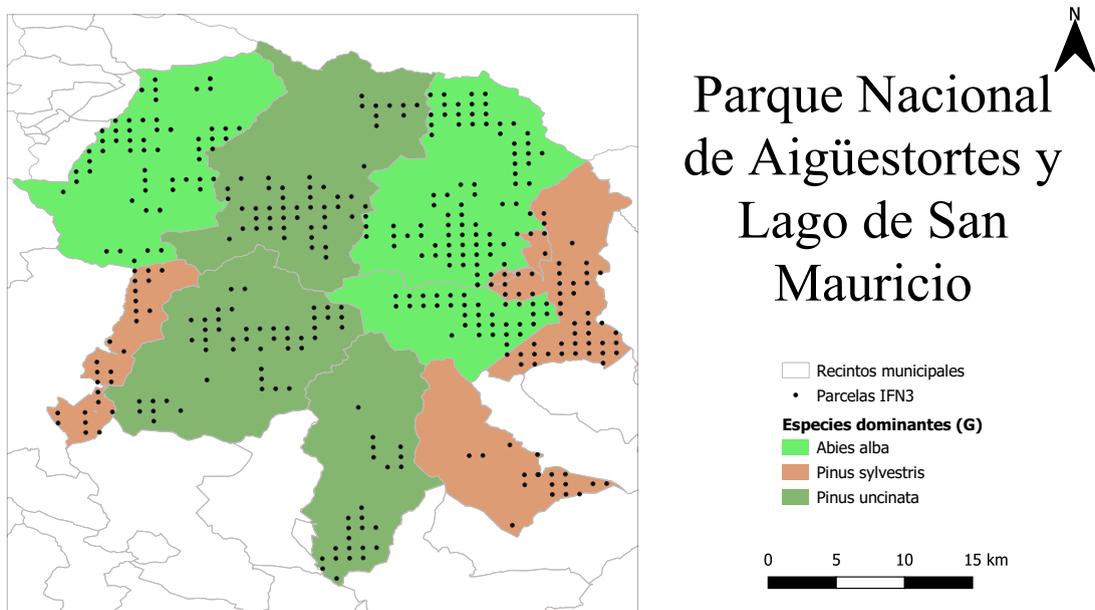


Figura M.2: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Aigüestortes y Lago de San Mauricio

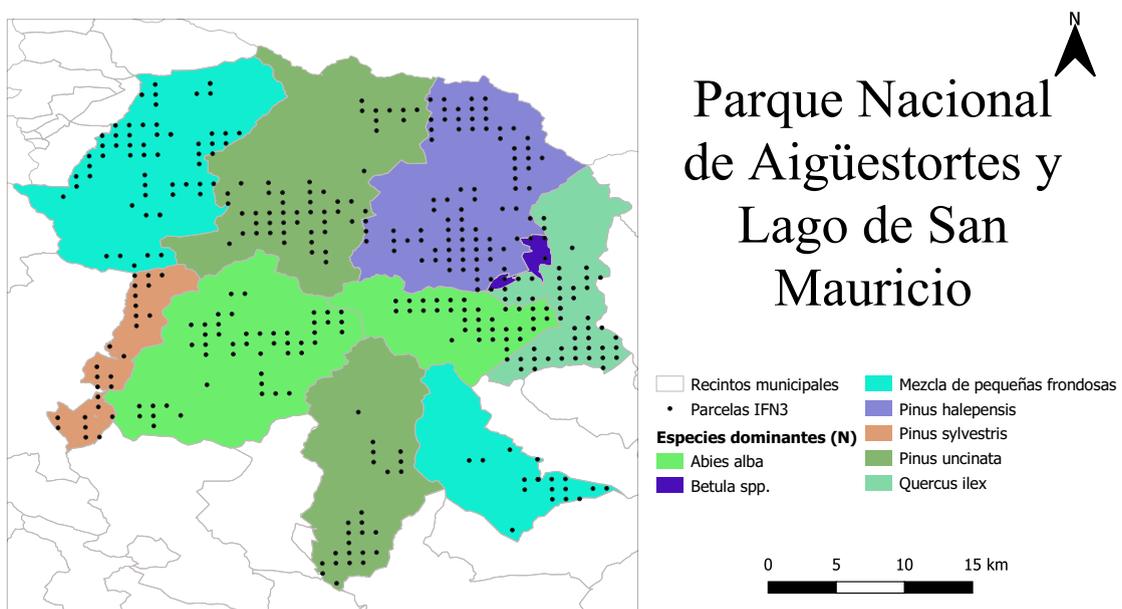


Figura M.3: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Aigüestortes y Lago de San Mauricio

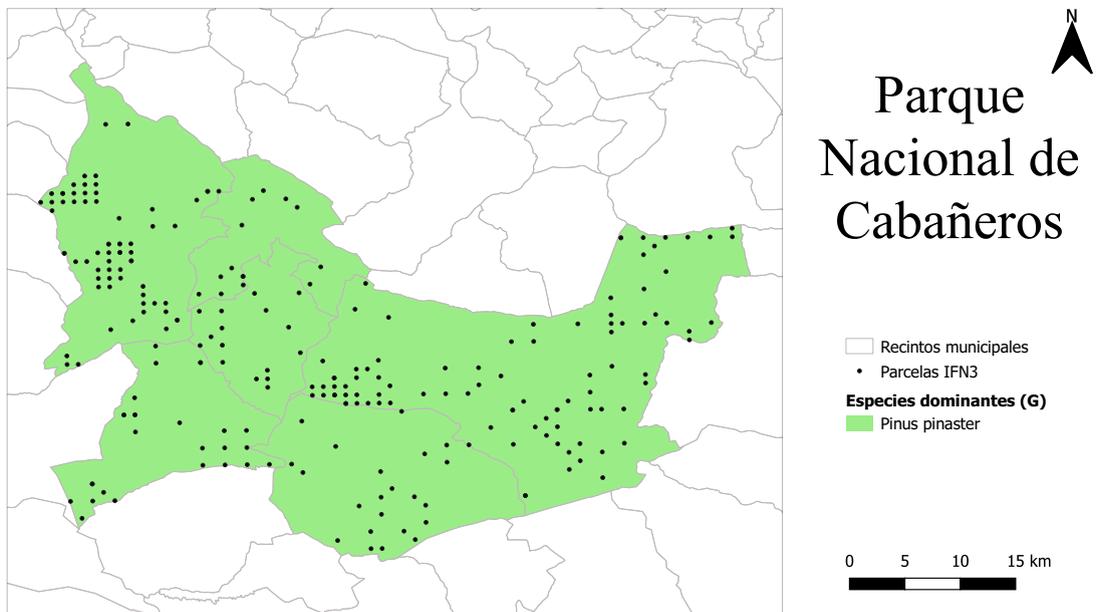


Figura M.4: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Cabañeros

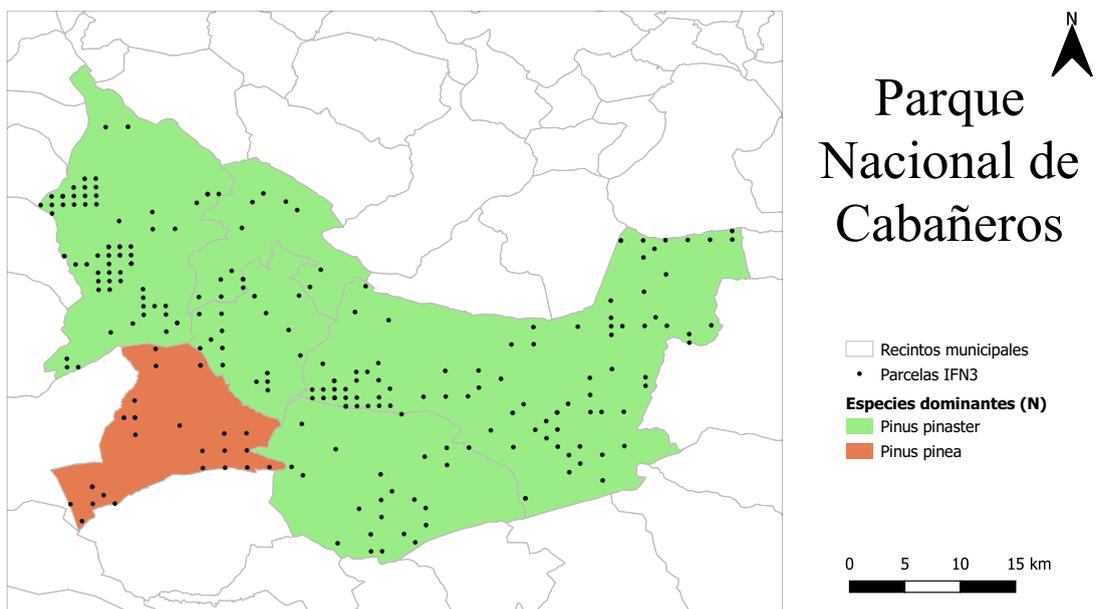


Figura M.5: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Cabañeros

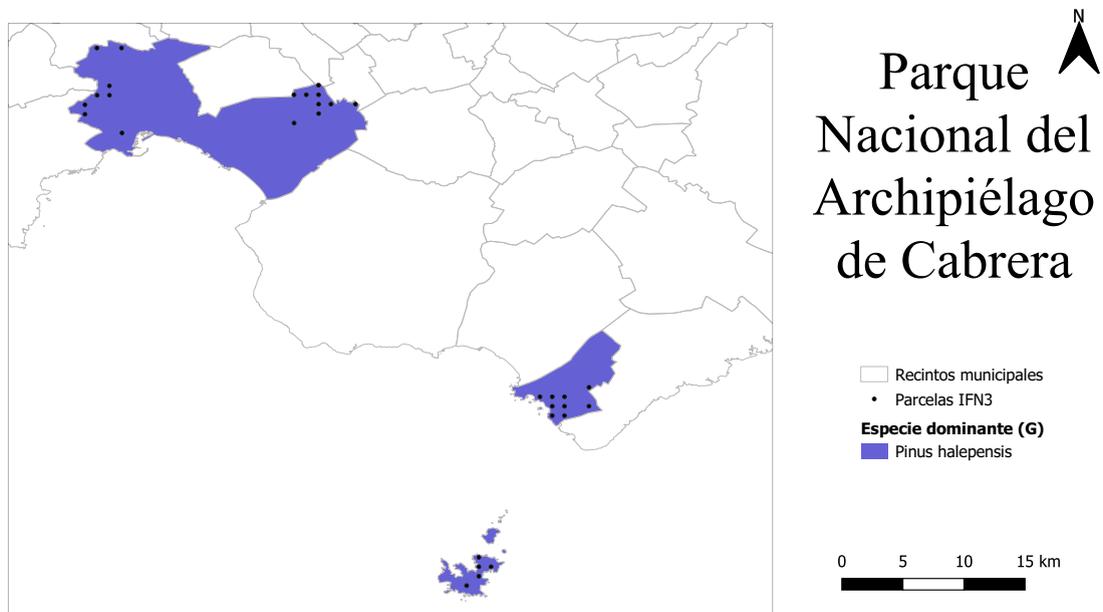


Figura M.6: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional del Archipiélago de Cabrera

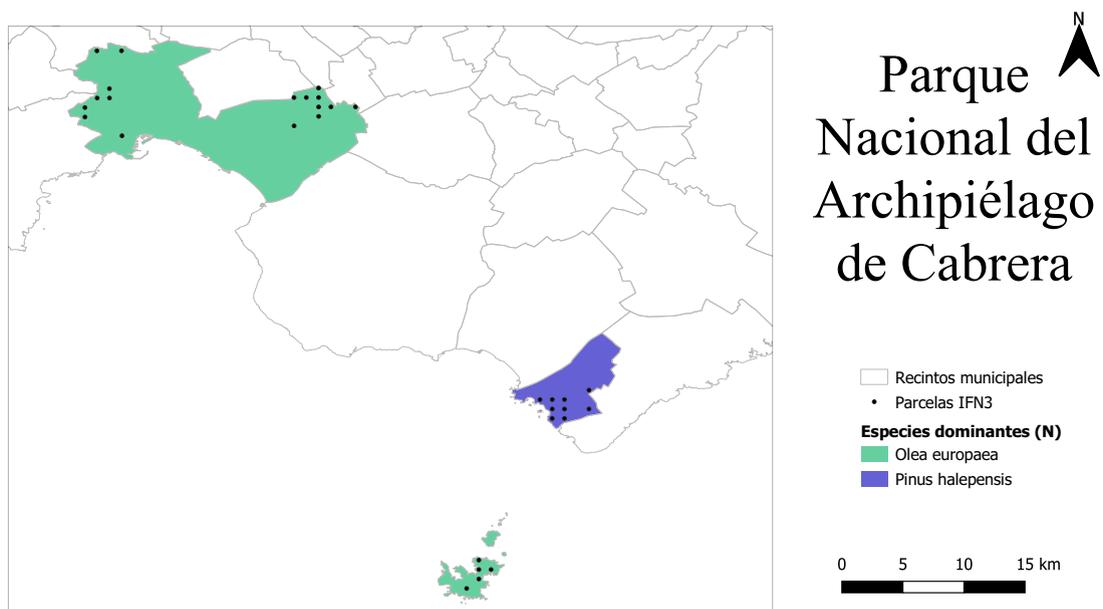


Figura M.7: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional del Archipiélago de Cabrera

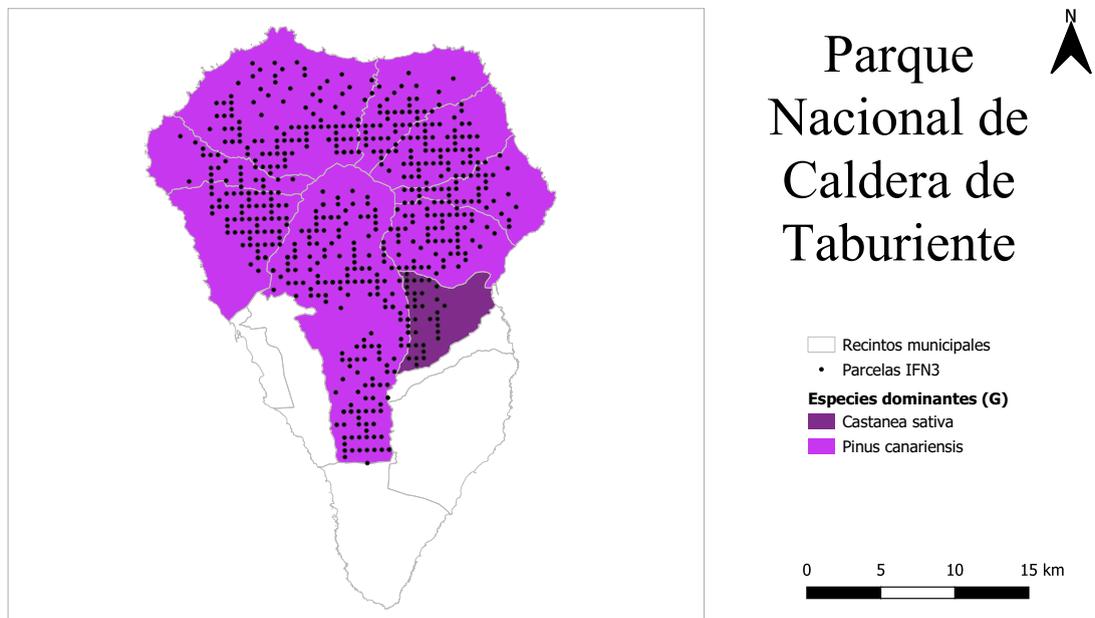


Figura M.8: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Caldera de Taburiente

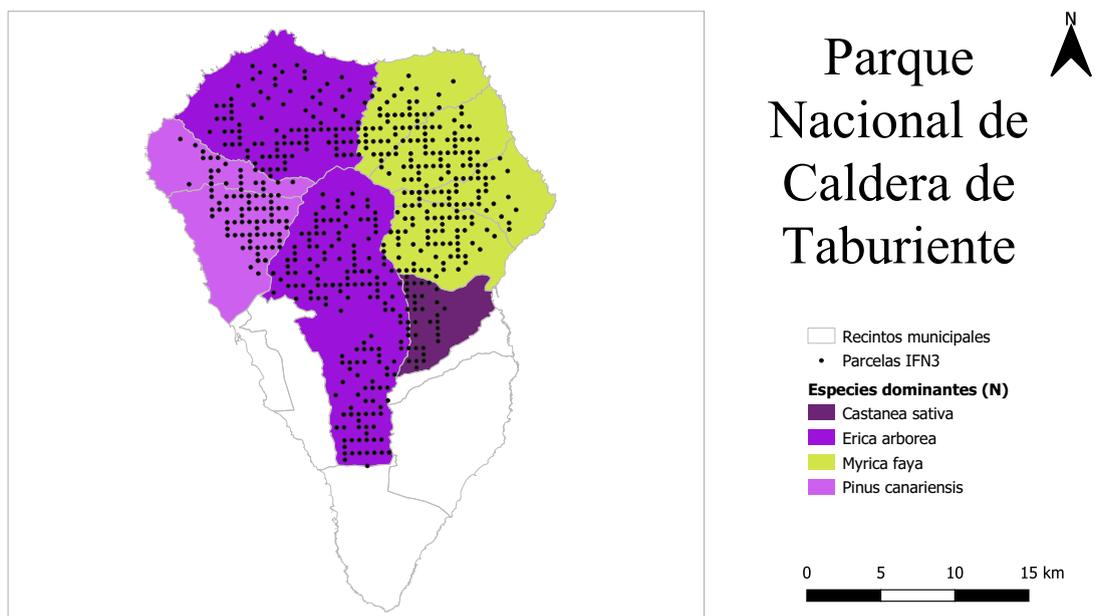


Figura M.9: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Caldera de Taburiente

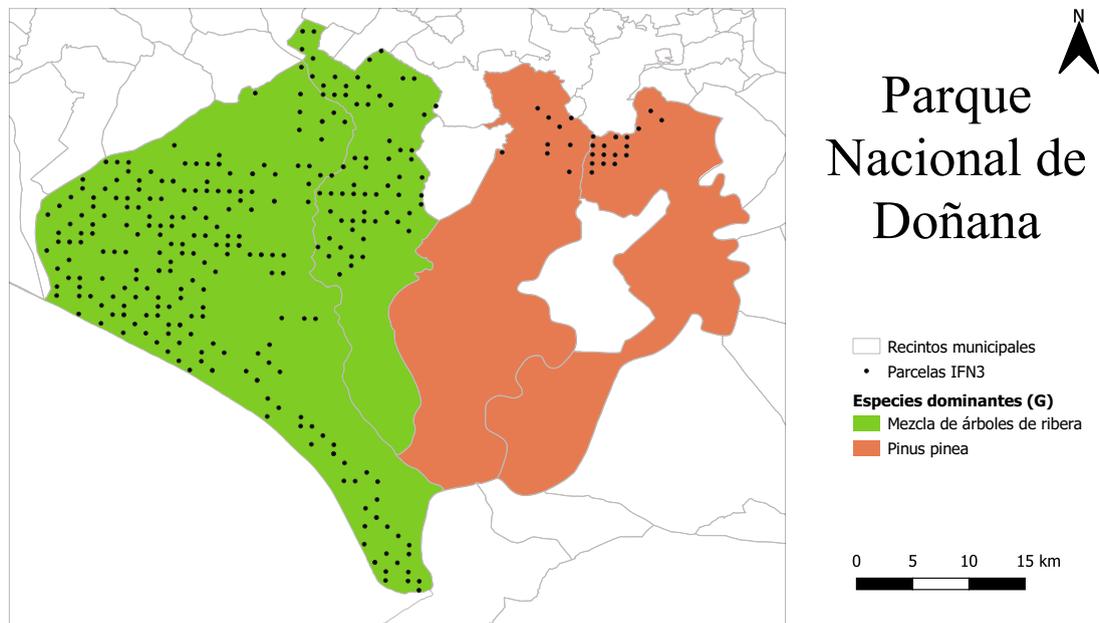


Figura M.10: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Doñana

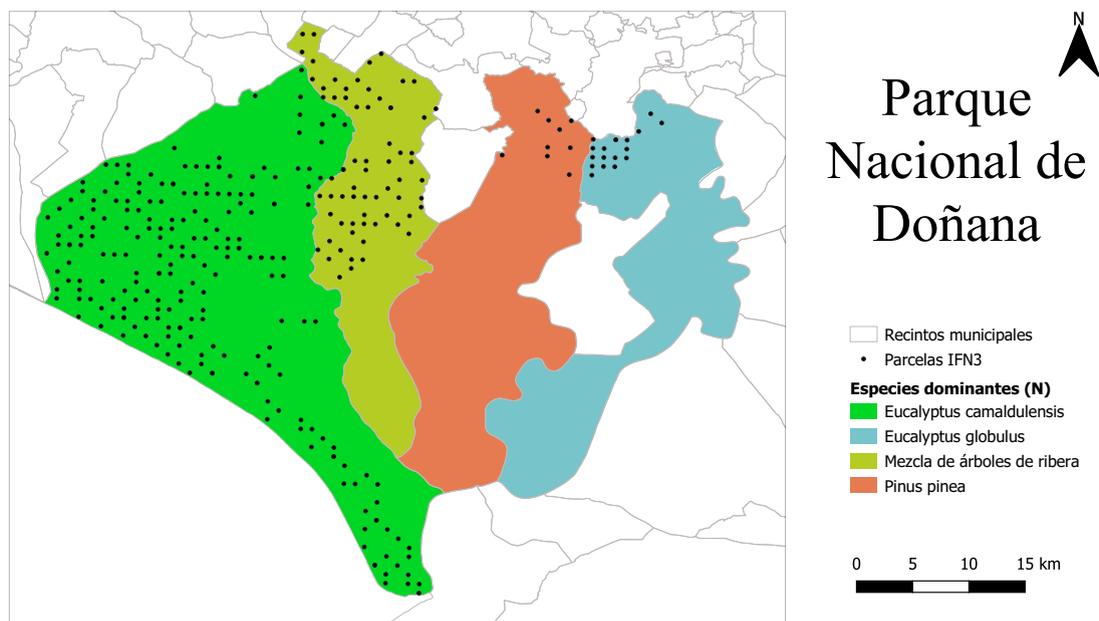


Figura M.11: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Doñana

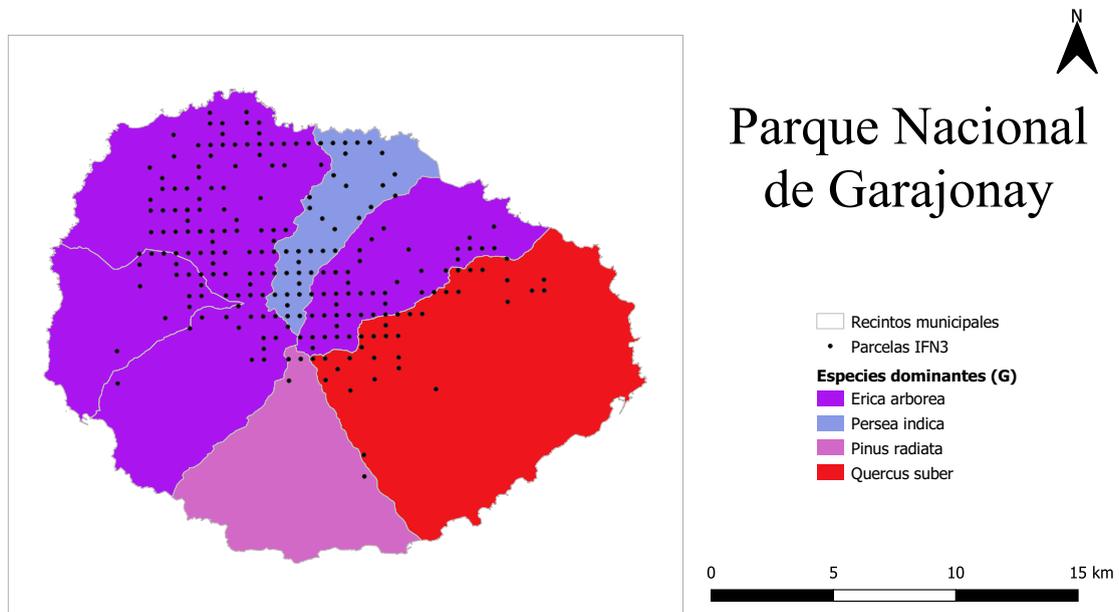


Figura M.12: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Garajonay

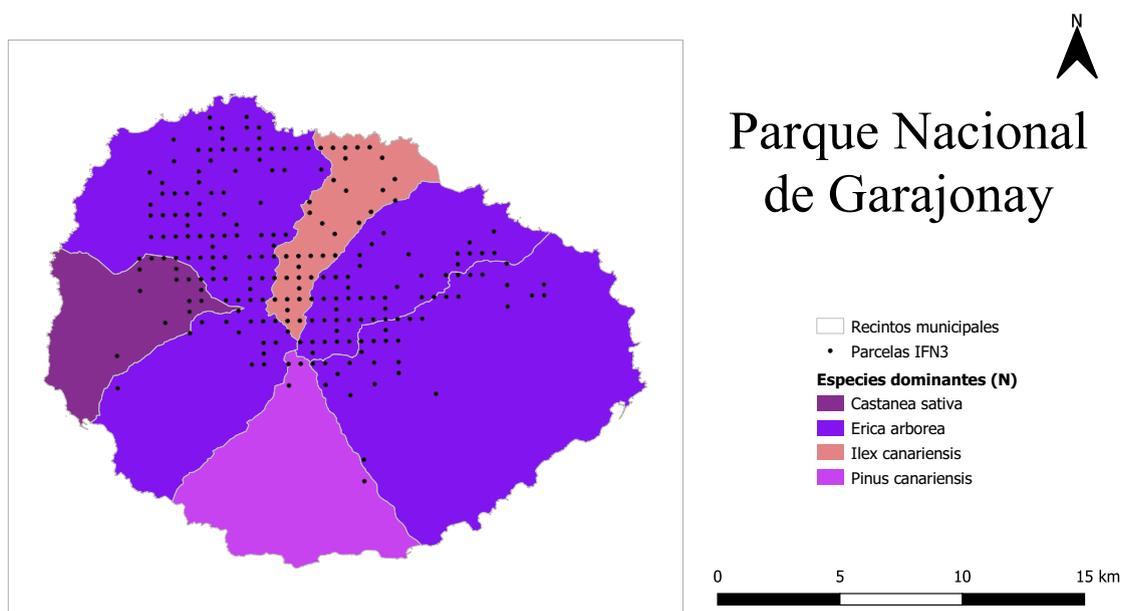


Figura M.13: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Garajonay

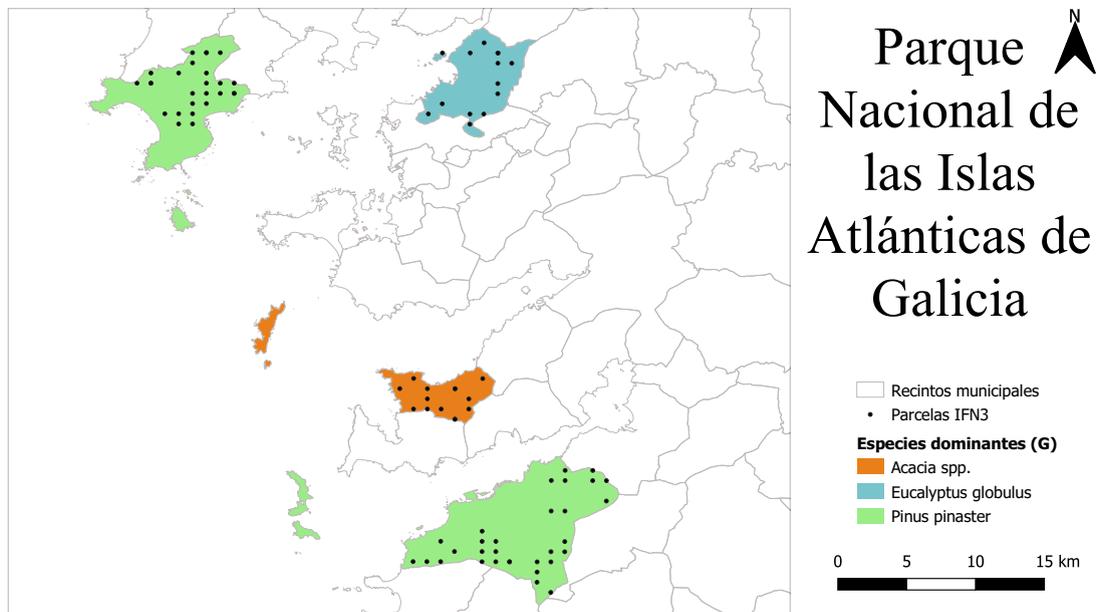


Figura M.14: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de las Islas Atlánticas

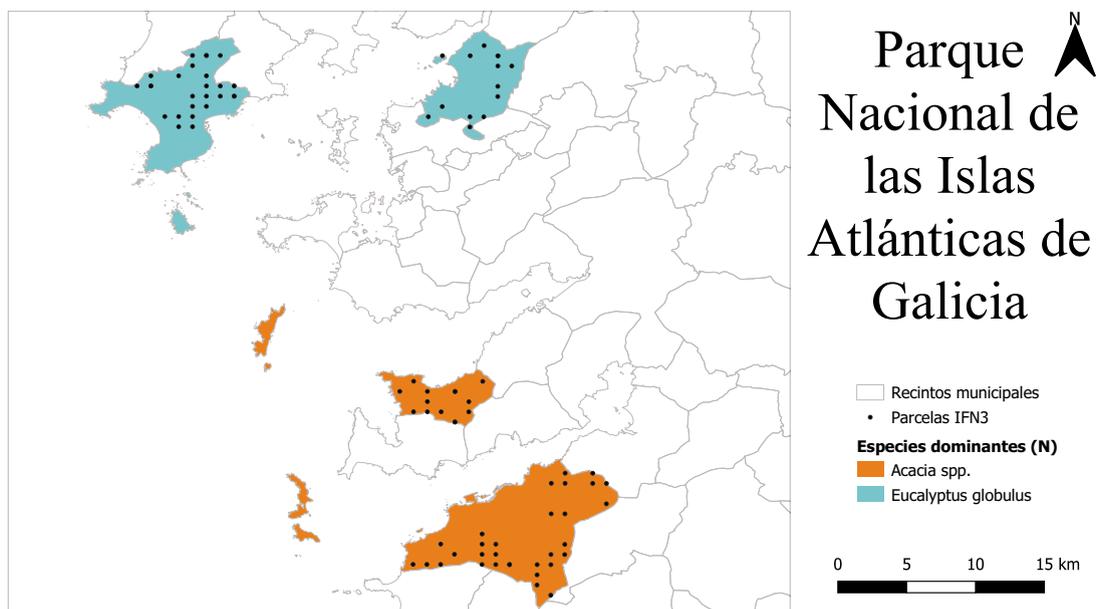


Figura M.15: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de las Islas Atlánticas

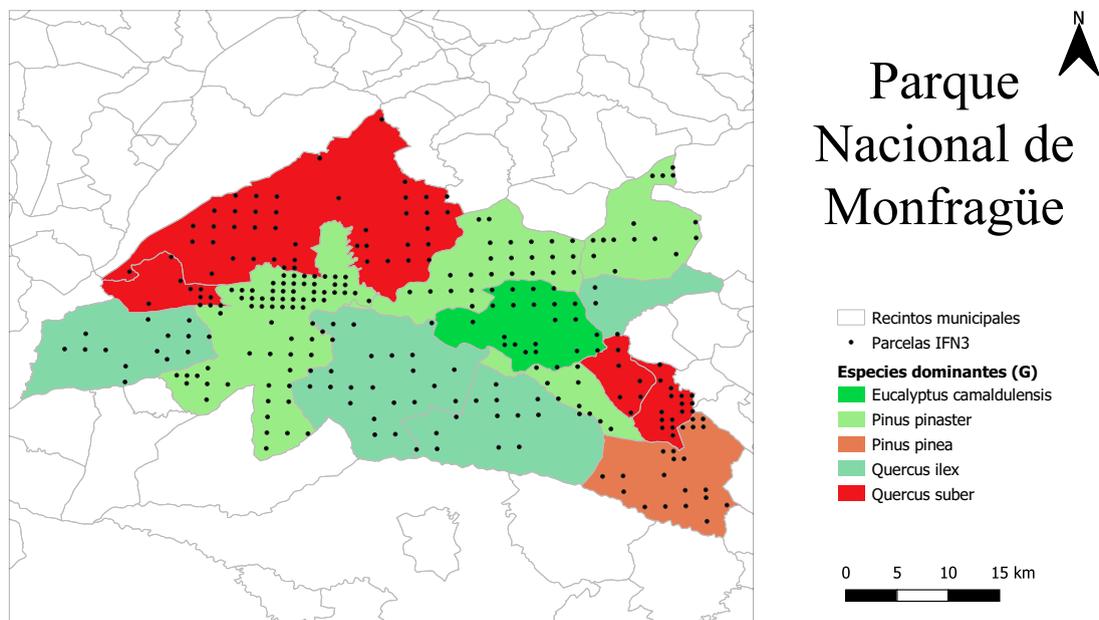


Figura M.16: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Monfragüe

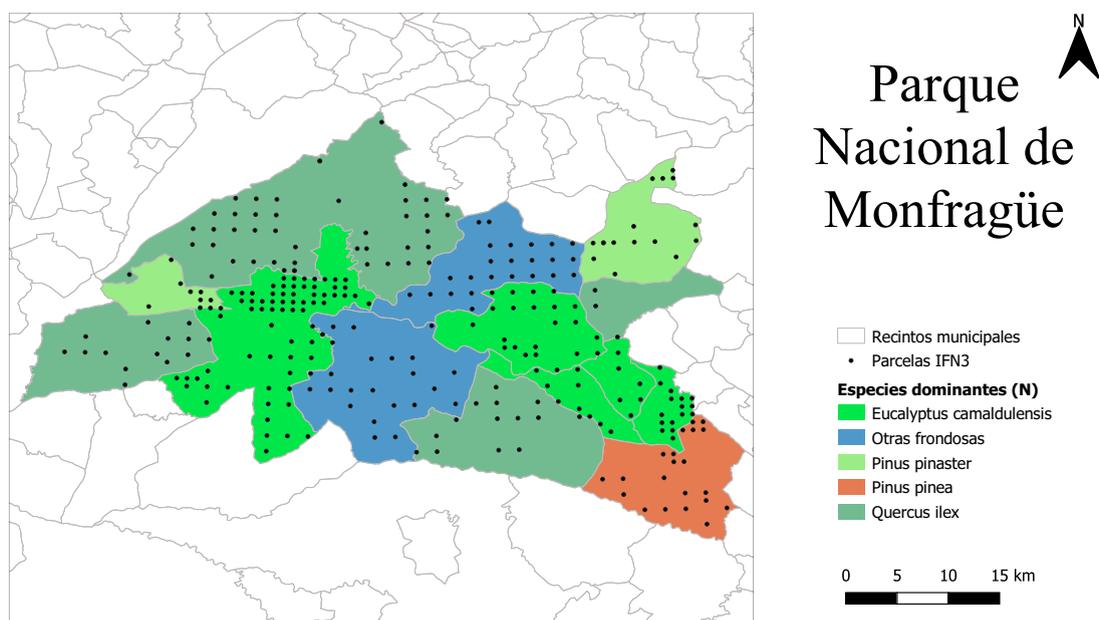


Figura M.17: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Monfragüe

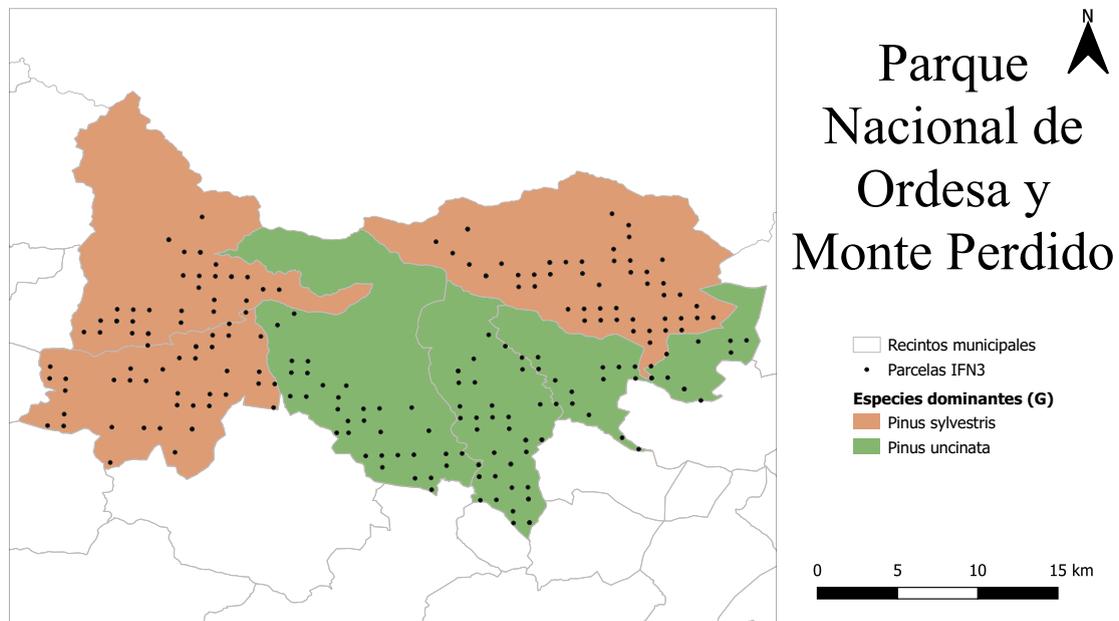


Figura M.18: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Ordesa y Monte Perdido

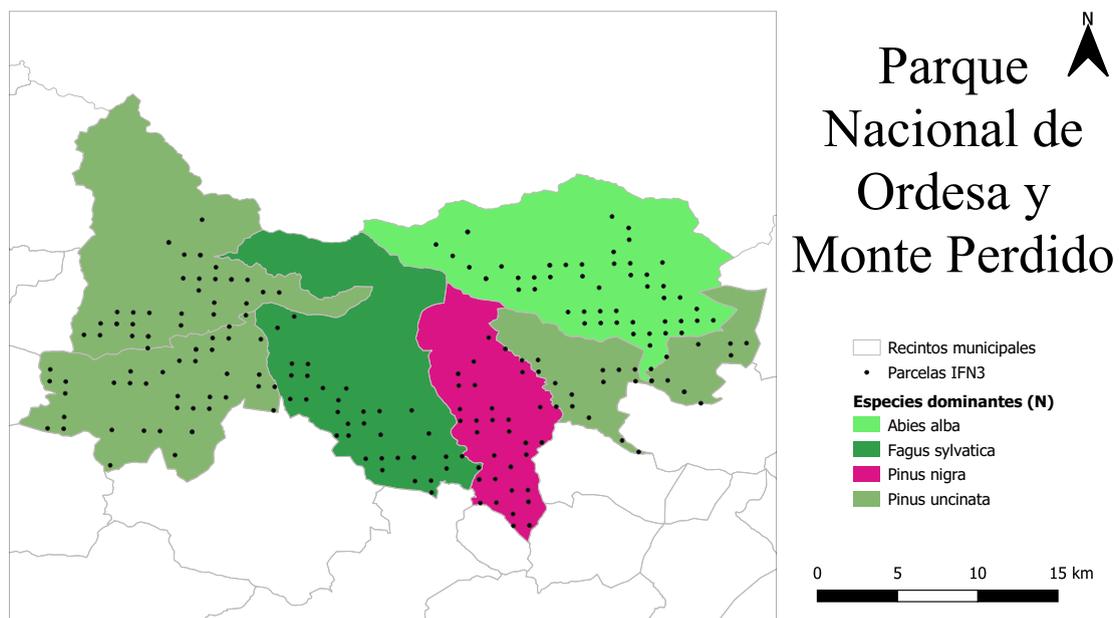


Figura M.19: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Ordesa y Monte Perdido

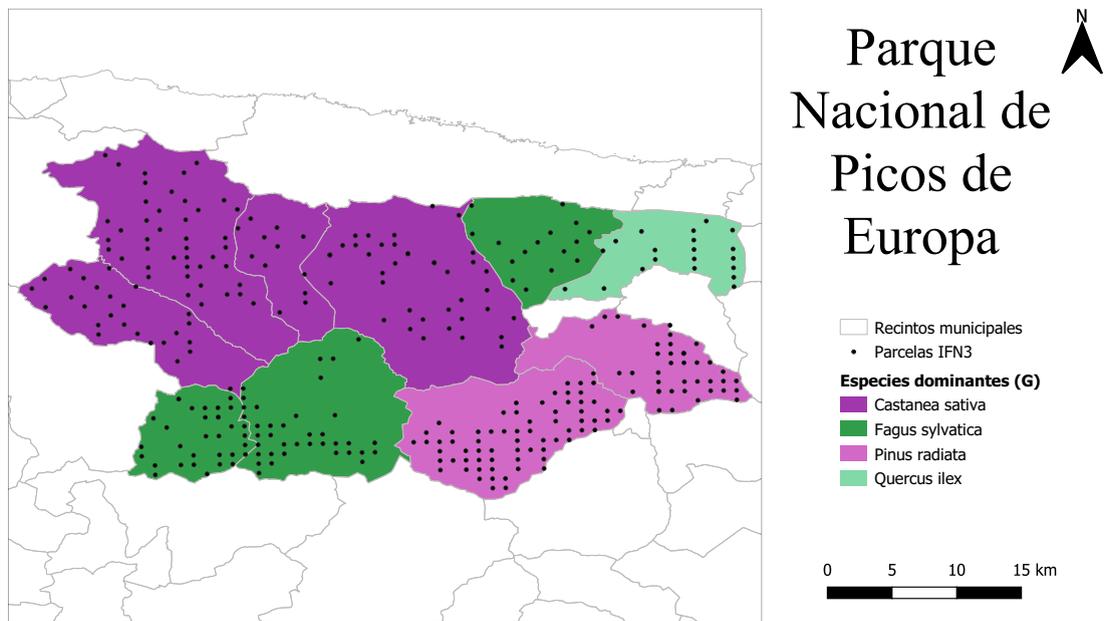


Figura M.20: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Picos de Europa

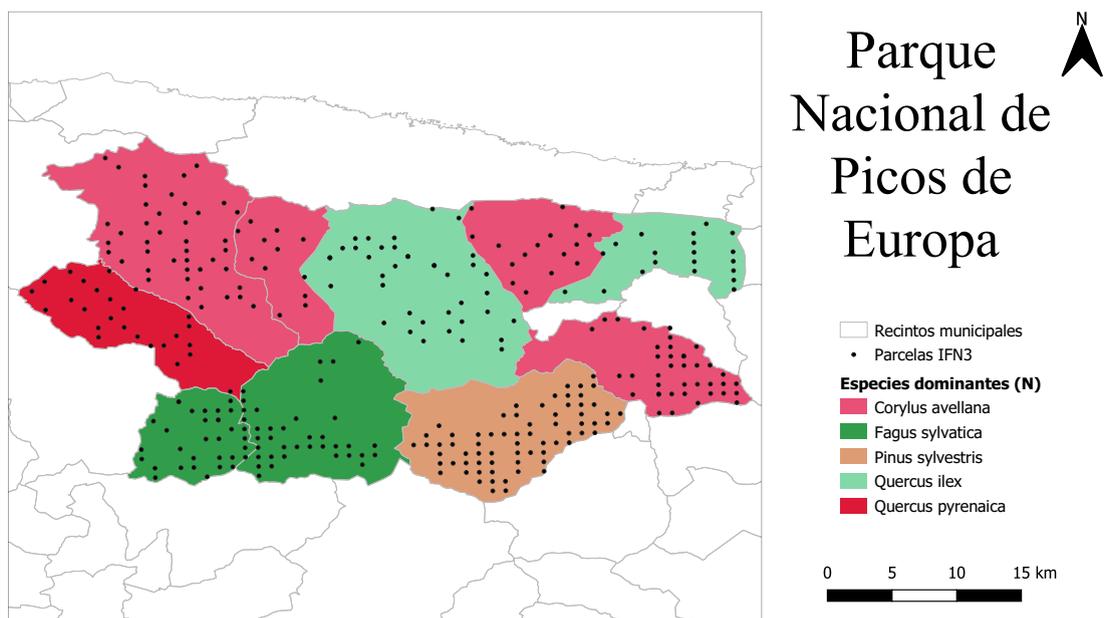


Figura M.21: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Picos de Europa

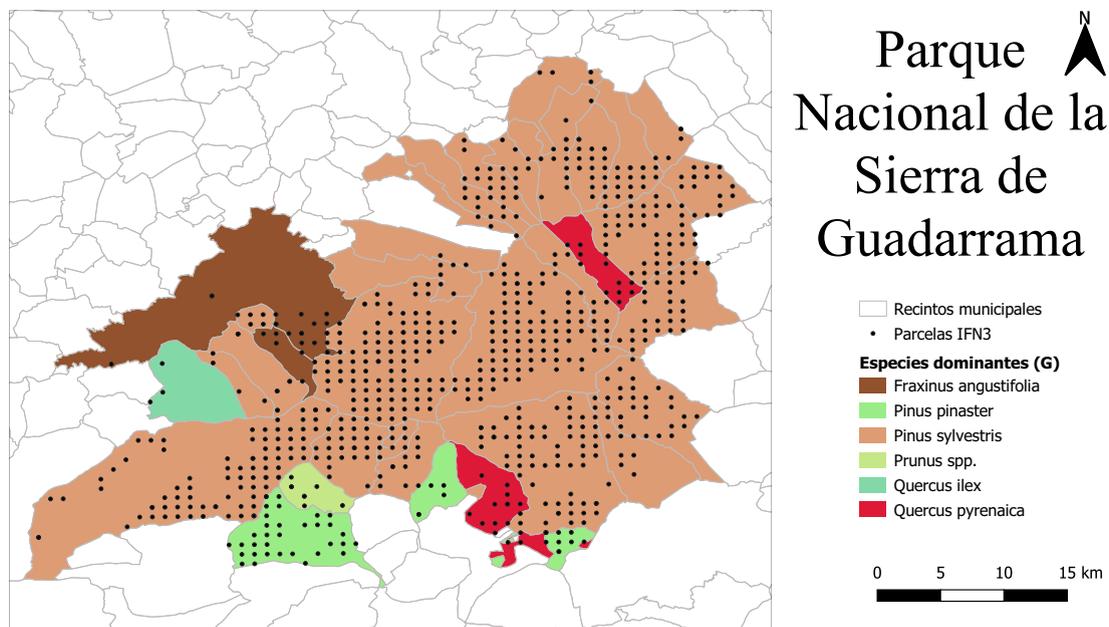


Figura M.22: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Sierra de Guadarrama

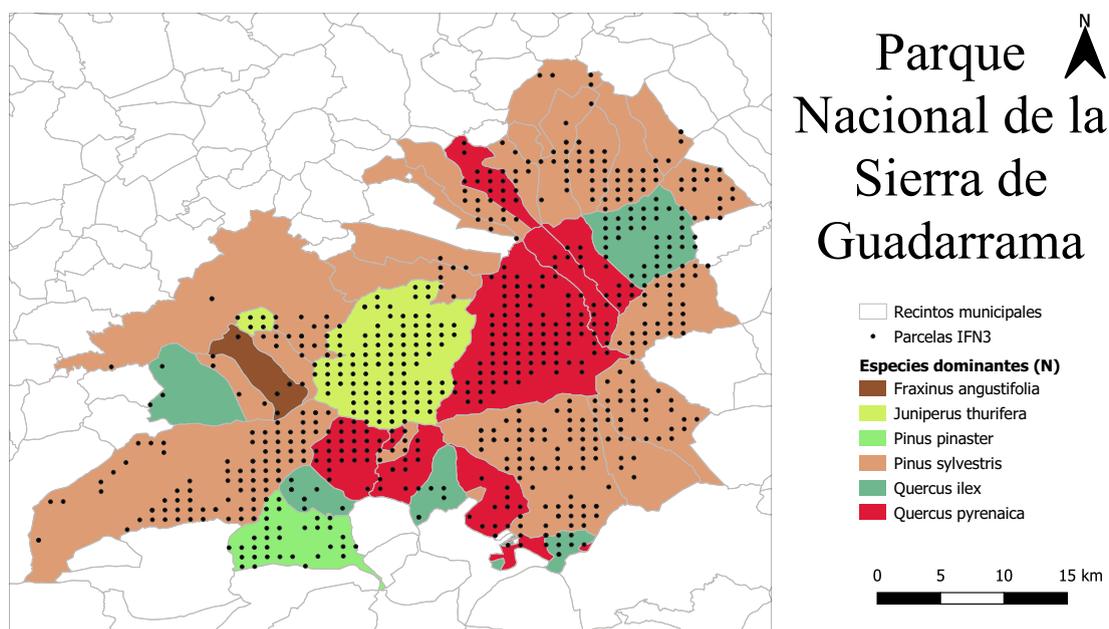


Figura M.23: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Sierra de Guadarrama

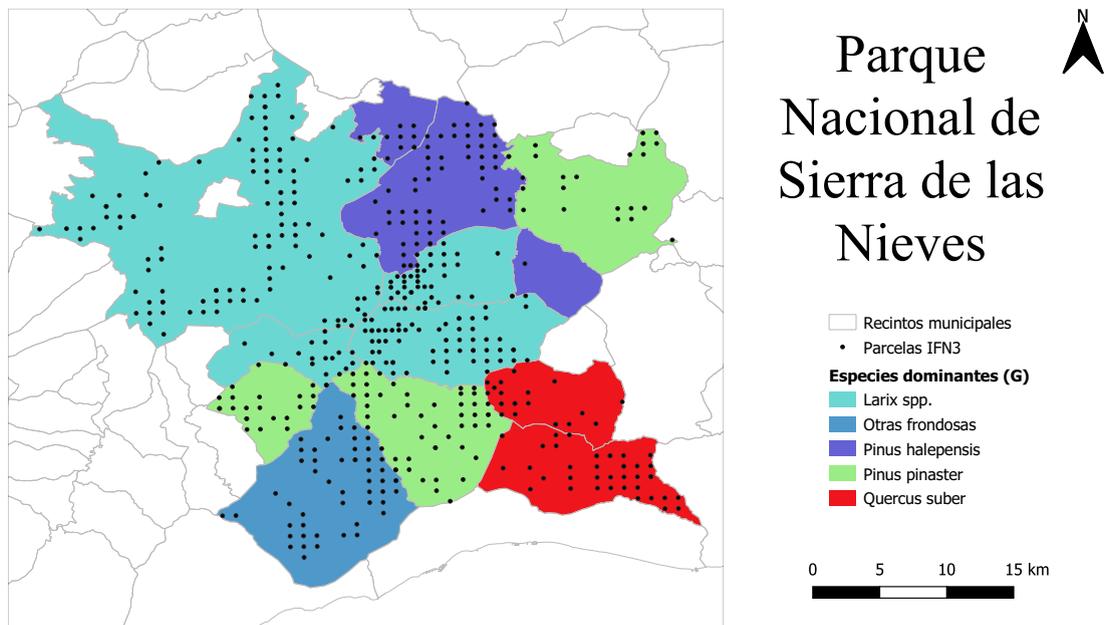


Figura M.24: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Sierra de las Nieves

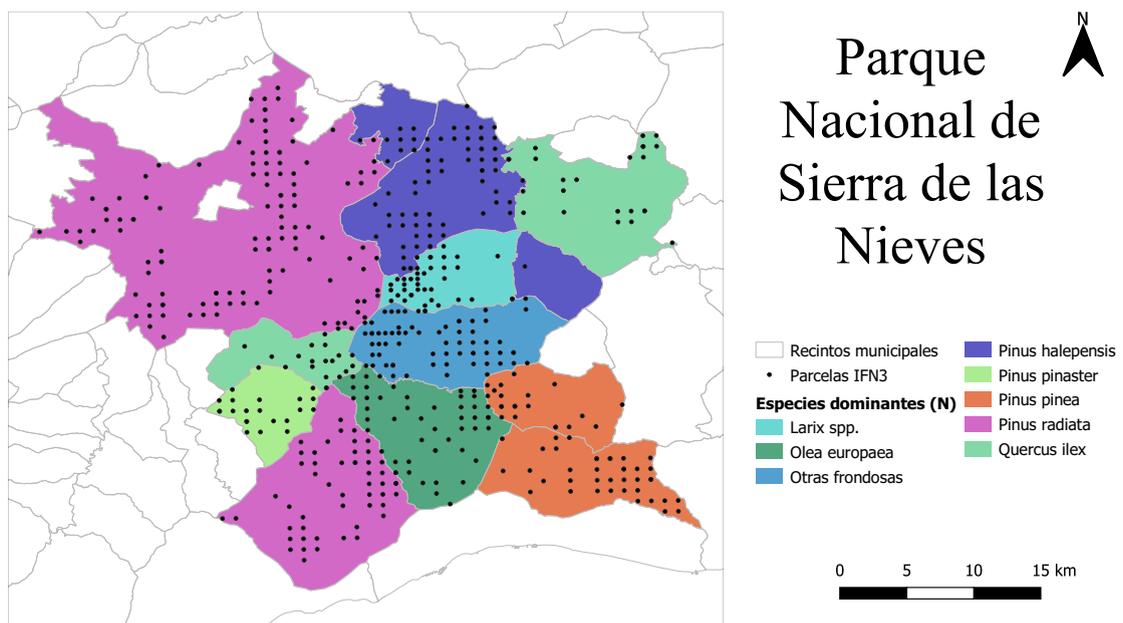


Figura M.25: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Sierra de las Nieves

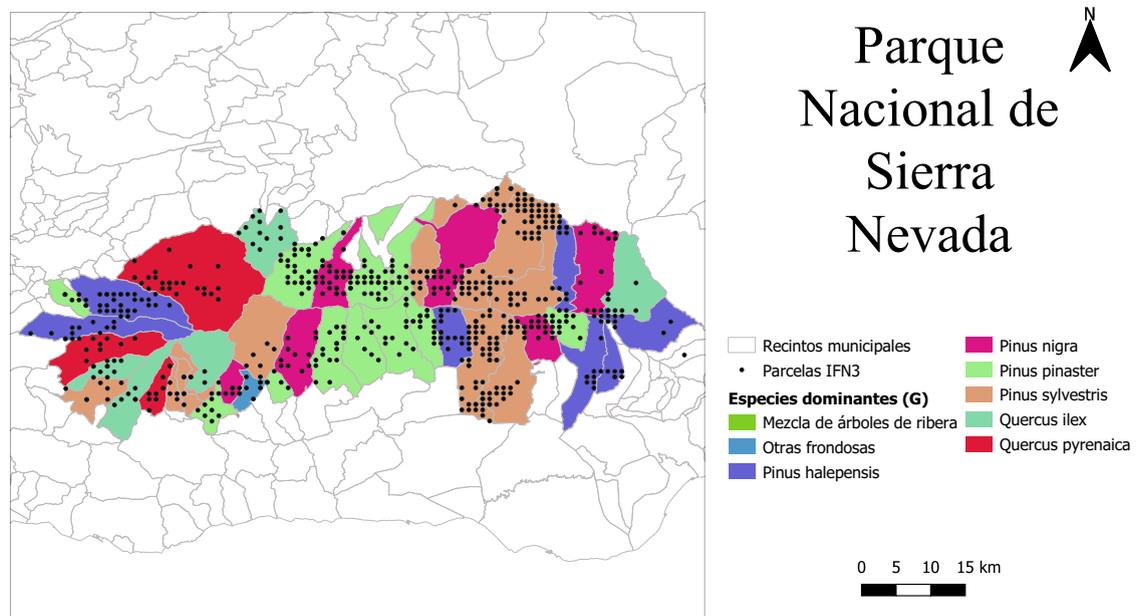


Figura M.26: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional de Sierra Nevada

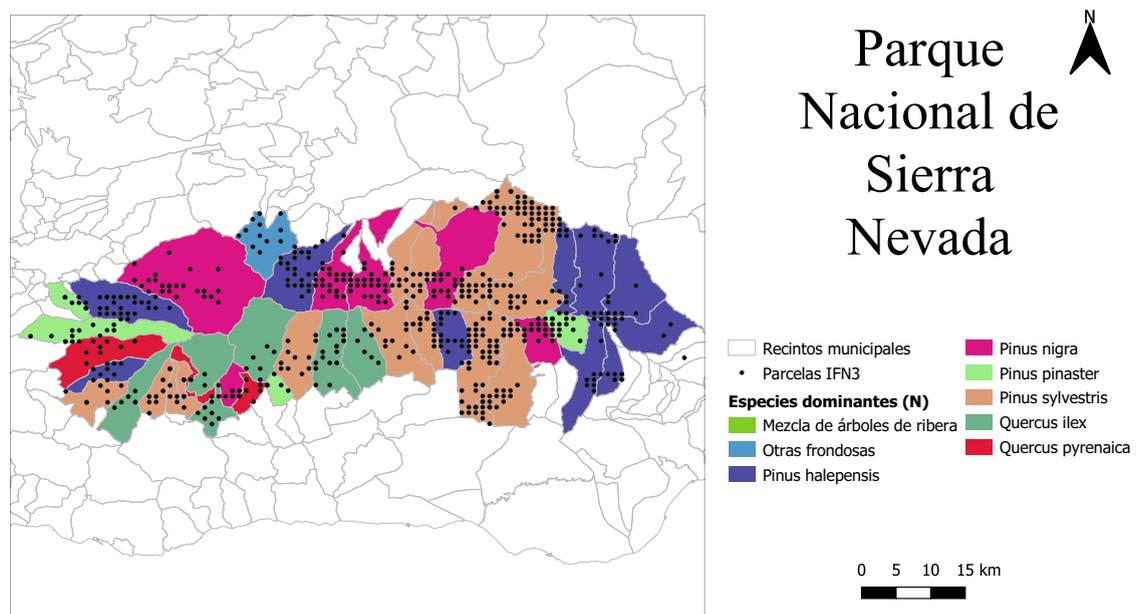


Figura M.27: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional de Sierra Nevada

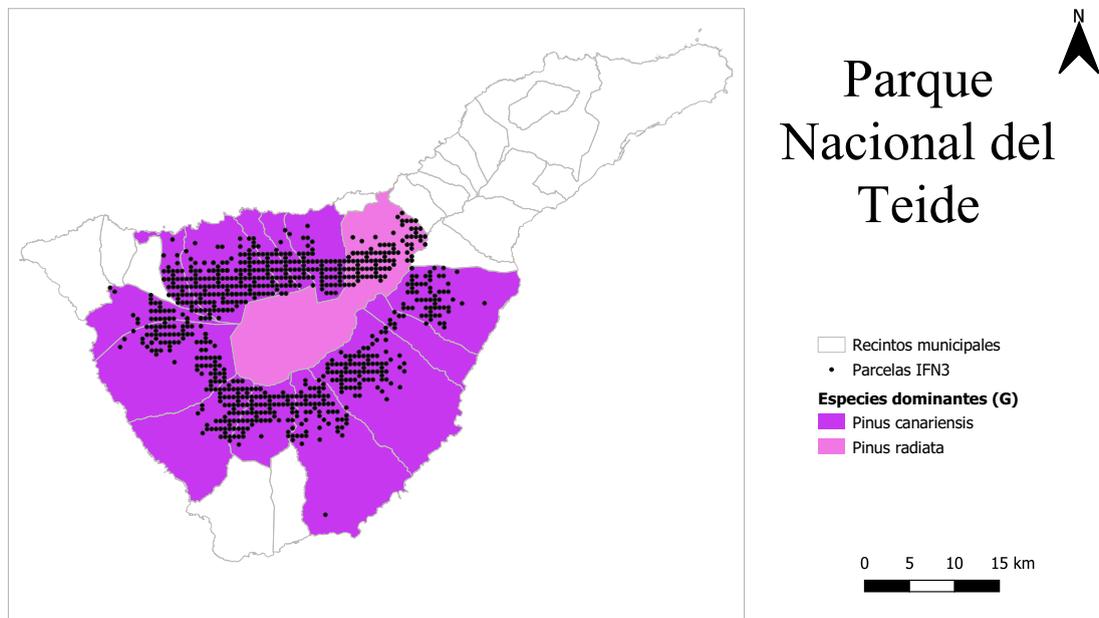


Figura M.28: Mapa de especies dominantes por municipio según el área basimétrica ( $m^2/ha$ ) y parcelas del IFN3 del Parque Nacional del Teide

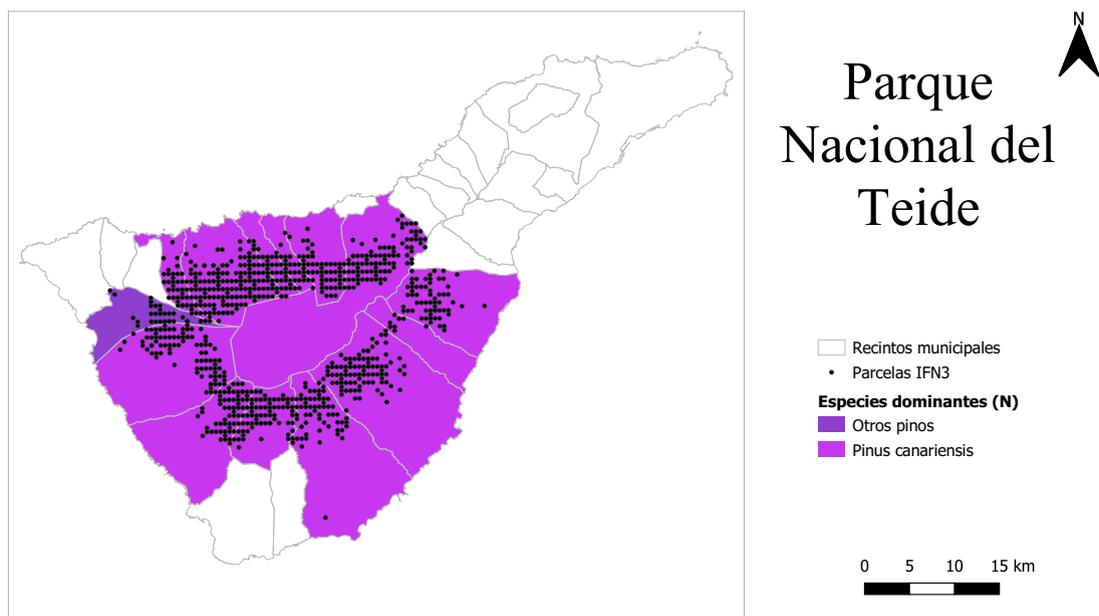


Figura M.29: Mapa de especies dominantes por municipio según el número de pies por hectárea y parcelas del IFN3 del Parque Nacional del Teide