

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE GEOGRAFÍA E HISTORIA



TESIS DOCTORAL

Evaluación de la huella digital del turismo de naturaleza
en los parques nacionales de España

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

Diana Carolina Barros Sulca

DIRECTORES

Javier Gutiérrez Puebla
Juan Carlos García Palomares

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE GEOGRAFÍA E HISTORIA

Doctorado en Geografía



***Evaluación de la huella digital del turismo de naturales en los
parques nacionales de España.***

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Diana Carolina Barros Sulca

Bajo la dirección de los Doctores
Javier Gutiérrez Puebla
Juan Carlos García Palomares

Madrid, 2023

Agradecimientos

La realización de esta tesis ha sido posible gracias al apoyo de varias instituciones y personas a las que quiero expresar mi más profundo agradecimiento. En primer lugar, quiero expresar mi agradecimiento al Grupo de Investigación Transporte, Infraestructuras y Territorio (tGIS) de la Universidad Complutense de Madrid (UCM), que me acogió durante el desarrollo de mi tesis. Han sido una parte fundamental en el desarrollo de esta investigación. También agradezco al Ministerio de Ciencia, Innovación y Universidades, el Fondo Europeo de Desarrollo Regional (FEDER) Proyecto DynMobility, RTI2018-098402-B-I00) y la Agencia Estatal de Investigación (proyecto TourMobility (PID2021-127499OB-I00, MCIN/AEI/ 10.13039/501100011033) por financiar parte de esta tesis.

Aprovecho para expresar mi agradecimiento a mi codirector el Dr. Javier Gutiérrez Puebla, por darme la oportunidad de unirme a su equipo y presentarme este nuevo campo de investigación que lo encuentro fascinante y que me ha abierto muchas puertas y oportunidades en mi carrera. Gracias por su paciencia, experiencia y apoyo constante en estos años, ha sido un honor poder trabajar bajo su dirección. Su tutoría ha sido fundamental para dar forma a mi crecimiento académico y personal, y estaré eternamente agradecida por su dedicación y apoyo.

También quiero expresar mi más sincero agradecimiento a mi segundo codirector, el Dr. Juan Carlos García Palomares, quien ha sido una pieza fundamental en el desarrollo de esta investigación sobre todo en los últimos años. Gracias a su orientación, confianza e infinita paciencia que me han permitido alcanzar este logro. Su tutoría y dedicación en esta tesis me han dado la confianza necesaria para enfrentar los desafíos a lo largo de este camino.

Gracias a todo el Grupo de Investigación tGIS e investigadores del despacho de la planta 12 del departamento de Geografía : Borja Moya Gómez ,Ana Condeço, Joaquín Osorio, Claudia Yubero, Julio César Gómez, Daniela Arias Molinares, Onel Pérez y Sofía Mendoza. Ha sido un privilegio ser parte de este grupo y estoy agradecida por haber tenido la oportunidad de conocer y compartir con personas tan talentosas y dedicadas.

Quiero extender mi agradecimiento a mis compañeros más cercanos y con los que tuve el honor de trabajar día a día : A Borja Moya Gómez, gracias por su ayuda para entender el mundo de los datos geolocalizados y explicaciones tan brillantes, que hicieron posible esta tesis. A Daniela Arias Molinares, mi compañera de mesa en el despacho que se ha convertido en una de mis mejores amigas, gracias por su ayuda y estar siempre presente, sus palabras de ánimo fueron clave para poder terminar esta tesis. A Julio César Gómez y Onel Pérez Fernández, que además de

compañeros de despacho fueron mis compañeros de piso, gracias por su amistad, sus consejos y opiniones.

Asimismo, me gustaría agradecer a mi familia y amigos por su apoyo constante y comprensión durante estos años. A mi madre Lorena por su amor y consejos, estoy profundamente agradecida por todo lo que ha hecho para ayudarme a alcanzar esta meta. A mi hermana Cristina por su cariño y brindarme su ayuda incondicional.

Finalmente quiero agradecer a Esteban, mi compañero de la vida por apoyarme y por siempre estar ahí para escucharme. Gracias por siempre creer en mí y motivarme a seguir adelante con mis sueños.

Índice

Agradecimientos.....	1
Índice de figuras:	6
Índice de tablas:	9
Resumen.....	11
Summary	13
1. Introducción.....	16
1.1. Interés y oportunidad de la investigación.....	16
1.2. Preguntas y objetivos de la investigación.....	19
1.2.1. Preguntas de investigación.....	19
1.2.2. Objetivos.....	22
Objetivo general	22
Objetivos específicos	23
1.3. Estructura de la tesis.	27
1.4. Artículos y Comunicaciones.....	30
2. Turismo en espacios protegidos y nuevas fuentes de datos.	33
2.1. Los espacios protegidos como destino turístico.....	34
2.1.1. La protección del medio natural - breve historia.....	34
2.1.2. El turismo de naturaleza y el uso público de los espacios naturales protegidos.	36
2.1.3. El turismo en parques nacionales y sus impactos.....	37
2.1.4. La capacidad de carga turística en los espacios protegidos.....	39
2.2. Las nuevas fuentes de datos.....	41
2.2.1. Big Data - la revolución de los datos.....	43
2.2.2. Redes Sociales y turismo.	45
3. El uso de datos geolocalizados de redes sociales en el estudio del turismo en espacios naturales protegidos.	51
3.1. Metodología de la revisión del estado del arte.	52
3.2. Descripción general de los artículos revisados.	56
3.3. ¿Cuántos visitantes van a los espacios naturales protegidos?	58
3.4. ¿De dónde vienen los visitantes?	61
3.5. ¿Cuándo visita la gente los espacios protegidos?.....	62
3.6. ¿Qué hacen los visitantes en los espacios protegidos?	64
3.7. ¿Por qué visitan los espacios naturales protegidos?	65
3.8. Conclusiones de la revisión de los artículos.....	68
4. Área de Estudio.	71
4.1. La red de Parques Nacionales de España: Breve historia.	71

4.2.	Características de la red de parques nacionales.....	73
5.	Datos utilizados: fuentes, procesos de descarga y tratamiento previo.....	89
5.1.	Fuentes de datos cartográficos y estadísticas oficiales.....	89
5.1.1.	Datos cartográficos.	90
5.1.2.	Estadísticas oficiales: Memorias del Organismo Autónomo de Parques Nacionales.	92
5.2.	Fuentes de los datos de redes sociales.	92
5.2.1.	Flickr.....	92
5.2.2.	Wikiloc.	93
5.3.	Procesos de descarga de los datos geolocalizados de redes sociales.	94
5.4.	Preprocesamiento y limpieza de los datos geolocalizados de redes sociales.	100
5.4.1.	Limpieza y preparación de datos de Flickr.	100
5.4.2.	Agregación espacial de las fotografías geolocalizadas.	101
5.4.3.	Limpieza y preproceso de los datos de Wikiloc.....	103
5.5.	Primera aproximación: exploración de los datos descargados.	105
5.5.1.	Las fotografías geolocalizadas de Flickr de la red de parques nacionales.	105
5.5.2.	Las rutas GPS de Wikiloc de la red de parques nacionales.....	110
6.	Lugar de origen de los visitantes de la red de parques nacionales españoles.	115
6.1.	Clasificación de las nacionalidades de los usuarios.....	115
6.2.	Validación del origen de los usuarios de Flickr como indicador del origen de los visitantes.	117
6.3.	Distribución de los visitantes por origen.....	119
6.4.	Países de origen de los visitantes extranjeros.....	119
6.5.	Comunidad autónoma de origen en los visitantes nacionales.....	123
7.	El ritmo temporal de los visitantes de la red de parques nacionales españoles.	127
7.1.	Metodología para el análisis temporal.....	127
7.2.	¿Los usuarios de Flickr reflejan la distribución mensual de visitantes?	130
7.3.	El ritmo mensual, semanal y diario de las visitas.	132
7.4.	Patrones de estacionalidad en las visitas a la red de parques españoles.....	135
7.5.	Tipología de parques según la temporalidad de sus visitas.	138
8.	Patrones espaciales del uso turístico de la red de parques nacionales españoles.....	142
8.1.	Metodología para el análisis espacial y la localización óptima de instalaciones de servicios.	143
8.2.	La distribución espacial de los visitantes de la red de parques nacionales.	147
8.3.	Análisis espacio - temporal de los visitantes de la red de parques nacionales.	153
8.4.	Análisis de autocorrelación espacial: identificando puntos calientes.	159

8.5.	Explorando la localización óptima de instalaciones de uso público en los parques nacionales – una aplicación del modelo de ubicación-asignación con datos geolocalizados de Flickr.	167
9.	Los itinerarios en la red de parques nacionales españoles.....	172
9.1.	Metodología para el análisis del uso de las redes de senderos y la distribución de los flujos.	173
9.2.	¿Existen diferencias importantes entre el movimiento de visitantes que practican senderismo y ciclismo?.....	177
9.3.	Flujos de visitantes y movimiento – una aplicación del análisis de redes con rutas GPS de Wikiloc y Flickr.....	184
10.	Conclusiones y futuras líneas de investigación.....	190
10.1.	Conclusiones.....	190
10.2.	Futuras líneas de investigación:.....	199
11.	Bibliografía.....	202
12.	Anexos.....	215
12.1.	Detalle de análisis de autocorrelación espacial:.....	215
12.2.	Descripción de las redes de senderos (grafos) y red de itinerarios de los parques nacionales:.....	230

Índice de figuras:

Figura 1. Diagrama de Preguntas de investigación alineadas con los objetivos de la investigación .	26
Figura 2. Grado de complejidad del cálculo de la capacidad de carga turística	41
Figura 3. Las cinco Vs definitorias del Big Data o datos masivos.....	45
Figura 4. ¿Qué pasa en el internet en 1 minuto 2020 y 2021?	46
Figura 5. Esquema del proceso de filtrado de artículos	54
Figura 6. Producción de artículos por año.	56
Figura 7. Distribución de artículos por fuente de datos utilizada.....	57
Figura 8. Artículos revisados de acuerdo con la categoría de espacio protegido analizado.	58
Figura 9. Distribución de los artículos por país del área de estudio.	58
Figura 10. Mapa de localización de los 15 parques nacionales españoles.....	71
Figura 11. Mapa de Picos de Europa.....	73
Figura 12. Mapa de Ordesa y Monte perdido	74
Figura 13. Mapa del parque nacional Cañadas del Teide.....	75
Figura 14. Mapa del parque nacional Caldera de Taburiente	76
Figura 15. Mapa del parque nacional Aigüestortes I Estany de Sant Maurici	77
Figura 16. Mapa del parque nacional Doñana	78
Figura 17. Mapa del parque nacional Tablas de Daimiel.....	79
Figura 18. Mapa del parque nacional Timanfaya	80
Figura 19. Mapa del parque nacional Garajonay	81
Figura 20. Mapa del parque nacional Archipiélago de Cabrera	82
Figura 21. Mapa del parque nacional Cabañeros.....	83
Figura 22. Mapa del parque nacional Sierra Nevada	84
Figura 23. Mapa del parque nacional Islas Atlánticas de Galicia	85
Figura 24. Mapa del parque nacional Monfragüe.....	86
Figura 25. Mapa del parque nacional Sierra de Guadarrama.....	87
Figura 26. Ejemplo de código para descarga de datos de Open Street Maps.....	91
Figura 27. Evolución de la red social Flickr.....	93
Figura 28. Proceso de descarga de fotografías geolocalizadas de Flickr	94
Figura 29. Proceso de web scraping (diagrama)	97
Figura 30. Ejemplo de recorte de fotografías geolocalizadas por límite administrativo.	100
Figura 31. Ejemplo de modelo de agregación de fotografías geolocalizadas.....	102
Figura 32. Modelo de conversión de archivos GPX a formato SIG (ejemplo).....	103
Figura 33. Modelo para el filtrado de rutas GPS de acuerdo con criterio de duración	104
Figura 34. Ejemplo de transformación de puntos GPS a líneas en el parque nacional Aigüestortes	104
Figura 35. Ilustración gráfica del cálculo de los parámetros de movimiento en ArcMap	105
Figura 36. Distribución espacial de las fotografías geolocalizadas de Flickr en la red de parques nacionales españoles (1 de 2)	108
Figura 37. Distribución espacial de las fotografías geolocalizadas de Flickr en la red de parques nacionales españoles (2 de 2)	109
Figura 38. Distribución espacial de las rutas GPS descargadas de Wikiloc.....	112
Figura 39. Distribución espacial de las rutas GPS de Wikiloc descargadas por web scraping de las actividades de senderismo y ciclismo	113
Figura 40. Modelo aplicado para la detección del país de origen de los visitantes de los parques nacionales	116

Figura 41. Ejemplo de sentencia SQL para selección y agrupación de usuarios por país (España).	116
Figura 42. Gráfico de dispersión/comparación de las variables visitantes extranjeros Flickr con visitantes extranjeros de la encuesta del INE.	118
Figura 43. Comparación de % de visitantes extranjeros obtenido de Flickr y % de visitantes extranjeros de la encuesta de alojamientos rurales (INE).	118
Figura 44. Distribución de los visitantes por origen (Nacional / Internacional) de acuerdo con datos de los usuarios de Flickr 2010-2018.	119
Figura 45. Distribución por país de los visitantes internacionales en los parques nacionales canarios.	120
Figura 46. Distribución por país de los visitantes internacionales en los parques nacionales de montaña.	120
Figura 47. Distribución por país de los visitantes internacionales en los parques nacionales insulares.	121
Figura 48. Distribución por país de los visitantes internacionales en los parques nacionales de humedales.	122
Figura 49. Distribución por país de los visitantes internacionales en los parques nacionales de bosque-paisaje mediterráneo.	122
Figura 50. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales canarios.	123
Figura 51. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales de montaña.	124
Figura 52. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales insulares.	124
Figura 53. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales de humedales.	125
Figura 54. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales de bosque-paisaje mediterráneo.	125
Figura 55. Modelo de agregación y cálculo de usuarios únicos por mes.	128
Figura 56. Gráficos de dispersión de la distribución mensual de % de visitantes y % de usuarios únicos de Flickr.	131
Figura 57. Distribución mensual de usuarios únicos de Flickr por parque nacional.	132
Figura 58. Distribución por días de la semana de usuarios únicos de Flickr por parque nacional.	133
Figura 59. Distribución diaria de usuarios únicos de Flickr por parque nacional.	134
Figura 60. Distribución mensual de visitantes en parques con estacionalidad alta.	136
Figura 61. Distribución mensual de visitantes en parques con estacionalidad media.	137
Figura 62. Distribución mensual de visitantes en parques con estacionalidad baja.	138
Figura 63. Mapa de ubicación de los parques y asignación de los clústeres.	140
Figura 64. Densidad de fotografías de Flickr por Km2 en el grupo 1 (parques con valores altos de densidad máxima).	148
Figura 65. Densidad de fotografías de Flickr por Km2 en el grupo 2 (parques con densidad máxima entre 500 y 2000 fotos/km2).	149
Figura 66. Densidad de fotografías de Flickr por Km2 en el grupo 3 (parques con densidad máxima baja menor a 500 fotos/km2).	150
Figura 67. Densidad de fotografías geolocalizadas por estación del parque nacional Aigüestortes.	154
Figura 68. Densidad de fotografías geolocalizadas por estación del parque nacional Teide.	155
Figura 69. Densidad de fotografías geolocalizadas por estación del parque nacional Sierra de Guadarrama.	156

Figura 70. Densidad de fotografías geolocalizadas por estación del parque nacional Islas Atlánticas de Galicia	157
Figura 71. Densidad de fotografías geolocalizadas por estación del parque nacional Monfragüe .	158
Figura 72. Mapas del índice local de autocorrelación espacial LISA (1 de 2).	161
Figura 73. Mapas del índice local de autocorrelación espacial LISA (2 de 2).	162
Figura 74. Distribución de los clústeres HH por tipo de equipamiento.....	165
Figura 75. Resultados de los modelos de ubicación-asignación: (1) 1 establecimiento adicional, (2) 3 establecimientos adicionales, (3) 5 establecimientos adicionales, (4) 10 establecimientos adicionales.	168
Figura 76. Ejemplo ilustrativo del proceso de map matching o ajuste de rutas.....	174
Figura 77. Herramienta para implementar algoritmo de map matching utilizando Model Builder de ArcGIS pro	175
Figura 78. Diagrama de flujo de trabajo para obtención de densidades de uso de senderos.....	175
Figura 83. Ejemplo de detección de paradas en las rutas GPS a partir del índice local de Moran. .	177
Figura 79. Mapas de densidad de uso de la actividad de senderismo (Grupo 1).....	180
Figura 80. Mapas de densidad de uso de la actividad de senderismo (Grupo 2).....	181
Figura 81. Mapas de densidad de uso de la actividad de ciclismo (Grupo 1).....	182
Figura 82. Mapas de densidad de uso de la actividad de ciclismo (Grupo 2).....	183
Figura 84. Distribución mensual de los itinerarios según actividad.	185
Figura 86. Mapa de flujos de visitantes desde puntos de acceso para las actividades de turismo (Flickr), Senderismo (Wikiloc) y Ciclismo (Wikiloc).....	188
Figura 87. Despliegue de datos geolocalizados de las tres redes sociales examinadas en el parque nacional Teide.	194
Figura 88. Mapa de clústeres LISA del parque nacional Aigüestortes.....	215
Figura 89. Mapa de clústeres LISA del parque nacional Cabañeros.	216
Figura 90. Mapa de clústeres LISA del parque nacional Tablas de Daimiel.	217
Figura 91. Mapa de clústeres LISA del parque nacional Sierra de Guadarrama.....	218
Figura 92. Mapa de clústeres LISA del parque nacional Monfragüe.	219
Figura 93. Mapa de clústeres LISA del parque nacional Ordesa y Monte Perdido.....	220
Figura 94. Mapa de clústeres LISA del parque nacional Picos de Europa.	221
Figura 95. Mapa de clústeres LISA del parque nacional Sierra Nevada.....	222
Figura 96. Mapa de clústeres LISA del parque nacional Teide.	223
Figura 97. Mapa de clústeres LISA del parque nacional Timanfaya.	224
Figura 98. Mapa de clústeres LISA del parque nacional Caldera de Taburiente.	225
Figura 99. Mapa de clústeres LISA del parque nacional Garajonay.....	226
Figura 100. Mapa de clústeres LISA del parque nacional Archipiélago de Cabrera.	227
Figura 101. Mapa de clústeres LISA del parque nacional Islas Atlánticas de Galicia – Isla Cíes.....	228
Figura 102. Mapa de clústeres LISA del parque nacional Islas Atlánticas de Galicia – Isla Ons.....	228
Figura 103. Mapa de clústeres LISA del parque nacional Doñana.....	229
Figura 104. Red de grafos y red de itinerarios del P.N. Archipiélago de Cabrera.....	230
Figura 105. Red de grafos y red de itinerarios del P.N. Aigüestortes.	231
Figura 106. Red de grafos y red de itinerarios del P.N. Tablas de Daimiel.	231
Figura 107. Red de grafos y red de itinerarios del P.N. Monfragüe.	232
Figura 108. Red de grafos y red de itinerarios del P.N. Caldera de Taburiente.....	233
Figura 109. Red de grafos e itinerarios en la isla Cíes.	234
Figura 110. Red de grafos e itinerarios en la isla Ons.....	234

Índice de tablas:

Tabla 1. Impactos negativos del uso público de los espacios naturales protegidos.	38
Tabla 2. Consulta de búsqueda estandarizada para ambas bases de datos de artículos científicos.	53
Tabla 3. Artículos revisados organizados de acuerdo con preguntas sobre el turismo en espacios naturales protegidos.	55
Tabla 4. Artículos que abordaron la pregunta ¿Cuántos visitantes?	60
Tabla 5. Artículos que abordaron la pregunta ¿De dónde vienen los visitantes?	62
Tabla 6. Artículos que abordaron la pregunta ¿Cuándo visita la gente los espacios naturales protegidos?	64
Tabla 7. Artículos que abordan la pregunta ¿Qué hacen los visitantes en los espacios protegidos?	65
Tabla 8. Artículos que abordan la pregunta ¿Por qué visitan los espacios naturales protegidos? ...	67
Tabla 9. Disponibilidad datos visitantes por año y mes	92
Tabla 10. Sistema de referencia espacial para cada parque de la red nacional.....	95
Tabla 11. Descripción de los atributos de los metadatos de las fotografías de Flickr.....	96
Tabla 12. Descripción de los atributos de las rutas de Wikiloc descargadas por web scraping.	97
Tabla 13. Descripción de los atributos de las rutas de Wikiloc descargadas de la página web	99
Tabla 14. Tamaño óptimo de lado del hexágono de agregación por parque nacional	102
Tabla 15. Fotografías eliminadas durante los procesos de limpieza y depuración de datos.	105
Tabla 16. Detalle de filtrado de fotografías geolocalizadas de Flickr por parque nacional.....	106
Tabla 17. Detalle por parque de las fotografías geolocalizadas y usuarios de Flickr. (datos depurados)	107
Tabla 18. Rutas obtenidas de la descarga de Wikiloc por parque nacional	110
Tabla 19. Detalle de primer proceso de depuración de rutas de Wikiloc.....	110
Tabla 20. Procesos de limpieza y filtros aplicados a las rutas de Wikiloc	111
Tabla 21. Detalle de rutas por actividad y procesos de limpieza de rutas GPS del parque nacional Teide.....	113
Tabla 22. Índice de estacionalidad mensual y semanal de los 15 parques nacionales	136
Tabla 23. Valor de las variables utilizadas para el análisis k-medias para cada parque	138
Tabla 24. Comparación post-hoc de ANOVA unidireccional entre grupos	139
Tabla 25. Detalle de los grupos obtenidos a partir de K-medias	139
Tabla 26. Tabla resumen de densidad de fotos de Flickr por parque nacional.....	151
Tabla 27. Superficie y número de zonas con densidades superiores a 300 fotografías/Km2.	152
Tabla 28. Índice global de autocorrelación espacial por parque nacional	159
Tabla 29. Detalle del número de hexágonos por clústeres HH, LL, HL y LH y superficie que ocupan en los parques.	163
Tabla 30. Número de equipamientos y puntos de interés por tipo.....	164
Tabla 31. Hexágonos HL por tipo de equipamiento.....	165
Tabla 32. Clústeres HH que atrajeron el mayor número de usuarios en los parques nacionales... ..	166
Tabla 33. Demanda cubierta de 2 instalaciones existentes	169
Tabla 34. Demanda cubierta de las instalaciones seleccionadas en todos los escenarios.....	169
Tabla 35. Estadísticas rutas senderismo.	178
Tabla 36. Estadísticas rutas ciclismo	178
Tabla 37. Características del movimiento de los itinerarios de Flickr y Wikiloc.....	185
Tabla 38. Frecuencia de paradas o puntos de interés en los Itinerarios de senderismo (Wikiloc). ..	186
Tabla 39. Frecuencia de paradas o puntos de interés en los Itinerarios de ciclismo (Wikiloc)	186
Tabla 40. Frecuencia de paradas o puntos de interés en los Itinerarios de turistas (Flickr).	187

Tabla 41. Clústeres de valores altos y atípicos altos en el parque nacional Aigüestortes.	215
Tabla 42. Clústeres de valores altos y atípicos altos en el parque nacional Cabañeros.	216
Tabla 43. Clústeres de valores altos y atípicos altos en el parque nacional Tablas de Daimiel.	217
Tabla 44. Clústeres de valores altos y atípicos altos en el parque nacional Sierra de Guadarrama.	218
Tabla 45. Clústeres de valores altos y atípicos altos en el parque nacional Monfragüe.	219
Tabla 46. Clústeres de valores altos y atípicos altos en el parque nacional Ordesa y Monte Perdido.	220
Tabla 47. Clústeres de valores altos y atípicos altos en el parque nacional Picos de Europa.	221
Tabla 48. Clústeres de valores altos y atípicos altos en el parque nacional Sierra Nevada.	222
Tabla 49. Clústeres de valores altos y atípicos altos en el parque nacional Teide.	223
Tabla 50. Clústeres de valores altos y atípicos altos en el parque nacional Timanfaya.	224
Tabla 51. Clústeres de valores altos y atípicos altos en el parque nacional Caldera de Taburiente.	225
Tabla 52. Clústeres de valores altos y atípicos altos en el parque nacional Garajonay.	226
Tabla 53. Clústeres de valores altos y atípicos altos en el parque nacional Archipiélago de Cabrera.	227
Tabla 54. Clústeres de valores altos y atípicos altos en el parque nacional Islas Atlánticas de Galicia.	228
Tabla 55. Clústeres de valores altos y atípicos altos en el parque nacional Doñana.	229

Resumen

Los parques nacionales son destinos destacados para el turismo y la recreación debido a su valor estético y las oportunidades recreativas que ofrecen. Sin embargo, el crecimiento de la actividad turística en estos espacios puede generar presión sobre los ecosistemas que protegen. Bajo esta perspectiva, la investigación del uso turístico de los parques nacionales se ha establecido como una de las líneas prioritarias para la investigación sobre estos espacios (Eagles, 2014).

Tradicionalmente, los datos para analizar el uso público de los parques nacionales se obtienen de encuestas y datos provenientes de dispositivos de conteo de vehículos y personas, lo que puede ser muy exigente en términos de tiempo y costos. En este contexto, los datos de redes sociales, con su creciente volumen de datos disponibles, han surgido como una fuente confiable e interesante para comprender las interacciones humanas. Una línea emergente de investigación se ha desarrollado en torno al uso de datos geolocalizados de redes sociales para estudiar el comportamiento de los visitantes en los espacios naturales protegidos. A pesar de estos avances, existe un aspecto relevante que ha recibido menos atención; el aprovechamiento de la geolocalización en el uso de técnicas de análisis y estadística espacial.

El objetivo de esta tesis es identificar y cartografiar los distintos aspectos que modelan el uso turístico y recreativo de los espacios naturales protegidos a través del análisis geográfico y estadístico de los datos geolocalizados provenientes de dos redes sociales, Flickr y Wikiloc, y tomando como área de estudio la red de parques nacionales de España conformada por 15 parques que representan la diversidad de ecosistemas presentes en el territorio español.

En el desarrollo de esta investigación se han aplicado técnicas de análisis espacial y estadístico para extraer características de los visitantes y estudiar su comportamiento espacial y temporal en los parques nacionales. Entre los indicadores derivados de este análisis destacan la densidad de uso de las redes de senderos según el tipo de actividad, la identificación de puntos de interés con mayor afluencia y la distribución espacial-temporal de los visitantes. Además de las técnicas espaciales, se emplearon métodos de estadística descriptiva e inferencial para identificar los patrones temporales de los visitantes y obtener información sobre aspectos como su país de residencia.

En primer lugar, se exploró el lugar de origen de los visitantes, basándose en los perfiles de usuarios de Flickr, lo que ha permitido identificar que, en la mayoría de los parques nacionales, la

demanda proviene mayoritariamente de visitantes nacionales. Sin embargo, este patrón se invierte en los parques Canarios.

En segundo lugar, se identificaron los patrones temporales de visitación. Para ello, se utilizó el componente temporal almacenado en los metadatos de las fotografías geolocalizadas para determinar la distribución mensual, semanal y diaria de los visitantes de la red de parques nacionales. A partir de estas distribuciones se identificaron los patrones de estacionalidad turística en la red de parques nacionales y se aplicaron técnicas de agrupamiento para clasificar los parques de acuerdo con los patrones temporales de visita. En general los parques nacionales presentan tres patrones temporales; parques con estacionalidad alta (mayoría de visitas en una estación), parques con afluencia regular durante el año (estacionalidad baja) y parques con estacionalidad media, donde se presentan varios picos de visitas a lo largo del año.

En tercer lugar, se ha explorado la distribución espacial de los visitantes empleando la geolocalización de los datos de Flickr. Se han identificado puntos de interés con mayor o menor afluencia mediante indicadores de autocorrelación espacial local, evidenciando la existencia de clústeres con valores significativos, es decir, áreas de alta concentración de visitantes. Estos clústeres varían en tamaño y número, destacando los parques de Ordesa, Picos de Europa, Teide y Sierra de Guadarrama como aquellos en los que los clústeres abarcan una mayor extensión. Por último, esta distribución espacial se ha empleado como indicador de la demanda en modelos de localización-asignación, lo que ha permitido determinar las ubicaciones óptimas para puestos de información en el Parque Nacional Teide.

Finalmente, se llevó a cabo el análisis de los datos geolocalizados de Wikiloc para calcular la intensidad de uso de la red de senderos de dos actividades; senderismo y ciclismo. Los resultados revelan un uso desigual de las redes de senderos en todos los parques, permitiendo distinguir entre senderos altamente transitados y senderos subutilizados. Como aplicación práctica de estos datos, se han aplicado técnicas de análisis de redes para inferir los puntos de acceso de los visitantes y calcular los flujos de visitantes en los puntos de interés y senderos en el Parque Nacional Aigüestortes.

Los resultados obtenidos en el marco de esta investigación validan el potencial de las nuevas fuentes de datos para estudiar el comportamiento espacial y temporal de los visitantes de parques nacionales y abren oportunidades para desarrollar nuevas investigaciones.

Summary

National parks serve as prominent tourist and recreational destinations, owing to their aesthetic allure and the wide range of leisure opportunities they offer. However, the escalating influx of tourists in these areas places considerable stress on the ecosystems they are designed to protect. Consequently, research into tourism within national parks has gained recognition as a top research priority (Eagles, 2014).

Traditionally, data for analyzing public use of national parks has been gathered from surveys and data collected from vehicles and people counting devices. However, this approach can be resource-intensive both in terms of time and costs. In this context, social media data, with its burgeoning volume of accessible information, has emerged as a dependable and intriguing source for comprehending human interactions. A growing area of research has evolved around the utilization of geotagged data from social networks to analyze visitor behavior in protected natural spaces. Despite these advancements, one crucial aspect has received limited attention: the use of geotagged data for spatial analysis and statistics.

The main objective of this thesis is to identify and map the multifaceted factors that influence the tourist and recreational usage of protected natural areas through geographic and statistical analysis of geotagged data derived from two social networks, Flickr and Wikiloc. The research focuses on the Spanish network of national parks, comprising 15 parks that represent the diversity of ecosystems within the Spanish territory.

Throughout this research, spatial and statistical analysis techniques have been applied to extract insights about visitors and explore their spatial and temporal behavior within national parks. Key findings include the density of trail network usage based on different activities and the identification of tourism hotspots. In addition to spatial techniques, descriptive and inferential statistical methods have been employed to unearth temporal patterns of visitors, including their countries of origin.

Firstly, the investigation delves into the visitors' origins using Flickr user profiles, revealing that, in most national parks, most visitors are domestic. However, this pattern is inverted in the Canarian parks.

Secondly, temporal visitation patterns are identified. To achieve this, the temporal component embedded in the metadata of geotagged photographs is employed to discern the monthly, weekly, and daily distribution of visitors to the national park network. These distributions unveil patterns of tourist seasonality, categorizing parks into three main groups: those with high seasonality, those with consistent year-round visitation, and those with medium seasonality, characterized by multiple visitation peaks.

Thirdly, the research explores the spatial distribution of visitors utilizing geotagged data from Flickr. It identifies areas with varying visitor influx, employing local spatial autocorrelation indicators, uncovering clusters with notable concentration of visitors. These clusters differ in size and number, with parks like Ordesa, Picos de Europa, Teide, and Sierra de Guadarrama displaying larger cluster areas. This spatial distribution is then utilized as an indicator of demand in location-allocation models to pinpoint optimal locations for information posts within Teide National Park.

Finally, the analysis of Wikiloc GPS tracks is conducted to assess the intensity of trail network usage for hiking and cycling. The results reveal disparate usage levels of trail networks across all parks, helping distinguish highly traveled trails from underused ones. Network analysis techniques are subsequently applied to deduce visitor access points and compute visitor flows at points of interest in Aigüestortes National Park.

The outcomes achieved in the context of this research substantiate the promise of novel data sources for examining the spatial and temporal conduct of national park visitors. This not only reaffirms the reliability and relevance of these new data sources but also unveils new avenues for further research and exploration.

1. Introducción

1. Introducción.

1.1. Interés y oportunidad de la investigación.

A nivel mundial, los parques nacionales son considerados como un símbolo de orgullo nacional debido a su alto valor ecológico y cultural. Tienen un doble objetivo: conservar los ecosistemas que albergan y promover la difusión de sus valores ecológicos a través de la educación y el uso recreativo (Dudley, 2008). Este doble propósito, junto con su representatividad, ha hecho de los parques nacionales uno de los principales destinos para el turismo de naturaleza (Balmford et al., 2009). En España la actividad turística en los parques nacionales ha presentado un crecimiento notable en los últimos diez años, observándose un incremento del 34% en la afluencia a los parques hasta llegar a las 15,5 millones de visitas en el 2017 de acuerdo con los datos presentados por el Organismo Autónomo de Parques Nacionales (MAPAMA & OAPN, 2017). A pesar de la pandemia, la red de parques nacionales recibió 9,3 millones de vistas en 2020, según la última memoria del MAPAMA & OAPN (2020).

El turismo en los parques nacionales se ha configurado como una actividad que posibilita el desarrollo económico de las comunidades aledañas, la conservación de la naturaleza y la educación en valores ambientales a sus visitantes (Ballantyne et al., 2011; Scheyvens, 1999). Sin embargo, la intensificación del turismo en los parques nacionales supone presión creciente sobre los recursos que soportan esta actividad. Así, por ejemplo; el incremento de visitantes puede suponer un aumento de los desechos producidos dentro del área del parque y sus instalaciones (Eagles et al., 2002; Valentine, 1992). Asimismo, un mayor número de visitantes implicarían niveles más altos de ruido lo que afecta a la vida silvestre y la interrupción de sus ciclos naturales (Bultjens et al., 2005; Orams, 1996). Actividades turísticas como el senderismo y ciclismo también pueden provocar la erosión del suelo y la destrucción de la flora, afectando los servicios ecosistémicos que brindan, como la regulación del suelo (Farrell & Marion, 2001). De forma indirecta la actividad turística en los parques nacionales puede contribuir a los problemas ambientales globales como el cambio global ya que el transporte de los visitantes desde su lugar de procedencia involucra la emisión de CO₂ a la atmosfera (Lin, 2010).

Por lo tanto, aunque las actividades turísticas en los parques nacionales tienen resultados positivos, es indispensable medir y analizar el uso turístico de estos espacios a fin de contribuir a su gestión sostenible (Newsome et al., 2012). Este proceso involucra el seguimiento periódico de indicadores claves sobre las características de los visitantes. En este sentido, la investigación es un aliado

importante que puede proveer nuevos conocimientos y métodos para mejorar el monitoreo de la actividad turística en dichos espacios. Autores especializados en el turismo de espacios protegidos con extensa experiencia en el tema han establecido que el estudio del uso turístico y su monitoreo son líneas prioritarias para la investigación del turismo de naturaleza (Balmford et al., 2009; Eagles, 2014; Newsome et al., 2012).

De forma convencional, los datos utilizados en la investigación del turismo en los espacios naturales protegidos se han obtenido mediante métodos tradicionales como encuestas, entrevistas y últimamente con la toma de datos de receptores GPS (Eagles, 2014; Shoval & Ahas, 2016), lo que supone un alto costo tanto en recursos económicos como de tiempo y que son empleados en el diseño de muestreos, recolección y toma de datos (Balmford et al., 2015). Además, en la mayoría de los casos, los datos carecen de suficiente detalle y se limitan principalmente a conteos estáticos de visitantes (Schägner et al., 2016). Por ello se hace imperativo el estudio e identificación de nuevas fuentes de datos que permitan estudiar con mayor detalle y de forma continua el uso turístico de los parques nacionales.

El advenimiento de nuevas tecnologías de la información junto con una mayor accesibilidad a dispositivos móviles ha dado lugar a la generación masiva de datos provenientes de diversos canales, tales como, redes sociales, transacciones bancarias móviles, sistemas de geoposicionamiento, registros de llamadas, etc., y que se conocen como Big Data (Gandomi & Haider, 2015). Estos datos, en particular los provenientes de redes sociales, proporcionan nuevas oportunidades y desafíos para la investigación de distintos fenómenos sociales, ya que el uso de tales tecnologías implica la generación masiva de datos por parte de sus usuarios en tiempo real. Gran parte de estos datos poseen una referencia espacial y se los identifica como información geográfica voluntaria (IGV) (Goodchild, 2007). La IGV resulta de gran interés para la investigación geográfica, pues mediante su análisis es posible el identificar los patrones y distribuciones espaciales de distintos fenómenos sociales (Stienmetz & Fesenmaier, 2016). De ahí que, en los últimos años, se han publicado un número considerable de estudios que utilizan IGV para estudiar diferentes aspectos de la actividad turística y sus destinos. En este sentido, el turismo urbano ha sido el más explorado (García-Palomares et al., 2015; Kádár & Gede, 2013), dejando a otros tipos de turismo, como el de naturaleza pendiente.

Al momento, el uso de la IGV en la investigación del turismo en espacios protegidos se encuentra en un estado temprano que tuvo sus inicios en el 2013, por lo que los estudios son escasos y se han centrado en la validación de los datos como indicadores fiables de la afluencia de visitantes

(Sessions et al., 2016; Wood et al., 2013). Sin embargo, hay algunos aspectos relevantes que han sido menos explorados como el análisis espacio - temporal de la distribución de los visitantes de los parques nacionales, y que constituye el principal foco de interés de esta investigación. Concretamente en esta tesis se analizan las dimensiones espacial y temporal del uso turístico de la red de parques nacionales españoles a través de la IGV de dos redes sociales (Flickr y Wikiloc). Se recopila y procesa la IGV para extraer características de los visitantes como el país de residencia. Se identifican los patrones de estacionalidad turística en la red de parques nacionales. Se analiza la distribución espacial de los visitantes para determinar los puntos de interés con mayor y menor concentración turística. Y finalmente se integran las dimensiones espacial y temporal de la IGV para calcular los itinerarios y flujos de dos tipos de visitantes: los visitantes con objetivo de hacer turismo reflejados en los datos de Flickr y los visitantes con objetivo de realizar actividades deportivas como el senderismo y el ciclismo, reflejados en los datos de Wikiloc.

Al ser un campo emergente, los estudios han analizado con mayor profundidad la IGV proveniente de redes sociales de fotografía (Panoramio, Instagram y Flickr) para el análisis de datos de puntos enfocándose en el ritmo temporal (Sessions et al., 2016; Walden-Schreiner, Leung, et al., 2018) y representatividad de los turistas (Hamstead et al., 2018; Oteros-Rozas et al., 2017; Schägner et al., 2016; Sessions et al., 2016; Sinclair et al., 2020; Wood et al., 2013). Y han dejado de lado la dimensión espacial de este tipo de dato, que en esta investigación se desarrolla a profundidad. Con menor atención se ha estudiado la IGV proveniente de redes sociales especializadas en deportes al aire libre (Strava, GPSies y Wikiloc) y que consisten en datos de rutas GPS de dispositivos y teléfonos inteligentes, que se han utilizado para mapeo y cuantificación de diferentes actividades turísticas y medición del uso de senderos (Campelo & Nogueira Mendes, 2016; Gosal et al., 2019; Norman & Pickering, 2017).

Actualmente asistimos a un momento de cambio de paradigma, en el cual las nuevas fuentes de datos tienen un papel predominante ya que por un lado nos permiten analizar con mayor detalle los fenómenos y por otro nos demandan el desarrollo o adaptación de métodos y técnicas para el tratamiento de los datos.

Las oportunidades que se presentan al trabajar con las nuevas fuentes de datos son diversas. Entre ellas podemos destacar que la alta granularidad espacial y temporal que caracteriza a la IGV de redes sociales permite realizar análisis espaciales más precisos, visualizar el comportamiento espacial de los visitantes de una forma dinámica, con datos siempre actualizados y estudiar los patrones de la evolución temporal del turismo en los parques nacionales. El acceso libre a los datos

presenta una ventaja al reducir el costo por adquisición de los mismos y permite su difusión en la comunidad científica. Todo esto evidencia la importancia del desarrollo de líneas de investigación enfocadas en el análisis del turismo en espacios protegidos utilizando IGV de redes sociales.

1.2. Preguntas y objetivos de la investigación.

En el apartado anterior se ha presentado el interés y oportunidad del tema tanto para la investigación como para la gestión de espacios naturales protegidos. A continuación, se presentan las preguntas de la investigación que guiarán esta investigación y los objetivos a los que están ligadas.

1.2.1. Preguntas de investigación.

PI.1: ¿Por qué es relevante estudiar el uso turístico de los espacios protegidos?

Los espacios naturales protegidos desempeñan un rol notable tanto para la conservación del patrimonio natural como por los beneficios sociales y económicos que pueden proveer a través su oferta recreativa y de las actividades económicas asociadas a ella. Por ello estudiar el uso turístico de los espacios protegidos es importante para determinar el grado de presión que soportan y también para medir el beneficio potencial del turismo en dichos espacios. Sin embargo, los datos necesarios para estudiar a fondo el desarrollo del turismo en los espacios protegidos, son limitados. Aspectos como la distribución espacial de los turistas durante su visita y la intensidad de uso de los equipamientos de recreación se han investigado con menor detalle y presentan por lo tanto oportunidades para estudiar dichos aspectos bajo nuevos enfoques.

Desde el punto de vista de la gestión, el contar con nuevas metodologías de análisis del turismo en espacios protegidos podría ser un aporte crucial para los procesos de seguimiento de su uso público, ya que sin este seguimiento es imposible conocer la efectividad de la planificación y el cumplimiento de los objetivos establecidos en los planes de gestión (Eagles & McCool, 2002; Jones & Ohsawa, 2016). En el desarrollo de esta tesis, se explorarán los principales conceptos que definen la actividad turística en los espacios protegidos para poder identificar los aspectos potenciales de la actividad turística que se pueden investigar con las nuevas fuentes de datos geolocalizados.

PI.2: ¿Qué interés tiene el uso de fuentes de datos alternativas para el monitoreo del uso público de espacios protegidos como los parques nacionales?

Las nuevas fuentes de datos, entre ellas los datos geolocalizados de redes sociales han suscitado gran interés en varios campos científicos, debido a las ventajas que presentan en términos de resolución espacial y temporal, volumen y acceso (Goodchild & Li, 2012). En la actualidad, los datos geolocalizados se configuran como un nuevo recurso de la investigación de fenómenos sociales como el turismo (Ferreira, 2019). Sin embargo, en el caso del turismo de naturaleza, el uso de IGV es un tema aún en desarrollo, que ofrece la oportunidad de examinar aspectos poco estudiados como el comportamiento espacial y temporal de los visitantes de parques nacionales.

PI.3: ¿Qué fuentes de datos de redes sociales son más apropiadas para el análisis geográfico del turismo en espacios naturales protegidos?

Existe un número considerable y creciente de redes sociales en el internet, por lo que la selección de las fuentes adecuadas es fundamental. En el campo de investigación del turismo, los datos más utilizados han sido aquellos pertenecientes a las redes de Twitter y Flickr. Al inicio de esta investigación se plantearon 3 fuentes de datos; Flickr, Wikiloc y Twitter, asumiendo que cada una nos ofrecería diferentes aspectos del turismo que explorar y que su aplicabilidad dependería también del número de registros que se puedan obtener. A lo largo de esta investigación se planifica examinar los datos de cada fuente y en base a los resultados seleccionar las fuentes más apropiadas para el estudio del turismo en espacios naturales protegidos.

PI.4: ¿Permiten los datos de Flickr medir e identificar el ritmo temporal del turismo en la red de parques nacionales?

Es relevante conocer y estudiar la distribución temporal de los visitantes de la red de parques nacionales, porque nos proporciona indicadores sobre la evolución de la afluencia de los visitantes y establecer si los espacios están ganando influencia y notoriedad como destinos turísticos. Sin embargo, las distintas fuentes de datos de redes sociales presentan una diferente resolución temporal, por lo que pueden ser más o menos adecuadas para la realización de análisis temporales. Fuentes de datos con mayor resolución temporal deberían permitir obtener resultados más precisos a través de la desagregación de los datos a diferentes escalas temporales y de esta

forma obtener las distribuciones mensuales, semanales y diarias de la afluencia de los visitantes de la red de parques nacionales.

PI.5: ¿Puede el componente espacial de los datos de Flickr usarse como indicador de la distribución espacial del turismo en la red de parques nacionales?

Visualizar la distribución espacial de los visitantes en los parques nacionales es relevante para detectar patrones de concentración o dispersión espacial de la actividad turística. La geolocalización de la IGV podría ser utilizada como un indicador de cuáles son los lugares de interés de los visitantes, detectar nuevos sitios de interés y cuantificar el uso (densidad) de cada espacio.

El comportamiento espacial de los turistas y visitantes es un campo poco explorado debido a la falta de datos que permitan examinar la distribución geográfica del visitante o turista. En este sentido, se propone extraer la distribución geográfica de los usuarios de Flickr y utilizar herramientas de análisis geográfico para entender los patrones espaciales del uso turístico de los parques. El análisis geo visual (Adrienko & Andrienko, 2011) y el análisis exploratorio de datos espaciales, ESDA por sus siglas en inglés, (Anselin, 1999) son dos enfoques metodológicos que se han encargado de comprender e identificar patrones en grandes conjuntos de datos espaciales, y que se utilizaran en esta investigación.

PI.6: ¿Pueden los datos de Flickr y Wikiloc ser indicadores de la movilidad turística dentro de los parques nacionales?

La movilidad de los visitantes de espacios naturales protegidos es un tema poco explorado, debido a la falta de datos que permitan contar con un seguimiento de los visitantes durante su visita al parque. En los últimos años, la investigación de la movilidad turística se ha apoyado en datos provenientes de receptores GPS (Orellana et al., 2012; Shoval & Isaacson, 2009). Sin embargo, tanto el diseño del estudio como la toma de datos están limitados a la disponibilidad de recursos. La IGV de Flickr y Wikiloc podría utilizarse para examinar distintos aspectos de la movilidad turística como: construcción de itinerarios, estimar flujos de visitantes desde accesos y medir el uso de la red de senderos de cada parque.

PI. 7: ¿Qué tipo de aplicaciones prácticas para el manejo y planificación de espacios pueden ofrecerse a partir de la huella digital?

La formulación de esta pregunta se basa en la idea de explorar que tipos de aplicaciones para la gestión del turismo en espacios protegidos se pueden llevar a cabo usando la IGV de Flickr y Wikiloc. Esta pregunta se responderá exponiendo dos usos potenciales de la IGV en aplicaciones prácticas para la gestión del uso turístico. En primer lugar, se propone desarrollar una aplicación que hace uso de modelos de localización óptima para determinar las localizaciones y la demanda de equipamientos de uso público como puntos de información turística. Una segunda aplicación consistirá en la cuantificación de los flujos de visitantes desde diferentes puntos de acceso de un parque nacional, mediante el uso de análisis de red y minería de patrones secuenciales.

1.2.2. Objetivos.

Objetivo general

Esta tesis parte de la hipótesis de que los datos geolocalizados de redes sociales generados por sus usuarios cuando visitan un espacio protegido pueden representar el uso turístico de los espacios naturales protegidos ya que son registros continuos de las actividades y experiencias que los usuarios escogen publicar. Por lo tanto, el propósito de esta investigación es el evaluar la utilidad de las nuevas fuentes de datos geolocalizados, surgidas a partir de la revolución del Big data, para estudiar el uso turístico de los espacios naturales protegidos. Las nuevas fuentes de datos geolocalizados se han abierto campo en la investigación de diferentes fenómenos geográficos como la movilidad y el turismo y presentan oportunidades de estudio para analizar bajo nuevas perspectivas el uso turístico en los parques nacionales.

De esta forma, se formula el objetivo general de identificar los distintos aspectos que modelan el uso turístico y recreativo de los espacios naturales protegidos a través del análisis geográfico y estadístico de la IGV proveniente de dos redes sociales, Flickr y Wikiloc, tomando como área de estudio la red de parques nacionales de España conformada por 15 parques que representan la diversidad de ecosistemas presentes en el territorio español.

Objetivos específicos

A partir del objetivo general, se plantean varios objetivos específicos que permitirán abordar la temática estudiada y cumplir con el propósito de esta investigación.

OE 1: Construir un marco teórico sobre el desarrollo de la actividad turística en los espacios protegidos con énfasis en los parques nacionales.

Para sentar las bases de la tesis es indispensable como punto de partida, establecer un marco teórico actualizado sobre las dos temáticas en las que se enmarca esta investigación. En primer lugar, la temática del turismo de naturaleza, en la que se abordará de forma general el concepto de turismo de naturaleza y su relación con el uso público de los espacios protegidos, para luego ahondar en los parques nacionales como destino de turismo de naturaleza, los impactos positivos y negativos de esta actividad y la capacidad de carga turística.

La segunda temática que tratar tiene que ver con las nuevas fuentes de datos surgidas de la llamada revolución de los datos y su uso como fuente de datos para la investigación. Para ello se revisarán las diferentes fuentes de Big data, enfatizando en el Big data geolocalizado proveniente de redes sociales y su aplicación en la investigación del turismo.

OE 2: Realizar una revisión del estado del arte sobre el uso de datos geolocalizados de redes sociales para el análisis del turismo en espacios naturales protegidos.

A la par del marco teórico que define de forma global los fundamentos científicos en los que se basa esta investigación, se debe realizar una revisión exhaustiva y actualizada sobre los estudios que utilizan y analizan datos geolocalizados de redes sociales para el análisis de la demanda y uso turístico de los espacios naturales protegidos. Entrando en detalle, se examinará y sistematizará qué tipo de fuentes se han utilizado al momento, cuáles han sido los enfoques metodológicos aplicados, las herramientas utilizadas y las distintas aplicaciones que se han propuesto para aprovechar el potencial de esta nueva fuente de datos.

OE 3: Recopilar y depurar los datos geolocalizados de Flickr y Wikiloc.

La recopilación de datos procedentes de distintas redes sociales comprende varios procesos que van desde la selección de las redes sociales, el acceso a los datos, el desarrollo de métodos de descarga adecuados a cada fuente, su depuración y transformación a un formato común que permita su análisis en un sistema de información geográfica y finalmente su almacenamiento en una base de datos geográfica. La recolección de datos de las dos fuentes propuestas y su análisis exploratorio posterior permitirá evaluar la factibilidad de uso de cada fuente, al igual que la selección de aspectos relacionados con el uso público que se puedan extraer de las mismas.

OE 4: Analizar el contenido temporal de los datos de Flickr para identificar patrones temporales de los visitantes y construir indicadores de uso turístico como estacionalidad y distribución mensual y diaria.

Los datos provenientes de redes sociales se caracterizan por tener una alta granularidad temporal. Se cuenta con la marca de tiempo o “timestamp” de cada registro o contenido que un usuario publica en su cuenta. Esta marca de tiempo contiene el año, mes, día y hora de cada registro por lo que resultará útil para explorar y analizar los patrones temporales que caracterizan las visitas a los parques nacionales. En este sentido, se pretende agregar temporalmente los datos de fotografías geolocalizadas para conocer los ritmos de utilización de los parques naturales, analizando la distribución diaria, semanal y mensual de los visitantes. A partir del indicador usuarios únicos diarios se agregará la información a nivel de mes y año para calcular índices de estacionalidad y clasificar a los parques de acuerdo con su grado de estacionalidad. Y a partir del indicador de usuarios por hora se realizarán las visualizaciones del comportamiento diario de los visitantes, pudiéndose identificar los periodos de entrada, salida y de mayor concentración de visitantes en el día.

OE 5: Utilizar herramientas de análisis geográfico para evaluar la distribución espacial de los visitantes y los patrones de concentración y dispersión de la actividad turística.

El conocimiento de la distribución espacial de los visitantes de los parques naturales es de gran interés para la gestión de estos espacios. El análisis de datos geolocalizados de fotografías compartidas mediante herramientas de análisis exploratorio de datos espaciales (Anselin, 1999) constituye una interesante oportunidad para extraer; la distribución espacial de los usuarios, los

lugares donde se concentran la mayoría de los usuarios y la densidad de los usuarios dentro de cada parque. Para lograr este objetivo se emplearán las técnicas como el índice de autocorrelación espacial global y local, para identificar en primer lugar el grado de concentración de visitantes y luego las localizaciones donde se producen, y el análisis de densidad Kernel para estimar la densidad de visitantes en el territorio de cada parque.

OE 6: Aprovechar la granularidad espacial y temporal de Flickr y Wikiloc para identificar itinerarios de los visitantes de parques nacionales.

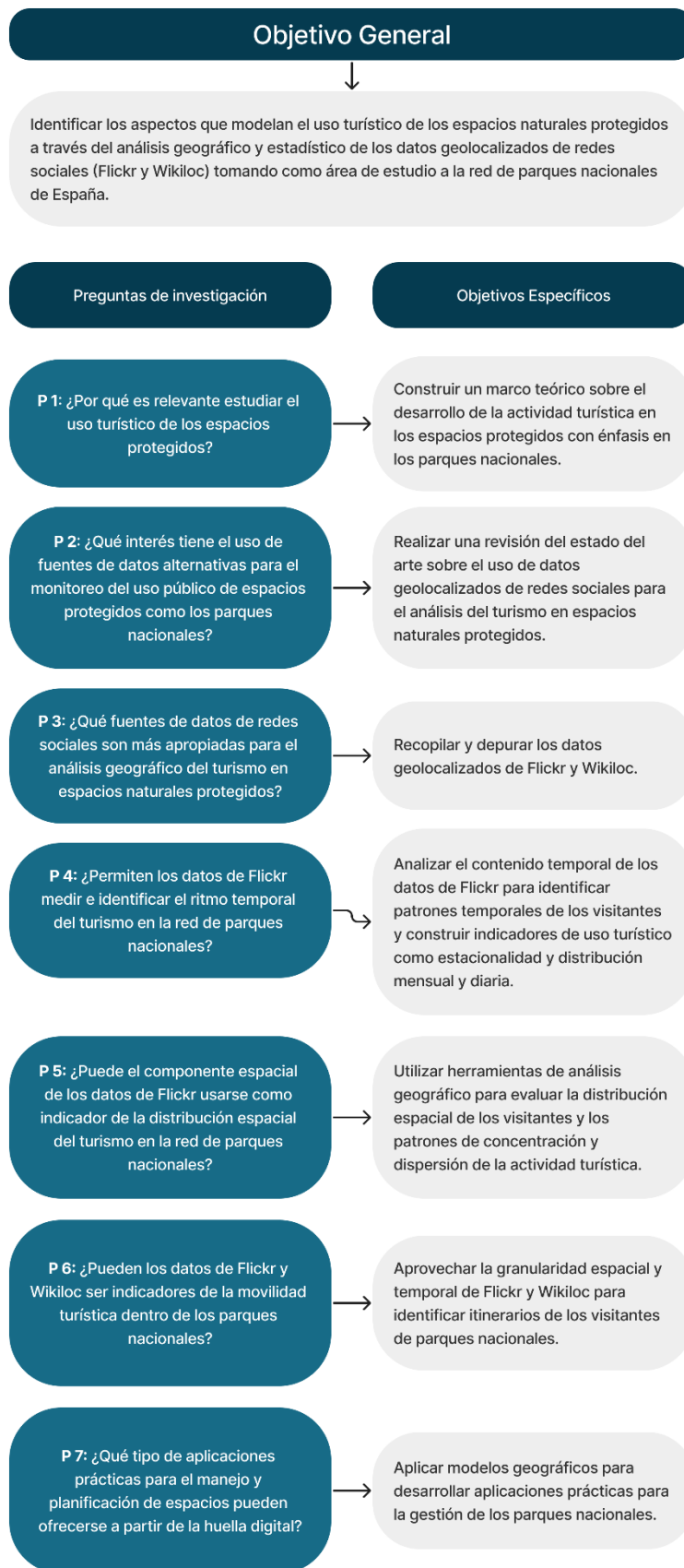
Un aspecto importante para la investigación del turismo en parques nacionales consiste en conocer las pautas de movilidad de los visitantes dentro de los parques, un aspecto hasta ahora poco estudiado y en todo caso con muestras pequeñas obtenidas a partir de dispositivos GPS repartidos entre visitantes de los parques. Un enfoque alternativo y de bajo coste es utilizar datos de redes sociales, que producen datos geolocalizados con alto nivel de detalle espacial y temporal, lo que permite hacer un seguimiento de los usuarios de estas redes durante su visita al parque y de esta forma reconstruir la secuencia y orden de los lugares visitados.

OE 7: Aplicar modelos geográficos para desarrollar aplicaciones prácticas para la gestión de los parques nacionales.

Como último objetivo empírico, se propone explorar las potenciales aplicaciones de la IGV para soluciones de gestión de los parques utilizando herramientas de análisis espacial. Se aplicarán técnicas de modelado geográfico como los modelos de localización óptima y el análisis de distancia cercana para identificar ubicaciones óptimas de nuevos puntos de información dentro del territorio de un parque nacional y en una segunda aplicación se obtendrán los flujos de visitantes desde los puntos de acceso de los parques.

La Figura 1 muestra la relación entre las preguntas de investigación y los objetivos de la tesis, y explica la estructura del documento que se describe en el siguiente apartado.

Figura 1. Diagrama de Preguntas de investigación alineadas con los objetivos de la investigación



1.3. Estructura de la tesis.

La presente tesis doctoral está estructurada en diez capítulos, que a su vez están organizados en secciones para presentar de forma organizada el trabajo realizado. El primer capítulo es el punto de partida de la investigación, y consiste en un bloque introductorio que presenta el interés y oportunidad de utilizar nuevas fuentes de datos provenientes del Big Data para el estudio del turismo en parques nacionales. En este capítulo se definen las preguntas de investigación y los objetivos que se han propuesto cumplir en esta tesis.

El segundo capítulo comprende el marco teórico, dividido en dos bloques temáticos. El primer bloque aborda el tema del turismo de naturaleza, centrándose en el turismo en espacios protegidos y profundizando en el origen de la protección de los espacios, el desarrollo de la actividad turística en dichos espacios y los efectos positivos y negativos que esta actividad produce. El segundo bloque se centra en las nuevas fuentes de datos, haciendo un recorrido desde el concepto de Big Data y los datos de redes sociales para luego ahondar en los datos geolocalizados de redes sociales y terminar exponiendo sus aplicaciones en el campo del turismo.

El tercer capítulo remata el marco teórico con una revisión exhaustiva del estado del arte sobre el uso de los datos geolocalizados de redes sociales en la investigación del turismo en espacios naturales protegidos. En la revisión se examina a detalle; los diversos tipos de redes sociales utilizadas, los aspectos del uso turístico examinados y los enfoques metodológicos que han propuesto en la literatura hasta el momento de elaboración de esta tesis. El capítulo concluye con una visión general de las ventajas y limitaciones del uso de estos datos en el análisis del turismo de espacios naturales protegidos. Esta revisión forma parte un artículo publicado en una revista de alto impacto (Q1) titulado: “Geotagged data from social media in visitor monitoring of protected areas; a scoping review” y que se publicó en 2021.

El cuarto capítulo consiste en la presentación del área de estudio, la red de parques nacionales españoles, compuesta por 15 espacios. Esta sección consiste en una breve descripción de las características de cada uno de los parques y su situación en relación con la actividad turística, con el fin de dar un contexto general para entender las características de estos espacios y su interés para esta investigación.

En el quinto capítulo se desarrolla la metodología empleada para la obtención y preparación de los datos geolocalizados provenientes de las dos fuentes clave en esta investigación: Flickr y Wikiloc.

En la primera sección del capítulo se detallan las fuentes y datos utilizados en esta tesis, explicando el acceso y descarga de los datos geolocalizados de las redes sociales Flickr y Wikiloc y también describiendo los datos secundarios, provenientes de fuentes como el IGN, OAPN y Open Street Maps que se utilizaron para la elaboración de cartografía y análisis espaciales. En el segundo apartado se describen los procesos de depuración y preparación de los datos para su uso en diferentes análisis espaciales. Finalmente, en la última sección, se proporciona una breve descripción de los datos una vez depurados, apoyándose en estadísticas descriptivas que permiten arrojar luz sobre la naturaleza y características de los datos geolocalizados obtenidos.

En el sexto capítulo se expone el análisis de los datos geolocalizados de Flickr para el estudio del origen de los visitantes a la red de parques nacionales. En este contexto, se desarrolla una metodología precisa para extraer y determinar el lugar de residencia de los visitantes de los parques nacionales españoles. Este capítulo se adentra en la exploración de la procedencia geográfica de los visitantes a los parques nacionales a través de sus actividades en Flickr

En el capítulo siete, se explora el componente temporal de los datos geolocalizados de Flickr para identificar los patrones temporales de la actividad turística en los parques nacionales a distintas escalas. Este capítulo se centra en la comprensión de cómo los patrones de actividad de los visitantes evolucionan en el tiempo, permitiendo una visión detallada de cómo se distribuye la afluencia turística en los parques nacionales a lo largo de diferentes momentos y períodos. Los resultados obtenidos de este capítulo forman parte de un artículo publicado en una revista de alto impacto (Q2) titulado: "Identifying Temporal Patterns of Visitors to National Parks through Geotagged Photographs" en 2020.

El capítulo ocho detalla la aplicación de técnicas de estadística espacial para identificar patrones espaciales del uso turístico en los parques nacionales a partir de datos geolocalizados de Flickr.

A través de la representación cartográfica y la disposición de información detallada, se expone de manera visual el comportamiento espacial de los visitantes en los parques nacionales, así como la identificación de los puntos de mayor interés dentro de dichos enclaves naturales. Parte de los resultados obtenidos en este apartado, se han publicado como un artículo científico en una revista de alto impacto (Q1) que se titula: "Using geotagged photographs and GPS tracks from social networks to analyse visitor behaviour in national parks" en 2020. El capítulo concluye con una aplicación práctica que aprovecha los datos de Flickr con la finalidad de identificar ubicaciones potenciales para nuevos puntos de información dentro de los confines de un parque nacional. Para llevar a cabo esta aplicación, se emplean modelos de localización óptima. El parque nacional Teide

actúa como el terreno de estudio, en el cual se exploran múltiples escenarios, variando la cantidad de nuevas ubicaciones, al mismo tiempo que se evalúa la eficacia de estas al abordar las necesidades informativas de los visitantes.

El capítulo nueve explora la movilidad de los visitantes dentro de la red de parques nacionales, usando los datos geolocalizados de Wikiloc. A través del análisis de redes y ajuste de las rutas descargadas de Wikiloc, se investiga y cuantifica el desplazamiento de los visitantes en seis parques nacionales. Este análisis se centra en las rutas generadas por actividades de senderismo y ciclismo. A través de la creación de mapas que representan la intensidad de uso, se logra exponer de manera visual y cuantitativa los patrones de comportamiento y los itinerarios seguidos por los visitantes en los parques nacionales. Como una culminación a este capítulo, se presenta una aplicación concreta de los datos obtenidos de Wikiloc en conjunción con información proveniente de Flickr. Esta aplicación tiene como objetivo analizar y representar los flujos de visitantes que se dirigen desde distintos puntos de acceso hacia un parque nacional en específico. Esta aplicación es llevada a cabo en el marco del parque nacional de Aigüestortes, utilizando los flujos de visitantes como una herramienta esencial para comprender las dinámicas de entrada y desplazamiento dentro del parque.

Finalmente, el capítulo diez discute los resultados obtenidos en relación con las preguntas de investigación y objetivos específicos planteados en el capítulo introductorio, señalando el trabajo realizado para la consecución de los objetivos y dar respuesta a las preguntas propuestas en esta investigación. A su vez, en este capítulo se presentan las conclusiones generales de este trabajo, destacando los principales hallazgos, la aplicación práctica de la metodología para la toma de decisiones sobre gestión del uso turístico y también indicar las limitaciones de uso de las nuevas fuentes de datos en el contexto del turismo de espacios protegidos. En el apartado final de este capítulo se presentan las futuras líneas de investigación que se han encontrado a partir del desarrollo de este trabajo, por ejemplo, la utilización de nuevas técnicas de minería de datos para analizar los componentes semántico y gráfico de los datos geolocalizados.

1.4. Artículos y Comunicaciones

Los resultados más significativos obtenidos durante el desarrollo de la investigación han sido publicados y difundidos en la comunidad científica en diferentes formatos.

Artículos:

- Barros, C., Gutiérrez, J., & García-Palomares, J. (2022). Geotagged data from social media in visitor monitoring of protected areas; a scoping review. *Current Issues in Tourism*, 0(0), 1–17. <https://doi.org/10.1080/13683500.2021.1931053>.

Indicios de Calidad: JCR Impact factor 2020: 7.430, Q1/D2 - 8/58 SSCI Hospitality, Leisure, Sport & Tourism. Número de citas: 6 en Google Scholar, 9 en Scopus.

- Barros, Carolina, Moya-Gómez, B., & Gutiérrez, J. (2020). Using geotagged photographs and GPS tracks from social networks to analyse visitor behaviour in national parks. *Current Issues in Tourism*, 23(10), 1291–1310. <https://doi.org/10.1080/13683500.2019.1619674>.

Indicios de Calidad: JCR Impact factor 2020: 7.430, Q1/D2 - 8/58 SSCI Hospitality, Leisure, Sport & Tourism. Número de citas: 53 en Google Scholar, 41 en Scopus.

- Barros, Carolina, Moya-Gómez, B., & García-Palomares, J. C. (2019). Identifying Temporal Patterns of Visitors to National Parks through Geotagged Photographs. *Sustainability*, 11(24), 1–16. <https://doi.org/10.3390/su11246983>.

Indicios de Calidad: JCR Impact factor 2019: 2.576, Q2/D5 - 53/123 en SSCI Environmental Studies. Número de citas: 21 en Google Scholar, 14 en Scopus

Comunicaciones en congresos:

- Barros C.(2017): Análisis de la Demanda Turística de los Parques Nacionales a través de las Redes Sociales. II-Taller de doctorado en Geografía. XXV Congreso de la Asociación de Geógrafos Españoles. 22- 31. ISBN: 978-84-944193-8-6.
<https://www.age-geografia.es/site/wp-content/uploads/2018/01/II-Taller-de-Doctorado-en-Geograf%C3%ADa.pdf>

- Barros C, Gutiérrez, J. (2018): El uso turístico de los parques nacionales a partir de redes sociales de fotografías geolocalizadas: el Parque Nacional del Teide. ISBN 9788491331469, págs. 710-720.
https://congresos.adeituv.es/imgdb//archivo_dpo62599.docx.
- Barros C, (2021): Modelling visitor's movements of nature-based tourism from social media data. The 10th MMV Conference: Managing ISSN 2535-2806. Pp. 340-342.
https://static02.nmbu.no/mina/publikasjoner/mina_fagrappport/pdf/mif73.pdf.

2. Turismo en espacios protegidos y nuevas fuentes de datos.

2. Turismo en espacios protegidos y nuevas fuentes de datos.

La aceleración de la urbanización en los últimos años ha propiciado un estilo de vida apegado a la ciudad y alejado de la naturaleza. Este fenómeno ha generado nuevas necesidades de recreación que demandan actividades y lugares que permiten el contacto directo con la naturaleza (Newsome et al., 2012). En este contexto, el turismo de naturaleza se enmarca como una actividad atractiva para satisfacer las necesidades de recreación y apreciación del medio natural. Por ello, los espacios naturales protegidos juegan un papel importante en el desarrollo de esta actividad ya que por sus características y nivel de conservación se han configurado como destinos predilectos por los turistas.

Sin embargo, el desarrollo de esta actividad puede implicar conflictos entre los objetivos de conservación y el uso público sobre todo cuando la demanda turística se incrementa, llegando a poner en riesgo la capacidad de carga de los espacios. De igual forma, la concentración espacial y temporal de los visitantes en los parques puede ocasionar problemas relacionados con la satisfacción de la visita y conflictos con las poblaciones aledañas a los parques.

Por lo tanto, es necesario contar con información actualizada y detallada sobre el uso público de los espacios protegidos con el fin de prevenir los impactos negativos derivados del uso público y potenciar los impactos positivos que benefician a la población. La información sobre el uso público de estos espacios es variable y depende de los recursos disponibles para cada área protegida, por lo que existen diversos limitantes a la hora de analizar el comportamiento de los visitantes.

Por otro lado, el uso generalizado de dispositivos integrados con sensores de localización como los teléfonos inteligentes, ha permitido a las personas compartir activamente su ubicación cuando usan plataformas de redes sociales (Hecht et al., 2011). Generando una gran cantidad constante de datos geolocalizados que pueden ser recopilados y utilizados para analizar el comportamiento de la población. Desde la perspectiva de la investigación del turismo en espacios protegidos, estas nuevas fuentes de datos se presentan como alternativas que podrían ayudar a caracterizar a los visitantes, determinar la demanda turística de diferentes espacios protegidos y complementar los datos tradicionales que se obtienen de la gestión de los mismos.

Este capítulo tiene como objetivo presentar los conceptos generales bajo los que se enmarca el tema de esta tesis. Por un lado, en esta investigación se aborda el tema del turismo en parques nacionales, el mismo que se contempla bajo el concepto de uso público en la regulación de los espacios naturales protegidos españoles. Por lo que en este capítulo se presentaran varios

conceptos que definen a la actividad turística en espacios protegidos, poniendo énfasis particular en los parques nacionales.

La segunda temática que aborda esta tesis es la aparición de las nuevas fuentes de datos originadas a partir de la revolución de los datos y la digitalización de las actividades. En esta segunda temática se describen los conceptos clave de datos masivos o Big Data, los datos geolocalizados generada a partir de la actividad de los usuarios de redes sociales y como se han utilizado para la investigación del turismo en general.

2.1. Los espacios protegidos como destino turístico.

2.1.1. La protección del medio natural - breve historia.

La idea de proteger y conservar lugares con paisajes únicos y excepcionales es universal y se encuentra presente en diversas culturas a lo largo de la historia. En la India hace más de 2000 años se establecieron “áreas de reserva” para la protección de recursos naturales como bosques, cuerpos de agua, etc. (Holdgate, 1999). Al mismo tiempo en China y Nepal, los bosques antiguos eran seleccionados para su resguardo. Las tradiciones de poblaciones locales delimitaron áreas especiales de protección como es el caso de los árboles sagrados en Ghana y las zonas “Tapu” en el Pacífico (Grommbridge, 1992). En tanto que, en Europa la protección de áreas naturales estuvo asociada al disfrute y recreo de la clase gobernante por lo que el objetivo era reservar áreas especiales para cotos de caza y jardines para los reyes y otros gobernantes (Eagles et al., 2002). Estos ejemplos reflejan que en un primer momento la protección de espacios naturales estuvo ligada a razones culturales y sociales más que a la preservación y conservación del medio ambiente per se.

Con el paso del tiempo y sumado al nacimiento y desarrollo del movimiento ecologista, estos sitios se han abierto para el uso y disfrute público, pasando de ser lugares exclusivos y de resguardo de recursos para una clase privilegiada hacia territorios de valor nacional cuyo propósito es conservar intacta su biodiversidad para el disfrute y admiración de la población presente y futura (Troitiño, 1995). De tal manera que ya a mediados del siglo XIX, cinco países se posicionaron como pioneros en la creación y protección de espacios naturales; Estados Unidos, Australia, Canadá, Sudáfrica y Nueva Zelanda, principalmente bajo la figura de parques nacionales.

Bajo las ideas románticas de conservación y convivencia armónica entre el ser humano y la vida animal y vegetal expresadas inicialmente por el poeta George Catlin, en 1872, el congreso de Estados Unidos crea el parque Yellowstone, el primer parque nacional de Estados Unidos con el

objetivo de dotar de un espacio natural para el uso y recreación de la población. A partir de esta designación, surgirán en todo el territorio americano varios parques nacionales tales como El Gran Cañón, Yosemite y El monte Rainier (West Sellars, 2009). Igualmente, en 1866 en Australia, se designaron a las cuevas Jenolan y las 2000 hectáreas circundantes como sitio protegido para su conservación y turismo. Adiciones de tierra posteriores llevaron a la creación del parque nacional Blue Mountain (Eagles et al., 2002). En 1885, Canadá tomando como referencia la legislación del parque Yellowstone, estableció el parque nacional Banff para proteger las fuentes de aguas termales del valle de Bow (Lothian, 1987). A partir de 1893, en Sudáfrica varias reservas forestales son creadas. Finalmente, en Nueva Zelanda, en 1894 se creó el parque nacional Tongariro mediante un acuerdo entre el gobierno y la comunidad Maori. Siendo hasta tiempos actuales un lugar de gran importancia espiritual para este grupo.

Estos primeros parques se crearon bajo ideas y atributos comunes. En primer lugar, su creación se debe a la iniciativa de los gobiernos nacionales de cada país. En segundo lugar, son espacios que se caracterizaban por tener una vasta superficie. En tercer lugar, estos sitios se dedicaron a la protección de un monumento natural determinante. Finalmente, estos parques se establecieron para el uso público de toda la población sin restricciones (West Sellars, 2009). Siendo así, que desde su origen se ha establecido a la visita de estos sitios como uno de los fundamentos primordiales para su creación y ha sentado las bases para su posterior organización y estructura en redes integrales de protección de ecosistemas.

A medida que el número de espacios protegidos crecía se hace evidente la necesidad de contar con un sistema de gestión y monitoreo de las mismas. En 1911, Canadá estableció el primer organismo mundial de parques, el Dominion Parks Bureau (DPB). En Estados Unidos, el servicio nacional de parques (USNPS) se fundó en 1916. En el mismo año, en Europa, España se situó como pionera de la protección de la naturaleza en el continente al aprobar la primera ley de parques nacionales en 1916, teniendo como modelo el parque nacional de Yellowstone en EE. UU (OAPN, 2011).

La expansión de espacios protegidos a nivel mundial toma forma en el siglo XX, cuando varios países alrededor del mundo empiezan a crear áreas protegidas dentro de sus territorios. Desde 1940 crece la iniciativa por crear áreas protegidas imitando los modelos canadienses y estadounidenses a nivel global. Sin embargo, la idea de formar una red integral de espacios protegidos tardó en desarrollarse hasta la década de 1940 cuando se comenzó a establecer un mayor número de espacios protegidos a nivel mundial (Eagles & McCool, 2002). Para 1962, se observó un alto incremento en el número de espacios protegidos, a raíz de la celebración del primer congreso mundial de parques en Seattle, EE. UU., ya que esta reunión significó el

establecimiento del 80% de las áreas protegidas actuales (Eagles et al., 2002). En la actualidad, la mayoría de los países en el mundo han aprobado legislaciones sobre los espacios naturales protegidos y para 2002 se contaba con cerca de 44000 espacios protegidos que representan el 10% de la superficie del planeta.

2.1.2. El turismo de naturaleza y el uso público de los espacios naturales protegidos.

La relación entre naturaleza y turismo es tan antigua como la historia de la conservación ambiental (Frost & Hall, 2009). El primer parque nacional "Yellowstone" fue creado no solo para cuestiones de conservación sino también para permitir la recreación y la apreciación de la naturaleza.

El "turismo de naturaleza" o ecoturismo se refiere a todas las actividades que giran en torno a la observación y apreciación de la naturaleza (Amo et al., 2006; Buckley, 2008). De acuerdo con la Organización Mundial del Turismo, el turismo de naturaleza es un tipo de actividad turística en el que la motivación esencial del visitante es observar, aprender, descubrir, experimentar y apreciar la diversidad biológica y cultural con una actitud responsable para proteger la integridad del ecosistema y fomentar el bienestar de la comunidad local (UNWTO, 2021).

Por otro lado, el turismo de naturaleza también se considera como un servicio ecosistémico cultural, ya que nos proporciona beneficios no materiales como la inspiración estética, la identidad cultural, recreación y la experiencia espiritual relacionada con el entorno natural (Paracchini et al., 2014).

El interés inicial en el turismo y la recreación en áreas naturales estaba relacionado con una idea romántica centrada en la búsqueda de significado espiritual y mental para las personas (Torres-Delgado & Saarinen, 2014). En general, la motivación de los primeros entusiastas de la naturaleza por visitar áreas naturales fue apreciar los elementos de la naturaleza y buscar la soledad. La demanda de turismo basado en la naturaleza creció rápidamente a fines de la década de 1980, lo que llevó a la creación de nuevos productos y tipos de turismo basados en los espacios naturales protegidos y el tipo de experiencias que ofrecen. Hoy en día la naturaleza es una importante atracción turística y los parques nacionales se encuentran entre los destinos más visitados para este tipo de turismo. Además, la Organización Mundial del Turismo señala que el turismo basado en la naturaleza es el sector de mayor crecimiento de esta actividad (Balmford et al., 2009). Los espacios naturales protegidos y particularmente los parques nacionales se han convertido en el principal destino para los turistas de naturaleza.

En España, las actividades turísticas que se pueden llevar a cabo en los espacios naturales protegidos se enmarcan bajo el concepto de uso público y el mismo establece el tipo de actividades permitidas y su desarrollo. EUROPARC define al uso público como: “El conjunto de programas, servicios, actividades y equipamientos que, independientemente de quien los gestione deben ser provistos por la Administración del espacio protegido con la finalidad de acercar a los visitantes a los valores naturales y culturales de éste, de una forma ordenada, segura y que garantice la conservación, la comprensión y el aprecio de tales valores a través de la información, la educación y la interpretación del patrimonio ” (Hernández de la Obra & Gómez Limón García, 2005).

Desde el punto de vista de la gestión, la Consejería del Medio Ambiente y el Organismo Autónomo de Parques Nacionales, son los entes que establecen las directrices de uso público y que se articulan a través de los Planes Rectores de Uso y Gestión.

2.1.3. El turismo en parques nacionales y sus impactos.

Los parques nacionales y el turismo se encuentran estrechamente relacionados pues una de las bases para la creación de estos espacios se apoya en su misión como lugares para la recreación y educación de las personas que los visitan. El turismo en los parques nacionales se ha configurado como una actividad económica idónea y que posibilita desarrollo económico de las comunidades aledañas, la conservación de la naturaleza y la educación en valores de la naturaleza a sus visitantes (Ballantyne et al., 2011; Scheyvens, 1999).

Por lo tanto, el turismo se convierte en un componente esencial tanto para la declaración de estos espacios como para su manejo y planificación (Eagles & Mccool, 2002). El turismo en parques nacionales tiene incidencia en aspectos fundamentales de los parques nacionales. Por ejemplo, desde la perspectiva ambiental, el turismo se configura como una herramienta de información y educación sobre la importancia de la conservación ambiental, el conocimiento de los valores naturales y culturales de cada parque y la concientización de la sociedad.

Los espacios naturales protegidos reciben a más de 8000 millones de visitantes al año en todo el mundo, de los cuales más del 80% corresponden a los espacios de Europa y Norteamérica (Balmford et al., 2015).

En España la actividad turística en los parques nacionales ha presentado un crecimiento estable y notable en los últimos diez años, observándose un incremento del 34% en la afluencia a los

parques hasta llegar a los 14,4 millones de visitas en el 2015 de acuerdo con los datos presentados por el Organismo Autónomo de Parques Nacionales (OAPN) (MAPAMA & OAPN, 2017).

Desde una perspectiva social, la visita a un parque nacional es una actividad que otorga bienestar y recuperación para los visitantes. Los parques nacionales son “un recurso de gran valor social y de atractivo para la población en general y un destino altamente preferido para evadirse, aprender o recuperar la experiencia y el contacto directos con la naturaleza” (López & Pardo, 2018).

En cuanto a la dimensión económica, el turismo en los parques nacionales puede ser una fuente directa de ingresos a través de la venta de entradas, pagos de servicios, etc. De forma indirecta a través de las actividades asociadas al uso recreativo (alojamiento, alimentación y transporte) son una fuente de ingreso para las comunidades cercanas (Zambrano et al., 2010).

Sin embargo, la intensificación del turismo en los parques nacionales supone presión sobre los recursos que soportan esta actividad, por ejemplo; un elevado número de visitantes supone un aumento de los desechos producidos dentro del área del parque y sus instalaciones (Eagles & McCool, 2002; Valentine, 1992). Asimismo, un mayor número de visitantes implican niveles más altos de ruido lo que afecta a la vida silvestre y el cambio de comportamiento y la interrupción de sus ciclos naturales (Buultjens et al., 2005; Orams, 1996).

Las actividades turísticas también pueden provocar la erosión del suelo y la destrucción de la flora, afectando los servicios ecosistémicos que brindan, como la regulación del suelo (Farrell & Marion, 2001).

De forma indirecta la actividad turística de los parques nacionales contribuye a los problemas ambientales globales como el cambio global ya que el transporte de los visitantes desde su lugar de procedencia involucra la emisión de CO₂ a la atmósfera (Lin, 2010).

Tabla 1. Impactos negativos del uso público de los espacios naturales protegidos.

Actividades de uso público	Acciones sobre los recursos	Impactos
Senderismo	Construcción y uso de infraestructura (senderos), Pisoteo en y fuera de los senderos	Cambios en el paisaje. Alteración del hábitat. Perturbación de la vida silvestre.
Ciclismo	Uso de bicicletas fuera de sendero	Reducción de cobertura vegetal. Erosión.
Escalada	Pisoteo en roca	Alteración del hábitat. Erosión.

Actividades de uso público	Acciones sobre los recursos	Impactos
Itinerarios 4x4	Contaminación acústica. Emisiones CO2	Degradación calidad del aire. Colisión con animales silvestres. Cambios comportamentales de fauna silvestre.
Observación de aves	Alimentación de fauna silvestre.	Cambios comportamentales de fauna silvestre. Desequilibrios en la cadena trófica
Natación, rafting, barcos a motor, remo	Uso del medio acuático. Acumulación de basura en el fondo de ríos, lagos.	Contaminación del agua.

Fuente: (González et al., 2021).

2.1.4. La capacidad de carga turística en los espacios protegidos.

Los efectos de la demanda de recreación en espacios naturales protegidos se han reconocido desde muy temprano en la historia de la conservación. En Europa, en los primeros bosques regulados, la caza se reservaba como un derecho exclusivo de la nobleza, de esta forma limitando el uso de estos espacios (Wagar, 1964).

El concepto de capacidad de carga aplicado a la actividad turística en los espacios naturales protegidos surge por primera vez a mediados de la década de 1930 en un estudio de visitantes realizado por el Servicio de parques nacionales de EE.UU, en varios espacios naturales protegidos de California (Sumner, 1936). En este estudio se plantea el problema de cómo mantener el uso público sin alterar las condiciones físicas de los espacios protegidos. Para abordar esta problemática se propuso la adaptación y aplicación del concepto de capacidad de carga utilizado para evaluar el uso de pastizales y bosques. Así, el concepto de capacidad de carga recreativa se define como el punto de saturación máximo de afluencia de visitantes, en el cual las características que hacen atractivo al espacio protegido se pueden mantener en el tiempo. Este primer concepto se basa en la hipótesis de que el aumento del número de visitantes provoca un mayor impacto ambiental medido en diferentes características físicas del espacio como la compactación del suelo, la destrucción de la vegetación, alteración del hábitat de la fauna y demás variables relacionadas.

Más adelante, varios autores se enfocaron en identificar otros factores, además de los ambientales, que podían influir en la capacidad de carga turística de un espacio protegido (Mc Cool & Lime, 2001). Especialmente, las investigaciones se centraron en aspectos sociales y de satisfacción de los visitantes, implicando que además de la capacidad de carga física de un espacio

protegido, es necesario identificar la capacidad de carga de un espacio con respecto a la calidad de la visita. En este enfoque, destaca el estudio de Wagar que define la capacidad de carga recreativa como el nivel de uso de un espacio protegido en el cual la calidad de la visita permanece constante (Wagar, 1964).

A partir de estas dos dimensiones, varios autores han desarrollado un amplio cuerpo de literatura, perfeccionando el concepto y desarrollando metodologías que permitan cuantificar la capacidad de carga (Manning, 2002). La evolución del concepto también ha derivado en la creación de una tipología que toma en cuenta distintas dimensiones del concepto y que establece cinco tipos de capacidades (Watson & Kopachevsky, 1996).

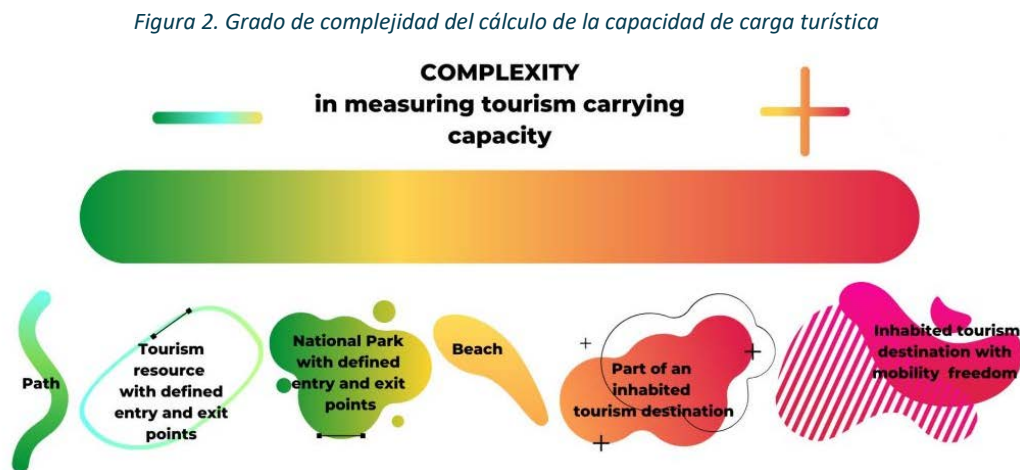
- Ecológica: nivel de actividad turística en el cual las características ambientales no se ven afectadas
- Física: capacidad espacial de un sitio y sus infraestructuras para desarrollar actividades recreacionales.
- Social: nivel de tolerancia de las poblaciones aledañas por la presencia y comportamiento de los visitantes.
- Económica: capacidad del espacio para desarrollar actividades turísticas sin excluir otras actividades económicas de la población local.
- Psicológica: nivel de actividad turística en el cual la experiencia de los visitantes no se deteriora.

A partir de esta clasificación se han llevado a cabo varios estudios enfocados en desarrollar metodologías y aplicaciones prácticas de los distintos tipos de capacidad de acogida. Entre estos trabajos, destaca el estudio de Miguel Cifuentes (1992), en el cual se propone una metodología para la determinación de la capacidad de carga turística en áreas protegidas y cuya aplicación se ha enfocado en parques nacionales (Cifuentes, 1992).

En este estudio se establecen un flujo de pasos a seguir para determinar la capacidad de carga turística. En primer lugar, se parte de analizar los objetivos del espacio protegido y si los mismos se alinean con el uso público y la oportunidad de desarrollo de actividades turísticas. Luego, se estudian los sitios de visita dentro del espacio protegido con el fin de identificar zonas de uso y zonas de conservación. A continuación, se identifican los factores específicos que influyen en el uso público de un espacio protegido, tomando en cuenta las particularidades de cada espacio. Finalmente, se determina la capacidad de carga para cada zona de uso público en base a tres niveles:

- Capacidad de carga física: es la relación entre el espacio disponible y la necesidad de espacio por visitante.
- Capacidad de carga real: se determina a partir de la capacidad de carga física y agregando factores de reducción que son particulares a cada espacio protegido.
- Capacidad de carga efectiva: es el límite máximo de visitas que se puede permitir dada la capacidad de manejo del espacio protegido.

Es importante señalar que, si bien la determinación de la metodología de Cifuentes permite cuantificar el número de personas por zona de visita y a distintos niveles, supone complejidad a la hora de obtener todos los aspectos que intervienen en su determinación, como se ilustra la siguiente figura.



El trabajo de Cifuentes ha sentado las bases de como determinar la capacidad de carga turística en distintos espacios protegidos y se ha configurado como la referencia para nuevos enfoques y metodologías, que se han aplicado en varios espacios protegidos (Rocha et al., 2021; Sayan & Atik, 2011).

2.2. Las nuevas fuentes de datos.

En nuestra sociedad, cada vez más digitalizada, las actividades de las personas ya sean presenciales o telemáticas, generan datos de forma constante. Las transacciones financieras y comerciales, las llamadas y mensajes de texto, los viajes y movilidad, entre otros, generan una cantidad ingente de datos en la internet, una especie de huella digital de nuestras vidas. A partir de julio 2021, el 65% de la población mundial contaba con acceso al internet y 79 zettabytes¹ de datos se habían

¹ 1 Zettabyte (ZB) = 1000 Exabytes = 1 millón de petabytes = 1000 millones de terabytes.

generado (Statista, 2021). Así, los datos se han transformado en un activo de gran valor para el desarrollo económico y social. En este sentido se desarrolla la llamada revolución de los datos.

Los nuevos datos pueden clasificarse de acuerdo con la naturaleza de sus fuentes. Con el foco puesto en su potencial interés para el turismo, de partida podemos distinguir como un grupo independiente aquellas fuentes de Big Data propias de este sector como las tarjetas inteligentes turísticas, los sistemas de información en destino o las páginas web de opiniones, como *TripAdvisor* o *Booking*. También podríamos englobar en este grupo los datos almacenados por las nuevas plataformas de alojamiento *peer to peer*, como *Airbnb*, que ofrecen información diversa acerca de los alojamientos geolocalizados.

Por otro lado, más allá de los datos provenientes de fuentes propias del sector turístico, hay multitud de datos provenientes de otras fuentes que están permitiendo abrir y explorar nuevas vías de análisis de la actividad turística. En primer lugar, podríamos destacar los datos procedentes de las redes sociales de la denominada *web 3.0*, caracterizadas por promover la interacción entre los usuarios. Las redes sociales son un actor clave en el rápido desarrollo y producción de datos debido a su creciente popularidad y uso. Según la plataforma Statista, había más de 1.600 millones de usuarios de redes sociales en todo el mundo en 2020 (Statista, 2021). Por ello, esta huella digital es particularmente relevante para el estudio de muchos ámbitos de un turismo en revolución. Una de las redes sociales más utilizadas en el ámbito académico es *Twitter*, dada la accesibilidad gratuita a una muestra amplia de sus datos (*tweets*), parte de los cuales están geolocalizados.

En segundo lugar, podemos destacar los datos provenientes de otro grupo de redes sociales que por su creciente expansión y por la mera naturaleza de sus datos, merecen una clasificación aparte: las redes sociales basadas en compartir fotografías como *Instagram*, *Flickr* o *Panoramio* —red ya extinta—, cuyos datos están geolocalizados. Estas redes sociales son de extraordinario interés para el turismo, dada la tendencia creciente de los turistas a realizar y compartir fotos en sus destinos.

Finalmente, podemos resaltar el gran valor de los datos provenientes de las aplicaciones de registro de rutas GPS, como *Wikiloc* o *Strava*, de gran utilidad para el estudio de las rutas seguidas tanto por turistas como por excursionistas, registradas con un nivel de detalle sin precedentes, gracias a la alta resolución espacio temporal del GPS, y con un nivel de significancia igualmente relevante, considerando la amplia muestra de datos que es posible obtener en los destinos turísticos más relevantes.

Junto a este nuevo y considerable flujo de datos, ha surgido la necesidad de transformar estos datos en información útil para la detección de patrones, tendencias y comportamientos, por lo cual nuevas técnicas de procesamiento de datos, apoyadas en tecnologías como la computación y

almacenamiento en la nube, han emergido con el fin de proporcionar respuestas al menor costo y tiempo posible.

En general, el proceso de extraer información de Big Data se divide en dos subprocesos, el manejo de datos y la analítica. El manejo de datos abarca todos los procesos y tecnología para adquirir y almacenar datos y prepararlos para su análisis y comprende tres etapas: 1) la adquisición de los datos, 2) la extracción, limpieza y filtrado de los mismos y 3) la integración y representación. En tanto que la analítica, también conocida como minería de datos, se refiere a las técnicas utilizadas para analizar y adquirir conocimiento del Big Data y que se divide en dos etapas: el modelamiento y la interpretación (Letouzé, 2012).

La forma en que se recolectan esos datos clasifican a la huella digital en dos categorías: pasiva y activa (Goodchild, 2007). Se habla de huella digital pasiva cuando los datos se extraen sin el conocimiento del usuario, por ejemplo, datos de navegación, direcciones IP, etc. En tanto que la huella digital activa se crea cuando un usuario comparte información deliberadamente a través de sitios web o redes sociales.

En definitiva, las nuevas fuentes de datos son variadas y ofrecen nuevas oportunidades para la investigación de la actividad turística. Aunque actualmente el acceso a datos más detallados o desagregados no se debe ya tanto a una barrera tecnológica como ética, establecida por los derechos en torno a la protección de datos personales, no cabe duda de que nuevas fuentes de datos cada vez más ricas seguirán incorporándose a las existentes, abriendo nuevas y prometedoras líneas de investigación acerca de la actividad turística.

2.2.1. Big Data - la revolución de los datos.

“Big Data” es un concepto con creciente popularidad que se utiliza para describir a las nuevas fuentes de datos originadas a partir de la revolución tecnológica y digital de la última década. El Internet, el desarrollo de las telecomunicaciones y la ampliación de la capacidad de almacenamiento han desencadenado una revolución en la producción y disponibilidad de datos de diversa índole.

Una definición más formal es la que propone De Mauro: "Big Data es el activo de información caracterizado por un volumen, velocidad y variedad tan elevados como para requerir una tecnología específica y métodos analíticos para su transformación en valor" (De Mauro et al., 2015). De forma similar, Gartner Inc., una importante firma de consultoría informática define al Big data como una información de gran valor con gran volumen, velocidad y variedad que demanda

formas innovadoras y rentables de procesamiento para mejorar su comprensión y la toma de decisiones (Gartner, 2017).

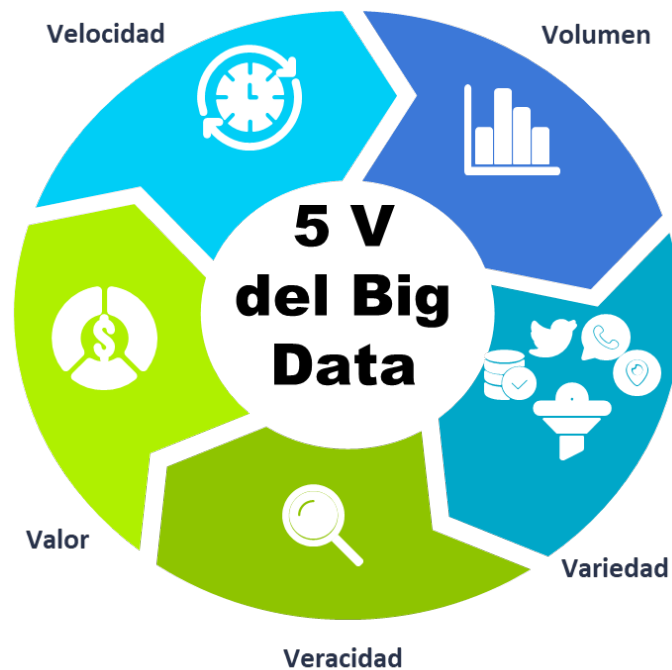
Existen tres categorías definitorias de Big Data que todos los académicos mencionan, las 3Vs (Volumen, Velocidad y Variedad) (Gandomi & Haider, 2015). El volumen se refiere a las cantidades masivas de datos que se tienen que procesar y almacenar, por lo que en términos de almacenamiento se habla de información que ocupa varios terabytes o petabytes (Schroeck et al., 2012). Un terabyte equivale a 1024 gigabytes, mientras que un petabyte equivale a 1024 terabytes. Sin embargo, el volumen en Big Data es un concepto relativo que depende del tiempo y el tipo de datos. En el futuro la capacidad de almacenar y generar datos se incrementará permitiendo de esta manera un mayor volumen de datos por lo que el Big Data de la actualidad puede no serlo en el futuro.

La característica de velocidad se refiere a la rapidez con la que los datos se generan y la velocidad a la que se deben procesar y analizar los mismos para obtener resultados fiables (Letouzé, 2012). La proliferación de dispositivos móviles como los teléfonos inteligentes que potencian el uso de redes sociales y el internet de las cosas ha dado lugar a una tasa de creación de datos inigualable y en constante crecimiento.

La variedad implica que los datos generados por estas nuevas fuentes de información tienen una estructura heterogénea, es decir los datos pueden ser; estructurados, semiestructurados y no estructurados. Nos referimos a datos estructurados cuando los mismos se encuentran organizados en filas y columnas con títulos y corresponden al que se suele encontrar en la mayoría de las bases de datos. En tanto que los datos no estructurados generalmente no tienen una estructura identificable y por lo tanto requieren mayor procesamiento para obtener información de los mismos (Demchenko & Membrey, 2013).

A estas tres características definitorias se han sumado dos nuevas características; Valor y Veracidad (Figura 3).

Figura 3. Las cinco Vs definitorias del Big Data o datos masivos



Fuente: Elaboración propia.

El valor se refiere a la capacidad de las empresas para extraer información clave de los datos, que permiten una mejor comprensión del negocio o clientes y de esta forma proporcionar servicios personalizados e información que antes no se conocía. De esta forma el valor puede ser interpretado como el valor monetario que representan los resultados generados del análisis de los datos masivos (Ishwarappa & Anuradha, 2015).

La veracidad se refiere a la exactitud de los datos. La incertidumbre en los datos puede ser causada por varias razones en los datos, que pueden ser cuestiones legales, problemas de privacidad, duplicación, etc.

El valor del Big Data radica en la capacidad de extraer información útil para la toma de decisiones, predicción de futuros comportamientos y modelar nuevos escenarios para lo cual es necesario procesos eficientes para convertir grandes volúmenes de datos rápidos y diversos en información significativa (Gandomi & Haider, 2015).

2.2.2. Redes Sociales y turismo.

El contenido generado por los usuarios de las redes sociales constituye una fuente representativa de Big Data. Por ejemplo, Facebook cuenta con más de 1860 millones de usuarios activos, los cuales suben diariamente cerca de 300 millones de fotos (Vranešević et al., 2019), Twitter cuenta con 320 millones de usuarios que cada segundo realizan 3935 (Global Data, 2017). Un aspecto

internet. Desde el punto de vista de la investigación, el análisis de datos de las redes sociales proporciona desafíos y oportunidades para generar conocimiento a partir del contenido digital generado a diario por sus usuarios.

Las redes sociales son una herramienta interactiva y multifacética que permite a sus usuarios comunicarse con otros y a la vez realizar un sin número de actividades en internet. Una cantidad considerable de datos personales son ingresados y almacenados en las diferentes redes sociales ya sea de forma consciente o inconsciente. Los datos geolocalizados generados por la interacción de sus usuarios han sido reconocidos como indicadores para entender y representar varios aspectos dentro de la geografía humana (Goodchild, 2007; Li et al., 2013). En este aspecto, las redes sociales como fuente de datos, tienen potencial para mejorar el conocimiento sobre las dinámicas temporales y espaciales de la actividad turística debido a su alta resolución e incluso pueden proporcionar datos en tiempo real (Chareyron et al., 2014).

En la investigación del turismo, la utilización de esta información tiene una historia reciente, pudiéndose encontrar artículos sobre este tema a partir del año 2007 (B. Zeng & Gerritsen, 2014). Algunas fuentes notorias de información geográfica voluntaria incluyen las redes sociales de mensajería (Twitter, Weebo), las redes sociales de fotografía (Instagram, Flickr, Panoramio), y las redes sociales especializadas (Wikiloc). Los principales temas de investigación que hace uso de estos datos comprenden: la identificación de puntos de interés o atractivo turístico, análisis de agrupamiento espacial, análisis de trayectorias y recomendación de rutas turísticas. A continuación, se describen algunos estudios revisados y organizados de acuerdo con su fuente.

En cuanto a aplicaciones de intercambio de fotografías y videos, la fuente más utilizada para la investigación es Flickr. Los metadatos de fotografías de Flickr se han empleado para medir la actividad turística en espacios urbanos mediante la diferenciación de fotografías tomadas por residentes y por turistas de acuerdo con la distribución temporal de las fotografías de cada usuario (Kádár & Gede, 2013). Otros estudios se han enfocado en identificar los flujos y patrones de movimiento de los visitantes a un destino, por ejemplo (Vu et al., 2015) emplean cadenas de Markov en fotografías geo localizadas de Flickr, para modelar las trayectorias de turistas en 10 atractivos turísticos de Hong Kong. En la misma línea, un estudio de (Z. Zeng et al., 2012) desarrolla un algoritmo para encontrar sitios turísticos potenciales a través de fotos geo localizadas de Flickr de la ciudad de Beijing. Nuevos enfoques metodológicos como Machine Learning, se han aplicado a fotografías geolocalizadas de Flickr para medir el atractivo de puntos de interés turístico en seis ciudades e identificar el comportamiento de usuarios que los visitan (Giglio et al., 2019). En otro estudio, las fotografías de Panoramio se han empleado para analizar la distribución y patrones

espaciales de concentración de actividad turística varias capitales europeas (García-Palomares et al., 2015). El componente temporal de fotografías geolocalizadas también se ha empleado para inferir los recorridos y movimientos de turistas y diferenciarlos de los residentes urbanos (Vu et al., 2015).

Asimismo, redes sociales como Twitter se han utilizado para analizar la actividad turística. Los tweets geolocalizados se han empleado para caracterizar aspectos espaciales, temporales y demográficos de los flujos turísticos en entornos urbanos (Chua et al., 2016). Técnicas como el análisis de agrupamiento o clustering se ha utilizado con tweets geolocalizados para mapear puntos calientes de actividad turística en Kioto (Oku et al., 2015). Los tweets geolocalizados también han demostrado ser útiles para la extracción de patrones de viaje en Austria y Florida (Hochmair & Cvetojevic, 2014). El análisis del contenido semántico de los tweets junto con el análisis espacial se ha aprovechado para estudiar la densidad de turistas con respecto a diferentes eventos en San Francisco (Brandt et al., 2016).

Varios autores han combinado múltiples tipos de huella digital, además de redes sociales para rastrear a los turistas de las ciudades. Por ejemplo, un estudio llevado a cabo con datos de Flickr y registros de teléfonos móviles de la ciudad de Roma ha permitido mapear la presencia y movimiento de turistas con gran detalle temporal y espacial (Girardin et al., 2008). Otro estudio combina las redes sociales de Twitter, Panoramio y Foursquare para describir la huella digital de los turistas en términos de consumo, conectividad y lugares turísticos en Madrid (Salas-Olmedo et al., 2018).

Todos estos trabajos académicos demuestran el interés del uso de los datos de redes sociales para la investigación. En este sentido, las redes sociales ofrecen oportunidades para aplicar nuevos enfoques metodológicos en distintas temáticas del turismo. De igual manera, los datos geolocalizados presentan ventajas interesantes como fuentes de datos. Una de las ventajas más señalada en las publicaciones revisadas es el bajo costo de adquisición de los datos en términos de acceso y tiempo. Gran parte de los datos generados en redes sociales son de acceso público, y muchas plataformas permiten la descarga de sus datos anonimizados.

Otra importante ventaja es la cobertura global de esas fuentes, ya que las redes sociales son plataformas globales, lo que permite el acceso a datos de usuarios de diferentes lugares del mundo y la posibilidad de realizar comparaciones entre diferentes áreas de estudio.

Otra ventaja importante de usar datos geolocalizados consiste en la relativa facilidad de replicar los datos generados por otros investigadores y la capacidad de tener una recopilación continua de datos (Stienmetz & Fesenmaier, 2016).

No obstante, los datos que se pueden recolectar de redes sociales también tienen limitaciones. Uno de los limitantes más señalados consiste en que los datos provenientes de redes sociales están sesgados por qué no toda la población usa redes sociales. Por eso, necesario antes de llevar a cabo cualquier estudio el conocer el perfil de la red social que se va a utilizar.

Otra limitante, está relacionado con la variabilidad del volumen de datos que se generan a nivel de cada país o localidad, lo que puede provocar escasez de datos e influenciar en la representatividad del estudio. En este sentido, es importante analizar la importancia de la red social que se va a emplear.

3. El uso de datos geolocalizados de redes sociales en el estudio del turismo en espacios naturales protegidos.

3. El uso de datos geocalizados de redes sociales en el estudio del turismo en espacios naturales protegidos.

Con el uso generalizado de dispositivos integrados con sensores de localización, las personas comparten activamente su ubicación cuando usan plataformas de redes sociales (Hecht et al., 2011). Haciendo de los datos geocalizados cada vez más accesibles para su uso y análisis. Con respecto al turismo de naturaleza y en particular en destinos como los parques nacionales, esta nueva fuente de datos se configura como una alternativa interesante para estudiar el comportamiento espacial y temporal de los visitantes. En este contexto los datos geocalizados de redes sociales permiten descubrir patrones espaciales y temporales del turismo de parques nacionales, de la misma forma en que se ha evidenciado para otros tipos de turismo como se mencionó en el capítulo anterior.

A pesar de que el uso de datos geocalizados de redes sociales para el análisis del turismo de naturaleza no presenta una producción científica tan abundante como la de otros tipos de turismo, existe un creciente número de publicaciones que abordan su uso y aplicación en la investigación. En este capítulo, se presenta una revisión de alcance de la literatura sobre el uso de datos geocalizados de las redes sociales para analizar el turismo en espacios naturales protegidos con el fin de mapear sistemáticamente la investigación realizada hasta el momento, así como para identificar brechas de investigación para futuras investigaciones. Los estudios revisados muestran resultados prometedores e indican que los datos de las redes sociales también pueden ser una fuente complementaria a los métodos tradicionales para aportar nuevos conocimientos sobre el comportamiento espacio-temporal de los visitantes. El contenido de este capítulo se encuentra publicado en una revista de alto impacto como un artículo de revisión de la literatura.

El uso de datos de redes sociales para la investigación es un campo relativamente nuevo, particularmente para el turismo en áreas protegidas. Sin embargo, el rápido crecimiento de la investigación con datos geocalizados de redes sociales ha dado lugar a un importante grupo de artículos que abordan esta temática. En consecuencia, el objetivo de este capítulo es mapear el estado actual del conocimiento sobre el uso de datos geocalizados de las redes sociales para analizar la actividad turística y el comportamiento de los visitantes de los espacios naturales protegidos.

El enorme potencial de los datos de las redes sociales ha generado un interés creciente en el uso de estas fuentes de datos para estudiar el turismo basado en la naturaleza, que se ve reflejado en el número creciente de publicaciones sobre este tema. Sin embargo, el uso de estas nuevas

fuentes de datos se encuentra en una etapa temprana y la realización de estudios más profundos sobre la aplicabilidad y efectividad de los datos de redes sociales es imperativa (Heikinheimo et al., 2017).

3.1. Metodología de la revisión del estado del arte.

Se realizó una revisión de alcance (scoping review) para comprender cómo se han utilizado los datos geolocalizados de las redes sociales en el análisis de diferentes aspectos del turismo en espacios naturales protegidos. Las revisiones de alcance se utilizan para: "examinar el alcance y la naturaleza de un área particular de investigación, resumir los hallazgos de la investigación existente e identificar las lagunas en la investigación (Arksey & O'Malley, 2005). Las revisiones de alcance también han sido reconocidas como adecuadas para áreas de investigación menos exploradas (Kastner et al., 2012). En consecuencia, se seleccionó este tipo de revisión, ya que el uso de datos geolocalizados en el turismo de espacios naturales protegidos es un tema relativamente nuevo y, por lo tanto, su investigación se encuentra en una etapa temprana. Además, una revisión de alcance permite una discusión más cualitativa sobre los beneficios y limitaciones de esta nueva fuente de datos.

El marco propuesto por Arksey y O'Malley está estructurado en cinco etapas: (1) Identificación de la pregunta de investigación, (2) Búsqueda de los estudios relevantes, (3) Selección de los estudios, (4) Visualización de los resultados mediante la elaboración de gráficos y (5) Cotejar, resumir y reportar los resultados.

Por lo que, en primer lugar, se definió la pregunta de investigación como: ¿Qué se sabe de la literatura existente sobre el uso de datos geolocalizados de las redes sociales para la investigación del turismo en espacios naturales protegidos?

En el segundo paso, se realizó la búsqueda de artículos científicos utilizando dos bases de datos de artículos científicos; Web of Science y Scopus como fuentes de información. La búsqueda se realizó el 3 de junio de 2020 y se centró únicamente en artículos de revistas publicados (excluyendo actas de congresos, capítulos de libros y artículos en prensa). En nuestra consulta de búsqueda se utilizaron diferentes combinaciones de palabras clave relacionadas con la temática de redes sociales y turismo de espacios protegidos, junto con los nombres de varias redes sociales. Buscamos estos términos en el título, tema y resumen. La sintaxis de búsqueda se estandarizó para ambas bases de datos (Tabla 2).

Tabla 2. Consulta de búsqueda estandarizada para ambas bases de datos de artículos científicos

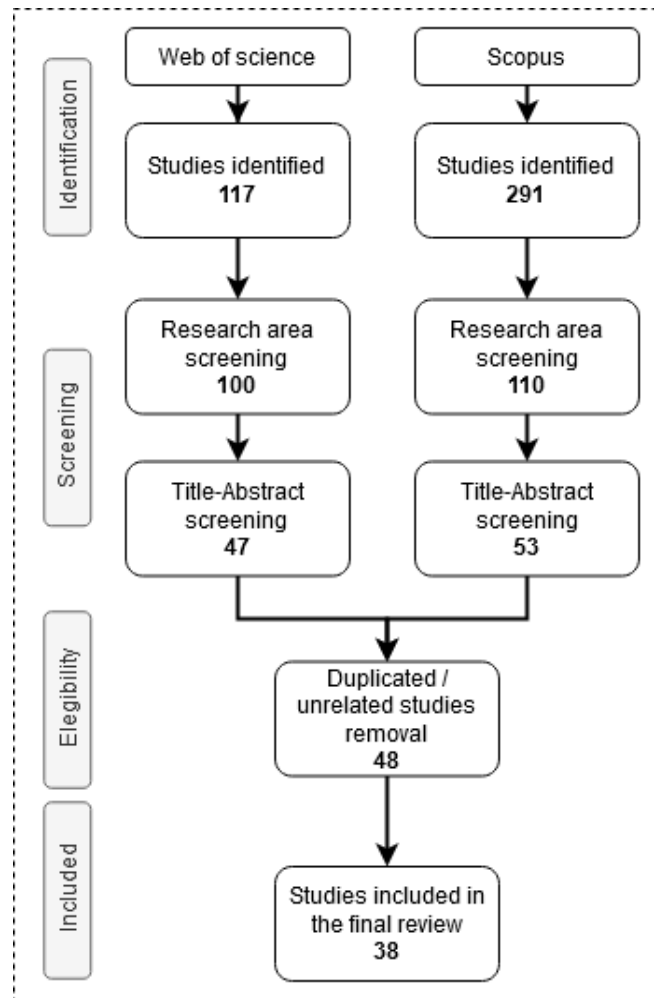
Base de datos	Sintaxis de la consulta de búsqueda
Web of Science	(TS=(("geotag*" OR "social media" OR "crowdsourc*" OR "user generated" OR "Big Data" OR Flickr OR Instagram OR Panoramio OR Wikiloc OR GPSies OR Strava OR Twitter OR Facebook OR Weibo) AND ("protected area*" OR "national park" OR "natural park" OR "nature based tourism" OR recreation OR visitor* OR "visitor monitoring") AND WC=(Geography OR "Biodiversity Conservation" OR Ecology OR "Environmental Sciences" OR Environmental Studies OR "Multidisciplinary Sciences")) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article)
Scopus	(TITLE-ABS-KEY (geotagged OR "social media" OR "crowdsourc*" OR "user generated" OR "Big Data" OR flickr OR instagram OR panoramio OR wikiloc OR gpsies OR strava OR twitter OR facebook OR weibo) AND ("protected area*" OR "national park" OR "natural park" OR "nature based tourism" OR recreation OR visitor* OR "visitor monitoring")) AND SUBJAREA (agri OR soci OR envi OR eart) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))

La búsqueda identificó 117 estudios potenciales en Web of Science y 291 en Scopus. A partir de estos resultados, realizamos un filtrado de tres niveles para seleccionar la literatura relevante. Primero, seleccionamos y eliminamos los estudios cuya área de investigación estaba fuera del alcance de esta revisión. Por ejemplo, estudios donde el turismo no era el tema principal, como publicaciones enfocadas en temas como la conservación de la biodiversidad, el marketing o gestión empresarial. La primera selección retuvo 100 estudios para Web of Science y 110 para Scopus.

Luego, realizamos una selección de títulos y resúmenes con el fin de retener los artículos que utilizan únicamente datos geolocalizados de redes sociales. Por ejemplo, se eliminaron los estudios que utilizaban datos geolocalizados de otras fuentes, como datos de teléfonos móviles u otro tipo de datos de fuentes colaborativas, como Open Street Maps y proyectos PPGIS. Después del filtrado por título, nos quedamos con 47 estudios para Web of Science y 53 para Scopus.

En un tercer paso, combinamos los resultados de ambas bases de datos y eliminamos los artículos duplicados, lo que resultó en 48 estudios. El paso final incluyó una revisión adicional de los 48 artículos identificados al excluir los estudios que no tenían un espacio protegido como caso de estudio, lo que resultó en 38 estudios que son el tema de esta revisión de alcance. Los artículos seleccionados finales se descargaron y exportaron al administrador de referencias de Mendeley para su posterior análisis y lectura.

Figura 5. Esquema del proceso de filtrado de artículos



Elaboración propia

En la cuarta y quinta etapa del estudio, todas las publicaciones incluidas se examinaron cuidadosamente y los hallazgos se clasificaron según el tipo de información relevante, teniendo como eje principal a los turistas o visitantes de los espacios naturales protegidos. Estos hallazgos se organizaron en seis temáticas en forma de preguntas que abordan los aspectos esenciales del turismo en espacios naturales protegidos (Tabla 3):

- 1) ¿Cuántos visitantes van a los espacios naturales protegidos?
- 2) ¿De dónde provienen los visitantes?
- 3) ¿A dónde van los visitantes dentro de los espacios naturales protegidos?
- 4) ¿Cuándo visitan el espacio natural protegido?
- 5) ¿Qué actividades realizan los visitantes?
- 6) ¿Por qué visitan el espacio natural protegido?

Tabla 3. Artículos revisados organizados de acuerdo con preguntas sobre el turismo en espacios naturales protegidos.

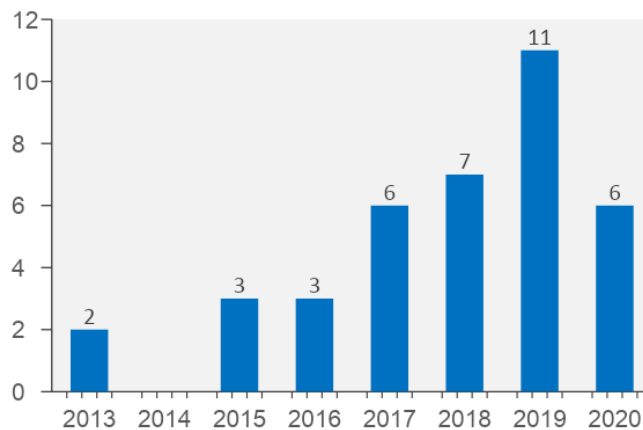
Artículo	¿Cuántos visitantes?	¿De dónde vienen?	¿A dónde van?	¿Cuándo visitan?	¿Qué hacen?	¿Por qué visitan?
(Orsi & Geneletti, 2013)			X			
(Wood, et al., 2013)	X	X		X		
(Richards & Friess, 2015)			X			X
(Willemen, et al., 2015)						X
(Levin, Kark, & Crandall, 2015)	X					
(Sonter, et al., 2016)	X	X				
(Campelo & Nogueira Mendes, 2016)					X	
(Sessions, et al., 2016)	X	X		X		
(Heikinheimo et al., 2017)	X	X	X	X	X	
(Tenkanen et al., 2017)	X			X		
(Norman & Pickering, 2017)	X				X	
(Hausmann et al., 2017)						X
(Levin, Lechner, & Brown, 2017)	X					
Márquez-Pérez J., et al (2017)					X	
(Hausmann et al., 2018)						X
(Fisher et al., 2018)	X					
(Walden-Schreiner, et al., 2018)	X		X	X		
(Mancini, Coghill, & Lusseau, 2018)				X	X	
(Walden-Schreiner, et al., 2018)			X			
(Martinez-Harms, et al., 2018)	X	X	X			
(Sinclair et al., 2018)		X	X			X
(Norman, Pickering, & Castley, 2019)	X			X		
(Kim, et al, 2019)	X		X			X
(Jurado Rota, et al., 2019)			X	X		
(Gosal, et al., 2019)				X		X
(Sinclair, et al. 2019)	X					X
(Clemente, et al., 2019)			X			X
(Callau, et al., 2019)			X			X
(Norman & Pickering, 2019)	X				X	X
(Barros, et al 2019)	X			X		
(Retka, et al., 2019)			X	X		
(Hausman et al., 2019)			X			X
(Rossi, et al., 2019)						X
(Pickering, et al., 2020)				X		X

Artículo	¿Cuántos visitantes?	¿De dónde vienen?	¿A dónde van?	¿Cuándo visitan?	¿Qué hacen?	¿Por qué visitan?
(Barros et al, 2020)	X	X	X	X	X	
(Do & Kim 2020)			X	X		X
(Sinclair, et al., 2020)	X	X				
(Chun et al., 2020)	X		X			

3.2. Descripción general de los artículos revisados.

En total se identificaron 38 artículos que emplearon datos geolocalizados de redes sociales para analizar o caracterizar el turismo en espacios naturales protegidos. La producción anual de publicaciones revela que existe una tendencia creciente en este campo de investigación (Figura 6). Aunque la investigación es relativamente reciente, el primer artículo se publicó en 2013, observamos un rápido crecimiento, ya que una cantidad considerable de los artículos revisados (46%) se publicaron en los últimos dos años. Los artículos fueron publicados en 21 revistas diferentes y cubren diferentes áreas de investigación; Ciencias ambientales (18), Geografía (7), Turismo (6), Ecología (4) y Conservación (2). De las 38 publicaciones, 30 utilizan datos de una única plataforma de redes sociales. Los estudios restantes combinaron datos de; el mismo tipo de plataforma (4), diferentes redes sociales (3) y combinados con otros datos de crowdsourcing (1).

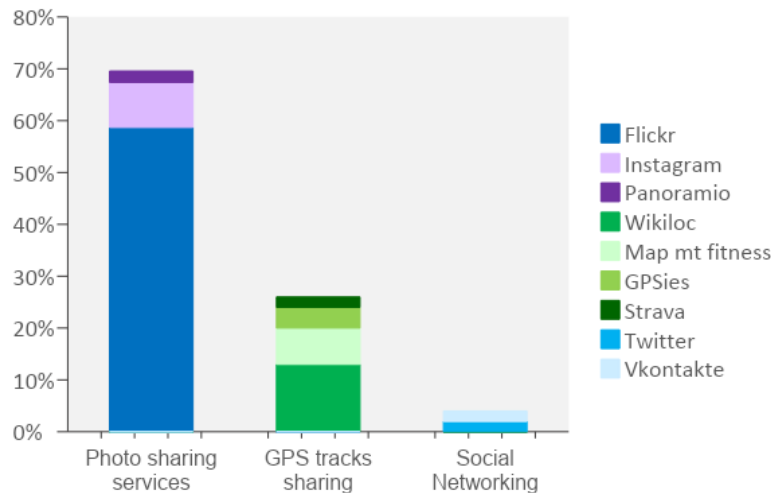
Figura 6. Producción de artículos por año.



La mayoría de los estudios utilizaron datos geolocalizados de redes sociales de intercambio de fotografías, siendo Flickr, el más utilizado dentro de este tipo de plataforma con un total de 29 artículos (Figura 7). Asimismo, existe un número creciente de estudios que en los últimos años han utilizado rutas GPS de plataformas de intercambio de rutas. En contraste con otros estudios de turismo, donde Twitter se usa ampliamente, parece que esta red social es poco adecuada para la investigación del turismo de espacios naturales protegidos, ya que solo un artículo la utilizó como fuente de datos.

El acceso a las interfaces de programación de aplicaciones (API) para la recopilación de datos parece ser un factor importante a la hora de seleccionar la fuente de datos. De los 38 estudios, 26 utilizaron APIs para recolectar y descargar sus datos. Se han comenzado a aplicar técnicas de recolección más novedosas como el web scraping, que consiste en recuperar contenido de páginas web a través de software o scripts automatizados, ya que dos artículos utilizaron este método (Callau et al., 2019; Jurado Rota et al., 2019). Los artículos restantes descargaron los datos manualmente o utilizaron una base de datos existente, como la base de datos InVEST de fotos geolocalizadas de Flickr.

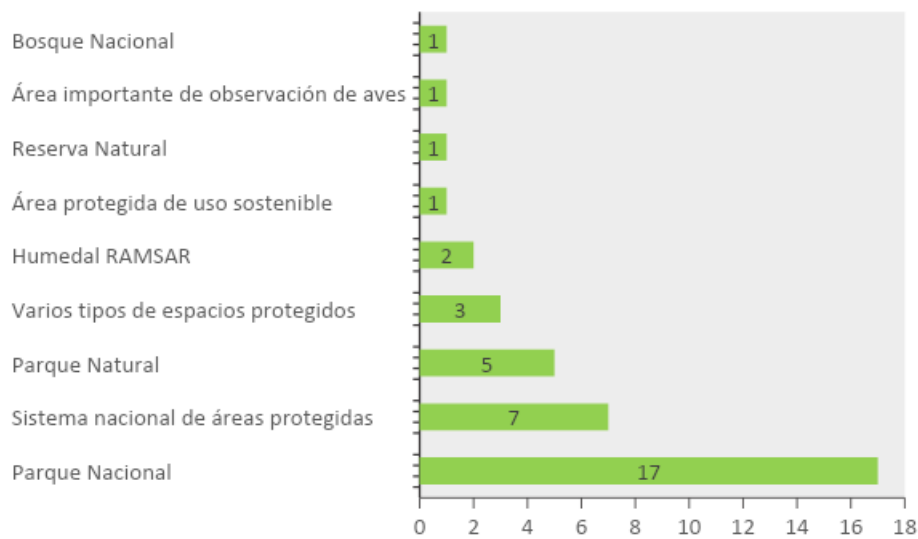
Figura 7. Distribución de artículos por fuente de datos utilizada.



Los estudios abordaron diversas categorías de espacios protegidos, siendo los parques nacionales el tipo más seleccionado para la investigación en los estudios revisados (Figura 8). Los parques nacionales se consideran los principales atractivos para los visitantes de los espacios protegidos porque muestran los ecosistemas más representativos dentro de un país. La etiqueta parque nacional es la etiqueta más conocida por turistas y visitantes (Barros et al., 2019; Livina & Reddy, 2017).

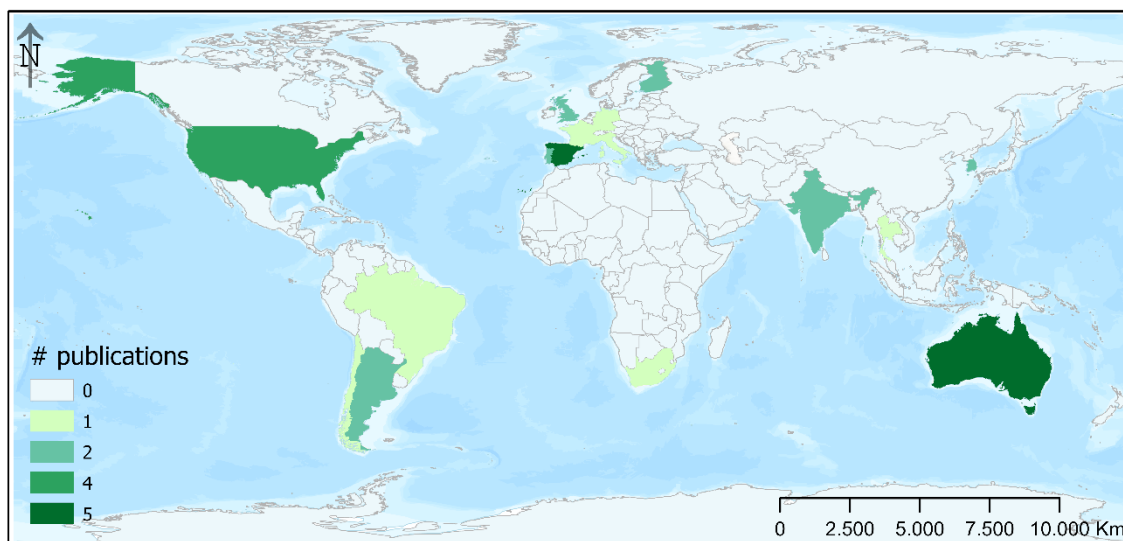
También hubo una cantidad considerable de estudios que abordaron todo el sistema de espacios protegidos de una región (Levin et al., 2017; Martinez-Harms et al., 2018; Sonter et al., 2016), un continente (Hausmann et al., 2017; Willemen et al., 2015) y el mundo (Levin et al., 2015; Wood et al., 2013).

Figura 8. Artículos revisados de acuerdo con la categoría de espacio protegido analizado.



A pesar de que los artículos revisados se han realizado con espacios protegidos ubicados en 17 países, la mayoría de los estudios se centraron en países del mundo desarrollado, con el 58% de las publicaciones elaboradas sobre espacios protegidos de Europa, América del Norte y Oceanía

Figura 9. Distribución de los artículos por país del área de estudio.



Mientras que, los artículos de las economías en desarrollo representan el 32% de las publicaciones revisadas con 2, 4 y 6 publicaciones de África, América Latina y Asia respectivamente. Además, se han realizado estudios que abarcan un alcance más amplio, por ejemplo, a escala global se han publicado 3 artículos y a escala transcontinental se encontró 1 artículo.

3.3. ¿Cuántos visitantes van a los espacios naturales protegidos?

La información de referencia más fundamental y necesaria para el control y la gestión del uso público de un espacio protegido es el número de visitantes que acuden en un intervalo de tiempo a dicho espacio (Cessford & Muhar, 2003). Las tasas y conteos de visitantes son el punto de partida para evaluar el impacto del turismo; por lo tanto, muchos resultados de la gestión dependen del conteo de visitantes para estimar indicadores como el flujo de visitantes, la presión de los visitantes, la capacidad de carga, entre otros. En este sentido, en este apartado reunimos aquellos artículos que han utilizado los datos geolocalizados de las redes sociales para cuantificar la presencia de los visitantes en un espacio protegido. En esta revisión 19 artículos abordaron esta cuestión utilizando datos de distintas redes sociales. De forma general, un gran porcentaje, el 78% de los artículos revisados utilizaron datos de Flickr para construir indicadores que reflejan el número de visitantes.

Un primer enfoque explorado en los artículos para cuantificar los visitantes consiste en comparar los conteos de visitantes oficiales con las estimaciones de visitantes derivadas del conteo de usuarios que han publicado fotos de los espacios protegidos en el mismo intervalo de tiempo. El objetivo de este enfoque es el de validar los datos geolocalizados como indicadores representativos de la presencia de visitantes en los espacios y luego continuar con un análisis más detallado de los patrones de visitas.

La mayoría de los estudios compararon las tasas de visitas anuales frente al conteo de usuarios anuales y establecieron una relación sólida entre las estimaciones de las redes sociales y el número de visitantes. El método preferido para validar los datos fue el análisis de correlación, utilizando los coeficientes rho de Pearson o Spearman (Tabla 4). Los resultados generales de las comparaciones arrojaron valores significativos de correlación positiva que oscilaron entre 0,22 y 0,91. Los modelos lineales generalizados también se utilizaron en menor medida para comparar cómo los datos geolocalizados junto con variables adicionales como el entorno ambiental y los factores socioeconómicos influyen en el número de visitantes en las áreas protegidas (Sessions et al., 2016). La mayoría de las publicaciones también coinciden en el uso de recuentos de usuarios en lugar de recuentos de fotos, tweets o rutas como un proxy para el número de visitas. Al considerar los usuarios se puede eliminar datos redundantes y reducir la desigualdad de participación de los usuarios (Tenkanen et al., 2017). Por lo tanto, el procedimiento estándar para cuantificar a los visitantes en base a redes sociales se ha establecido como el conteo del número de usuarios únicos por día y luego su agregación a diferentes escalas (mes y año).

En términos de cobertura espacial, los autores de los artículos han validado las estimaciones de los visitantes a nivel mundial (Levin et al., 2015; Wood et al., 2013), a nivel de país (Barros et al., 2019;

Chun et al., 2020; Ghermandi & Sinclair, 2019; Sinclair et al., 2020), a nivel regional (Martinez-Harms et al., 2018; Sessions et al., 2016; Sonter et al., 2016) y escala local (Barros et al., 2020; Kim et al., 2019; Walden-Schreiner, Leung, et al., 2018).

Un segundo enfoque que se ha analizado consiste en cuantificar la presencia de visitantes en la infraestructura de los espacios protegidos, como los senderos y caminos dentro. En este sentido, las rutas GPS de los sitios de uso compartido de rutas parecen ser más adecuadas para evaluar el uso de los senderos (Norman et al., 2019; Norman & Pickering, 2017). Por ejemplo, la cantidad de rutas cargadas en aplicaciones de fitness y al aire libre demostró ser un indicador de la popularidad de los senderos en espacios protegidos de Australia y presentaron una correlación alta con los datos de los contadores instalados en los senderos (Norman & Pickering, 2017).

Con menos frecuencia, la investigación también ha examinado el uso de los datos de diferentes redes sociales con el fin de identificar cuál de las plataformas funciona mejor como indicador proxy del número de visitantes (Tenkanen et al., 2017) y como indicador del uso de senderos (Norman et al., 2019).

Tabla 4. Artículos que abordaron la pregunta ¿Cuántos visitantes?

Referencia	Fuente de datos	Método	Resultados	Tipo de EP
(Wood, et al., 2013)	Flickr	GLM	$R^2 = 0.386$	Espacios Protegidos
(Levin, Kark, & Crandall, 2015)	Flickr	Pearson's r	$r = 0.80$	Espacios Protegidos
(Sonter, et al., 2016)	Flickr	GLM	$R^2 = 0.22$	Tierras de conservación
(Sessions, et al., 2016)	Flickr	GLM	$R^{2*} = 0.649$	Parques Nacionales
(Heikinheimo et al., 2017)	Instagram	Spearman rho	$r_s = 0.67$	Parques Nacionales
(Tenkanen et al., 2017)	Instagram, Flickr, Twitter	Spearman rho	$r_s = 0.84 - 0.75$	Parques Nacionales
(Norman & Pickering, 2017)	Wikiloc, GPSies, Map my fitness	GLM	$R^2 = 0.864 - 0.588$	Parques de Conservación
(Levin, Lechner, & Brown, 2017)	Flickr, Open Street Map y Wikipedia	Spearman rho	$r_s = 0.584$	Espacios Protegidos
(Fisher et al., 2018)	Flickr	Pearson's r	$r = 0.790$	Bosque Nacional
(Walden-Schreiner, et al., 2018)	Flickr	Spearman rho	$r_s = 0.312 - 0.249$	Parques Nacionales
(Martinez-Harms, et al., 2018)	Flickr	Spearman rho	$r_s = 0.74$	Espacios Protegidos
(Norman, Pickering, & Castley, 2019)	Map my fitness	GLM	$R^2 = 0.836 - 0.284$	Parques Nacionales
(Kim, et al, 2019)	Flickr	Pearson's r	$r = 0.9$	Parques Nacionales
(Sinclair, et al. 2019)	Flickr	Pearson's r	$r = 0.84$	Humedales

Referencia	Fuente de datos	Método	Resultados	Tipo de EP
(Norman & Pickering, 2019)	Map my fitness, Strava, Wikiloc	GLM	$R^2 = 0.369 - 0.027$	Parques Nacionales, Parques de Conservación
(Barros, et al 2019)	Flickr	Pearson's r	$r = 0,98 - 0,54$	Parques Nacionales
(Chun et al., 2020)	Flickr	Pearson's r	$r = 0.77$	Parques Nacionales
(Barros et al, 2020)	Flickr, Wikiloc	Pearson's r	$r = 0.84$	Parques Nacionales
(Sinclair, et al., 2020)	Flickr, Vkomtakte	Pearson's r	$r = 0.97$	Parques Nacionales

* Coeficiente en regresión negativa binomial

3.4. ¿De dónde vienen los visitantes?

Conocer las características de los visitantes es importante para la gestión y planificación de los espacios naturales protegidos. Los datos que caracterizan a los visitantes incluyen datos demográficos y sociales tales como la edad, el nivel educativo, el lugar de procedencia, entre otros (Eagles et al., 2002). Entre estas características, el lugar de origen de los visitantes es importante porque puede revelar la existencia de variabilidad en los patrones de visita y preferencias de acuerdo con el origen del visitante (Sinclair et al., 2020). A este respecto, las fotografías geolocalizadas se han utilizado para estimar la ubicación de la residencia de los visitantes de espacios naturales protegidos, destacando Flickr como la fuente de datos utilizada en la mayoría de los artículos revisados (Tabla 5).

La mayoría de las publicaciones utilizaron el atributo “home location”, que el usuario agrega a la hora de registrar su perfil en la red de Flickr, para inferir el origen de los visitantes.

Dado que el registro de este atributo en un perfil de usuario es opcional, no todos los usuarios ingresan este dato en su perfil. Según los estudios revisados, entre el 30% y el 52% de los usuarios registraron la localización de su residencia en su perfil.

Por ello, algunos autores se han enfocado en evaluar diferentes métodos de predicción para inferir el origen de los usuarios que no proporcionaron estos datos (Bojic et al., 2015). Estos métodos dependen de la extracción y análisis de toda la colección de fotos de los usuarios (Sinclair et al., 2020).

Uno de los métodos más populares es el estimar la ubicación de la residencia de un usuario como el lugar donde el usuario tomó el número máximo de fotos (Heikinheimo et al., 2017; Sinclair et al., 2018; Wood et al., 2013).

Otro enfoque empleado en los artículos revisados consiste en asumir el origen del usuario como el lugar donde un usuario pasó el número máximo de días activos (Sinclair et al., 2020).

Las técnicas de agrupación de K-means también se utilizaron para estimar la ubicación del hogar de los usuarios (Martinez-Harms et al., 2018).

Una vez que se ha estimado el lugar de origen de los visitantes por medio de los datos de redes sociales, este indicador se ha utilizado para diferentes aplicaciones como; la clasificación de visitantes en grupos similares (Barros et al., 2020; Sessions et al., 2016), la estimación del gasto por visitante de acuerdo a nacionalidad (Sonter et al., 2016), el mapeo de la distribución espacial de diferentes turistas (Heikinheimo et al., 2017; Sinclair et al., 2018) y la accesibilidad a espacios protegidos de acuerdo a la procedencia (Martinez-Harms et al., 2018).

En general, todos los artículos encontraron una fuerte correlación positiva entre el atributo de ubicación de la residencia del usuario y el país de origen de los visitantes reportado en encuestas y formularios de las administraciones de los espacios protegidos (Barros et al., 2020; Sonter et al., 2016; Wood et al., 2013), lo que demuestra que los datos de las redes sociales son representativos de los visitantes y revelan información sobre la procedencia de los turistas o visitantes.

Tabla 5. Artículos que abordaron la pregunta ¿De dónde vienen los visitantes?

Referencia	Fuente de datos	Método	Escala del indicador	Tipo de EP
(Wood, et al., 2013)	Flickr	ANCOVA	País	Espacios protegidos
(Sonter, et al., 2016)	Flickr	Regresión Lineal	Nacional, Local	Espacios protegidos
(Sessions, et al., 2016)	Flickr	Mann-Whitney U-Test	Internacional	Parques Nacionales
(Heikinheimo et al., 2017)	Instagram	Pearson Chi-square test	Región, país	Parques Nacionales
(Martinez-Harms, et al., 2018)	Flickr	Geocodificación	Provincia	Espacios protegidos
(Sinclair et al., 2018)	Flickr	Pearson's r	Internacional	Humedales
(Barros et al., 2020)	Flickr, Wikiloc	Pearson's r	País	Parques Nacionales
(Sinclair, et al., 2020)	Flickr, Vkontakte	Pearson's r	Nacional, Local, Internacional	Parques Nacionales

3.5. ¿Cuándo visita la gente los espacios protegidos?

Los datos sobre la distribución temporal de los visitantes se utilizan para medir los flujos de visitantes, el movimiento de visitantes en el interior del espacio y para identificar patrones estacionales de uso (Newsome et al., 2012). La distribución temporal también es un componente

esencial para medir la capacidad de carga. Por lo tanto, conocer la variabilidad temporal en el volumen de visitantes a cualquier espacio protegido es importante para examinar y mitigar los impactos de los visitantes (Hadwen et al., 2007).

Se ha demostrado que los datos geolocalizados de las redes sociales son útiles para explorar la variación temporal de las actividades humanas (Gutiérrez-Puebla, 2018). A este respecto el atributo de “timestamp” de una publicación en las redes sociales proporciona datos detallados de la fecha y la hora (Di Minin et al., 2015).

En el seguimiento de visitantes de espacios protegidos, este atributo abre oportunidades para explorar el comportamiento temporal de los visitantes con una granularidad más fina. Una cantidad considerable de los artículos revisados ha reconocido el valor del atributo “timestamp” como indicador potencial de la distribución temporal de los visitantes. En total, 14 estudios analizaron los patrones temporales de visitantes a través de las redes sociales (Tabla 6).

Sobre todo, la investigación se ha enfocado en los datos de tiempo de las fotos geolocalizadas. Tener una distribución temporal a diferentes escalas permite describir tendencias y patrones de visitas. Por ejemplo, utilizando modelos de regresión lineal y análisis de correlación, se ha validado el número agregado de usuarios de Flickr por mes como un proxy de los visitantes mensuales en parques nacionales de diferentes países (Barros et al., 2019; Heikinheimo et al., 2017; Sessions et al., 2016; Tenkanen et al., 2017; Walden-Schreiner, Leung, et al., 2018; Walden-Schreiner, Rossi, et al., 2018).

En la misma línea, se analizó la distribución temporal de los datos geolocalizados para encontrar diferencias en los patrones de visitas semanales durante eventos especiales o festivales (Wood et al., 2013) y para encontrar el mes y la hora preferidos de los visitantes (Do & Kim, 2020).

Conjuntamente, se han analizado las distribuciones temporales de rutas GPS ya que cuentan con una resolución más fina para documentar los cambios en la capacidad de carga de visitantes según el día de la semana y la hora del día en los senderos de los espacios protegidos. Por ejemplo, se utilizó la distribución semanal y diaria de las rutas GPS para describir la dinámica temporal de las diferentes actividades realizadas por los visitantes en los senderos de los espacios protegidos de Australia (Norman et al., 2019).

Un segundo grupo de artículos ha utilizado las distribuciones temporales de fotografías geolocalizadas para identificar patrones estacionales de uso. La estacionalidad es una característica clave en el turismo (Baum & Lundtorp, 2001). Con el fin de encontrar patrones de uso estacionales, los investigadores han aplicado varias herramientas como la creación de mapas de distribución

espacial de diferentes visitantes en diferentes franjas de tiempo (Gosal et al., 2019; Jurado Rota et al., 2019) y el cálculo de índices de estacionalidad (Barros et al., 2019).

En general, los hallazgos muestran que los datos geolocalizados se ajustan bien a los registros oficiales a nivel mensual (Mancini et al., 2018). Además, el atributo “timestamp” permite extraer datos con una resolución más fina lo que permite identificar y localizar los momentos de mayor presión de los visitantes.

Tabla 6. Artículos que abordaron la pregunta ¿Cuándo visita la gente los espacios naturales protegidos?

Referencia	Fuente de datos	Método	Indicador	Tipo de espacio
(Wood, et al., 2013)	Flickr	Agregación de datos	Usuarios por semana	Espacios protegidos
(Sessions, et al., 2016)	Flickr	GLM	Usuarios por estación y mes	Parques Nacionales
(Heikinheimo et al., 2017)	Instagram	Agregación de datos	Usuarios por mes	Parques Nacionales
(Tenkanen et al., 2017)	Instagram, Flickr, Twitter	Pearson r	Usuarios por mes	Parques Nacionales
(Walden-Schreiner, et al., 2018)	Flickr	Pearson r	Usuarios por mes y hora	Parques Nacionales
(Mancini, Coghill, & Lusseau, 2018)	Flickr	Análisis Wavelet	Usuarios por estación	Parques Nacionales
(Norman, Pickering, & Castley, 2019)	Map my fitness	Chi-square test	Rutas por día	Parques Nacionales
(Jurado Rota, et al., 2019)	Wikiloc	Agregación de datos	Rutas por día y mes	Humedal
(Gosal, et al., 2019)	Flickr	Agregación de datos	Usuarios por estación	Reserva Natural
(Barros et al. 2019)	Flickr	Índice de estacionalidad	Usuarios por mes, día y hora	Parques Nacionales
(Retka, et al., 2019)	Flickr	Agregación de datos	Usuarios por mes, día y hora	Espacio de uso sostenible
(Pickering, et al., 2020)	Flickr	Agregación de datos	Usuarios por mes, día y hora	Parques Nacionales
(Barros et al., 2020)	Flickr, Wikiloc	Pearson r	Usuarios por mes, día y hora	Parques Nacionales
(Do & Kim 2020)	Flickr	Agregación de datos	Usuarios por mes y hora	Parques Nacionales

3.6. ¿Qué hacen los visitantes en los espacios protegidos?

Los datos sobre las actividades realizadas por los visitantes brindan a la administración y a los investigadores información útil para diseñar programas de monitoreo específicos y para evaluar adecuadamente los impactos de los visitantes (Hadwen et al., 2007). Generalmente, las actividades permitidas en los espacios naturales protegidos incluyen la interpretación ambiental, los deportes basados en la naturaleza y las actividades recreativas (Marques et al., 2010). En este sentido, los

datos geolocalizados de las redes sociales pueden aportar nuevos conocimientos sobre qué actividades son populares entre los visitantes, ya que los usuarios de las redes sociales comparten lo que hacen y disfrutan (Syn & Oh, 2015). Del total de artículos revisados, solo siete publicaciones exploraron el uso de datos geolocalizados para estudiar las actividades de los visitantes durante su estancia en los espacios protegidos (Tabla 7).

La mayoría de los artículos utilizaron las rutas GPS de las redes sociales de intercambio de rutas para estimar y mapear la intensidad del uso de los senderos para el desarrollo de diferentes actividades deportivas como; el senderismo (Barros et al., 2020), el ciclismo de montaña (Campelo & Nogueira Mendes, 2016; Norman et al., 2019) y carrera (Norman & Pickering, 2017).

Asimismo, se han utilizado las rutas GPS para calcular los tiempos de recorrido para realizar caminatas en parques nacionales y naturales (Márquez-Pérez et al., 2017).

En contraparte, se han utilizado fotos geolocalizadas para extraer las actividades retratadas en las imágenes y evaluar su popularidad (Heikinheimo et al., 2017). También se han utilizado las fotos para mapear la concentración de diferentes actividades de observación de vida silvestre (Mancini et al., 2018).

Tabla 7. Artículos que abordan la pregunta ¿Qué hacen los visitantes en los espacios protegidos?

Referencia	Fuente de datos	Método	Indicador	Tipo de EP
(Campelo & Nogueira Mendes, 2016)	Wikiloc, GPSies	Clasificación Ráster	Intensidad de uso de ciclismo de montaña	Parque Natural
(Heikinheimo et al., 2017)	Instagram	Análisis de contenido	% de fotos por actividad	Parque Nacional
(Norman & Pickering, 2017)	Wikiloc, GPSies, Map my fitness	Densidad de líneas	Densidad de rutas por actividad	Parque Nacional
(Márquez-Pérez J., et al., 2017)	Wikiloc	MIDE	Tiempos de viaje por sendero	Parque Nacional
(Mancini, Coghill, & Lusseau, 2018)	Flickr	Densidad Kernel	% de fotos por actividad	Parque Nacional
(Norman, Pickering, & Castley, 2019)	Map my fitness	Densidad Kernel	Intensidad de uso de ciclismo de montaña	Parque Nacional
(Barros et al., 2020)	Flickr, Wikiloc	Agregación de rutas	Intensidad de uso de senderismo	Parque Nacional

3.7. ¿Por qué visitan los espacios naturales protegidos?

Comprender las experiencias que los visitantes buscan y obtienen de un espacio protegido es esencial para la gestión exitosa del turismo (Newsome et al., 2012). Medir las preferencias y las percepciones sigue siendo una tarea difícil para los administradores e investigadores del turismo en espacios naturales protegidos porque obtener esta información puede ser costoso (Moore &

Weiler, 2009). Los métodos tradicionales para analizar las preferencias de los visitantes incluyen la realización de encuestas, entrevistas, grupos focales, evaluación de expertos (Eagles & McCool, 2002; Newsome et al., 2012). En este sentido, el contenido de las publicaciones en las redes sociales puede ser útil para evaluar lo que los visitantes perciben como valioso para compartir en Internet, proporcionando así nuevos conocimientos sobre la percepción de los visitantes sus preferencias y las razones por las que visitan un espacio protegido.

Por ejemplo, publicar imágenes deja huellas digitales de las preferencias espaciales, y las personas que comparten fotos específicas consideran que el lugar representado no solo "vale la pena visitar" sino también "vale la pena compartirlo visualmente" (Gliozzo et al., 2016). De los artículos revisados, catorce analizaron el contenido de las publicaciones de redes sociales para extraer las preferencias de los visitantes (Tabla 8).

La mayoría de los estudios que se han revisado realizan análisis de contenido de las fotografías geolocalizadas dentro de un área protegida para clasificar y extraer etiquetas y atributos que son de interés para los visitantes. Por ejemplo, (Gosal et al., 2019) analizó y clasificó el contenido de las fotos de Flickr para crear una tipología de turistas de acuerdo con el interés revelado en las imágenes, como el paisaje, la vida silvestre, la religión y la recreación. Dada la cantidad de datos recopilados, algunos estudios han aplicado el análisis de contenido automatizado mediante el uso de modelos de aprendizaje automático como Google Cloud Vision (Gosal et al., 2019) y Amazon Rekognition (Do & Kim, 2020). Otros estudios que utilizaron conjuntos de datos más pequeños realizaron una clasificación manual de las imágenes (Clemente et al., 2019; Hausmann et al., 2018; Richards & Friess, 2015) y codificación de imágenes (Pickering et al., 2020).

Con menos frecuencia, la investigación ha utilizado los textos que se encuentran en la descripción y las etiquetas de las fotos para extraer las preferencias de los visitantes. En este sentido, métodos como el análisis semántico han sido útiles para identificar las preferencias culturales y naturales de los visitantes por temporada (Pickering et al., 2020).

Una segunda línea de investigación que han seguido los artículos que abordaron esta cuestión consiste en utilizar modelos de regresión lineal para identificar las variables que influyen en el atractivo de los espacios protegidos. En este sentido, las variables que afectan las preferencias de los visitantes de acuerdo con los autores revisados son; la accesibilidad (Hausmann et al., 2017; Norman et al., 2019), la densidad de población vecina (Hausmann et al., 2019), y el tipo de cobertura vegetal (Callau et al., 2019).

Además, se han utilizado modelos de regresión para probar si los cambios en las condiciones originales y la configuración de un sitio podrían influir en la percepción de los visitantes. Por

ejemplo (Sinclair et al., 2018, 2019), mostró que la calidad del agua en los humedales es un factor importante para aumentar las visitas, así como para mejorar las experiencias de los visitantes.

Los hallazgos expuestos en este apartado abren nuevas oportunidades para explorar el contenido de las publicaciones en las redes sociales y para extraer información cualitativa, como las experiencias, las motivaciones y los niveles de satisfacción de los visitantes, que pueden ser valiosos para la promoción de los espacios protegidos como destinos de recreación.

Tabla 8. Artículos que abordan la pregunta ¿ Por qué visitan los espacios naturales protegidos?

Referencia	Fuente de datos	Método	Resultados	Espacios Protegidos
(Richards & Friess, 2015)	Flickr	Análisis de contenido	% de fotografías clasificadas	Reserva Natural
(Willemen, et al., 2015)	Flickr	Análisis de contenido	Indicador de atractivo por especie	Espacios Protegidos
(Hausmann et al., 2017)	Instagram	GLM	Variables explicativas de la visita	Espacios Protegidos
(Hausmann et al., 2018)	Flickr, Instagram	Análisis de contenido	Preferencias de los visitantes	Parque Nacional
(Kim et al., 2019)	Flickr	GWR	Variables espaciales de la visita	Parque Nacional
(Sinclair et al., 2018)	Flickr	GLM	Variables explicativas de la visita	Humedal
(Sinclair et al., 2019)	Flickr	GLM	Variables explicativas de la visita	Humedal
(Gosal, et al., 2019)	Flickr	Análisis de contenido	Tipología de turistas	Reserva Natural
(Clemente, et al., 2019)	Flickr	Análisis de contenido	Preferencias sociales	Parque Natural
(Callau, et al., 2019)	Wikiloc	Análisis de contenido	Caracterización del paisaje	Parque Natural
(Norman & Pickering, 2019)	Map my fitness, Strava, Wikiloc	GLM	Variables explicativas de la visita	Parques Nacionales
(Hausman et al., 2019)	Flickr	GLM	Indicador de atractivo	IBA
(Pickering, et al., 2020)	Flickr	Análisis de contenido	Preferencias estacionales de los visitantes	Parque Nacional
(Do & Kim 2020)	Flickr	Análisis de contenido	Preferencias de los visitantes	Parque Nacional

3.8. Conclusiones de la revisión de los artículos.

La revisión de los estudios en este capítulo reveló las ventajas que existen en utilizar los datos geocalizados de redes sociales. La ventaja principal está asociada con el acceso libre y descarga rápida de los datos lo que resulta en la reducción de tiempos destinados a la obtención y preparación de los datos al igual que la reducción de los costos asociados a dichos procesos. Por ejemplo la utilización de datos geocalizados como fuente de datos en la investigación de turismo de parques redujo la necesidad de trabajo de campo para la recolección de datos (Orsi & Geneletti, 2013).

Otra de las ventajas reportada en todas las publicaciones revisadas está relacionada con la granularidad de los datos tanto a nivel espacial como temporal lo que ha permitido analizar el comportamiento de los visitantes con mayor detalle y que antes no era posible con los datos tradicionales. Esto resulta una oportunidad importante para la investigación de los ritmos temporales de los visitantes y su distribución en el espacio y que sobre todo puede ser muy útil para analizar espacios protegidos en países que no cuentan con un sistema de seguimiento de uso público y los datos disponibles son escasos y desactualizados.

Las nuevas fuentes de datos requieren de algún tipo de validación o contraste con datos secundarios para poder establecer su representatividad como indicadores del turismo en espacios protegidos (Heikinheimo et al., 2017). La oportunidad de contar con una cantidad considerable de datos reportados por los mismos visitantes, en sí es una ventaja importante ya que contamos con datos directos que reflejan la percepción y preferencias de los usuarios y la posibilidad de recopilar datos continuos y cada vez más abundantes.

Con respecto a las fuentes de datos geocalizados, las redes sociales de fotografías se han posicionado como la fuente predilecta para la investigación del turismo en espacios protegidos y en particular Flickr, debido a su acceso fácil y con pocas limitaciones. El uso de Twitter como fuente de datos geocalizados no parece ser viable para la investigación del turismo de naturaleza, lo que se refleja en una sola publicación disponible que utiliza esta fuente de datos. Sin embargo, hacen falta más estudios que verifiquen la utilidad de Twitter más allá del componente de localización. Por ejemplo, hace falta análisis semánticos de los mensajes y hashtags que mencionen a los espacios protegidos para entender la percepción de los visitantes y población en general sobre el turismo y la calidad de las experiencias.

La literatura revisada refleja que la investigación y uso de datos geocalizados de redes sociales se encuentra en una etapa temprana de desarrollo por lo que la mayoría de los estudios analizados

están enfocados en estudios exploratorios que buscan establecer por un lado la representatividad de los datos y por otro los métodos empleados durante la recolección, procesamiento y análisis. Por lo que futuras líneas de investigación sobre las aplicaciones prácticas, integración con otras fuentes de datos son requeridas.

Un desafío de los datos geolocalizados de redes sociales expuesto en los artículos revisados es el comportamiento dinámico y la variabilidad de la popularidad de las redes sociales como sucedió con la aplicación de rutad MapMyFitness (Norman & Pickering, 2017), el cierre de plataformas como Panoramio (Barros et al., 2019), el relanzamiento de aplicaciones como GPSies ahora denominada AllTrails (Teles da Mota & Pickering, 2020), los cambios en la accesibilidad a los datos como en el caso de Instagram y Strava (Di Minin et al., 2015) y la restricción de descarga masiva de datos como Wikiloc (Mínguez García et al., 2016). No obstante, el uso de datos geolocalizados de redes sociales, la creación de nuevas plataformas y la cantidad de datos y contenido generado por usuarios continúa creciendo por lo que será necesario la continua validación de los datos para la investigación y gestión turística.

Esta nueva fuente de datos ofrece nuevas oportunidades para la investigación de la actividad turística. A pesar de las limitaciones producidas por el sesgo de cada fuente, los estudios expuestos muestran su fiabilidad para explicar patrones espacio-temporales de la actividad turística en espacios naturales protegidos y nos permiten tener una visualización dinámica del comportamiento del turista en el destino. De esta forma, los datos geolocalizados de redes sociales se posicionan como un complemento a las fuentes tradicionales como encuestas y estadísticas de organismos oficiales.

A la par, los estudios expuestos en este capítulo demuestran el potencial de las nuevas fuentes de datos para extraer información valiosa de diversos componentes para la investigación. Desde la perspectiva de la geografía, contar con datos geolocalizados de alta granularidad temporalidad ofrecen la oportunidad de aplicar nuevos métodos y herramientas para el análisis de fenómenos geográficos.

A diferencia de fuentes tradicionales, las nuevas fuentes de datos proporcionan un registro continuo del comportamiento de los turistas a un costo muy bajo y de acceso fácil. Por lo tanto, estos datos pueden ser una alternativa para monitorear la actividad turística y contar con indicadores para la planificación y gestión de destinos con pocos recursos y de difícil acceso.

4. Área de estudio.

4. Área de Estudio.

4.1. La red de Parques Nacionales de España: Breve historia.

En España, la declaración de un espacio como Parque Nacional se produce mediante la Ley de las Cortes Generales (OAPN, 2011) (Plan Director de La Red de Parques Nacionales, 2011). El origen de la red de parques nacionales se remonta a 1918, con el establecimiento de la legislación de Parques Nacionales y la declaración de los 2 primeros parques; Covadonga (actual Picos de Europa) y Ordesa y Monte Perdido. Con la declaración de estos espacios se abren las puertas para la protección de los ecosistemas únicos y representativos del territorio español. A partir de la legislación de 1918, se irán declarando parques nacionales en los ecosistemas más sobresalientes, llegando a contar con 15 parques distribuidos por toda la geografía española.

En la actualidad, la red de parques nacionales es un sistema que representan los ecosistemas y paisajes excepcionales y endémicos de la geografía española.

Figura 10. Mapa de localización de los 15 parques nacionales españoles.



A fin de que un espacio sea declarado como Parque Nacional, este debe reunir los requisitos establecidos por el Plan Director de la Red de Parques Nacionales. De acuerdo con el plan director se establecen seis criterios para la declaración:

- **Representatividad**, es decir el parque debe ser una muestra significativa del sistema natural al que representa.
- **La Extensión** del parque debe ser amplia para permitir la evolución natural y el funcionamiento de los procesos ecológicos.
- El espacio debe presentar **escasa intervención humana**.
- El territorio debe ser **continuo** y sin elementos que rompan la armonía de los ecosistemas.
- El parque no puede incluir en su interior **núcleos urbanos**.
- El territorio debe estar rodeado por un territorio susceptible de ser declarado **zona periférica de protección**.

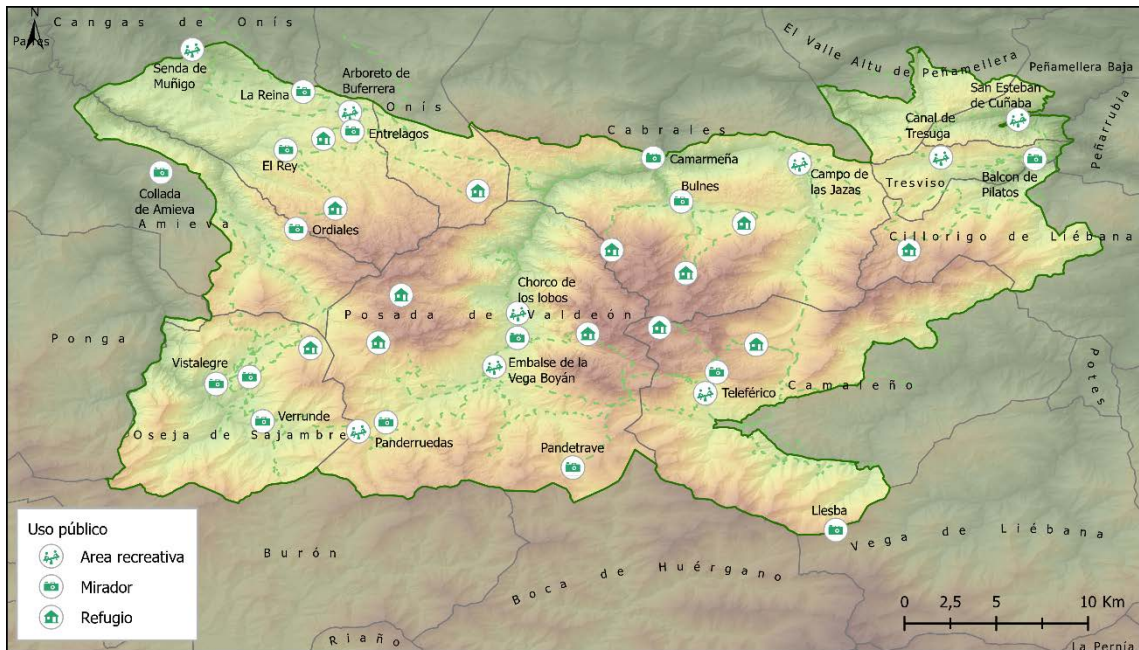
En el siguiente apartado, se describen en orden cronológico de creación las principales características de los 15 parques nacionales que forman la red.

4.2. Características de la red de parque nacionales.

Picos de Europa

El parque nacional Picos de Europa fue declarado en 1918 bajo el nombre de parque nacional de Covadonga con una extensión de 67.455 hectáreas. Se encuentra situado en la Cordillera Cantábrica y cubre 11 municipalidades y 3 provincias: Asturias (27.355 Ha), Cantabria (15.381Ha) y León (24.719Ha). El rasgo principal de los Picos de Europa son las formaciones de caliza de sus tres macizos, delimitados por profundos desfiladeros modelados a lo largo de dos millones de años por las cuencas de los ríos Sella, Dobra, Cares, Duje y Deva. El amplio rango altitudinal del parque hace posible una gran diversidad de paisajes naturales entre los que se pueden observar; bosques atlánticos, praderías de montaña y formaciones de matorral, así como roquedos propios de las grandes moles calcáreas de los tres principales macizos montañosos. A su vez esta gran diversidad aloja una gran biodiversidad que incluye al oso pardo, una de las especies más representativas de la cordillera cantábrica.

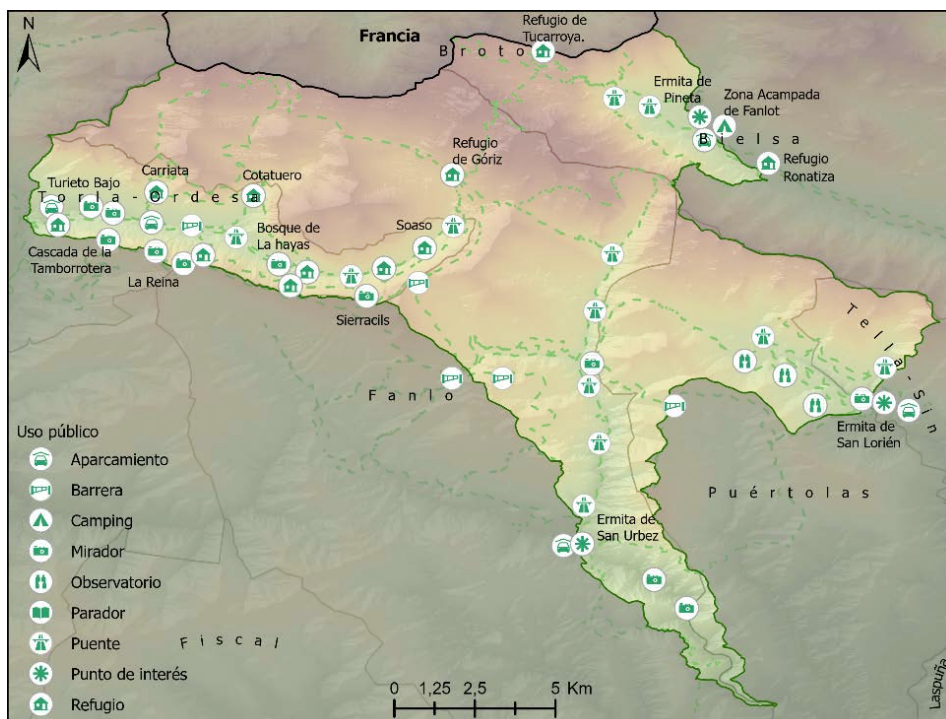
Figura 11. Mapa de Picos de Europa



Ordesa y Monte Perdido

Fundado en 1918, el parque nacional Ordesa y Monte perdido se localiza en los Pirineos centrales en la provincia de Huesca, de la comunidad autónoma de Aragón. El parque cuenta con una superficie de 15.696,20 hectáreas y fue declarado Patrimonio de la Humanidad de la Unesco en 1997. El paisaje del parque se encuentra dominado por el macizo Monte Perdido, la montaña calcárea más alta de Europa con 3355 msnm, desde donde derivan los valles de Ordesa, Pineta, Añisclo y Escuaín. Es un espacio de grandes contrastes entre las zonas de cumbres donde predomina la aridez extrema de los desiertos kársticos, pasando por los pastos suprafoestales hasta llegar a los escarpados valles o cañones de origen glaciar donde el agua en forma de cascada está siempre presente. Esta gran diversidad paisajística ha dado lugar a una elevada concentración de especies vegetales entre las que se encuentra especies endémicas como el zapatito de Venus (*Cypripedium calceolus*). El quebrantahuesos, un ave rapaz amenazada es la especie animal más representativa de este espacio (MAPAMA & OAPN, 2017).

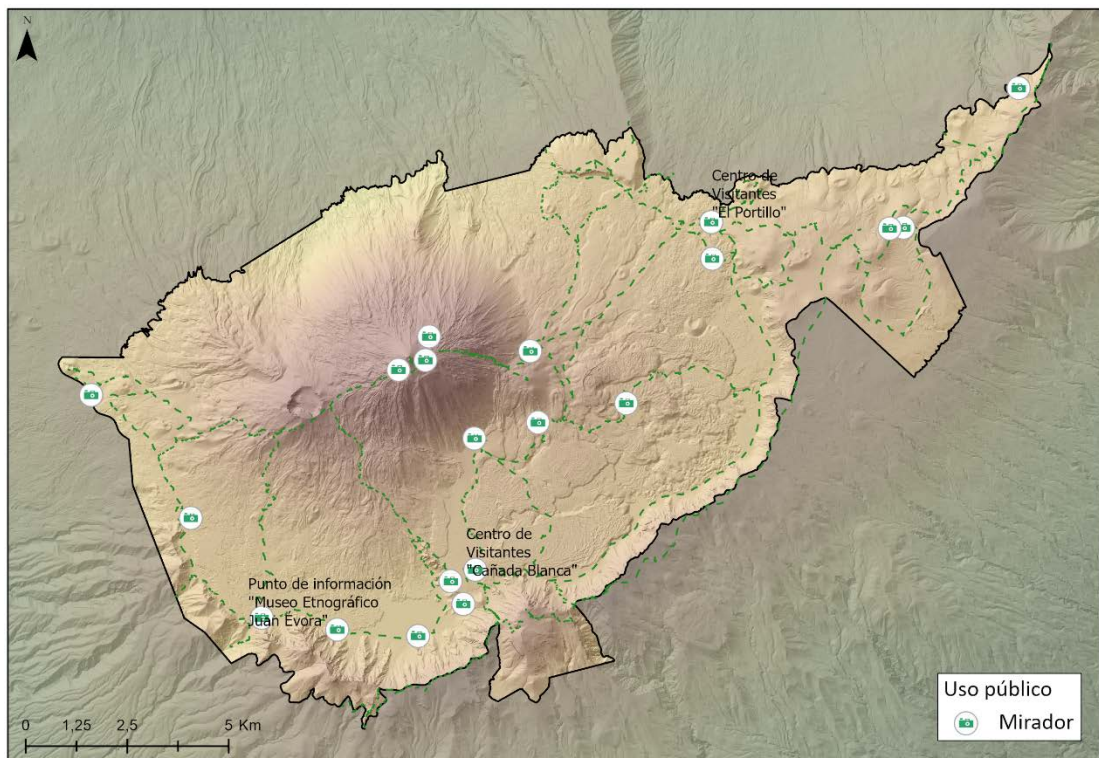
Figura 12. Mapa de Ordesa y Monte perdido



Cañadas del Teide

El Parque Nacional Cañadas del Teide, declarado Patrimonio de la Humanidad en 2007 por la Unesco, se localiza en el centro de la isla Tenerife en las Canarias. Tiene una superficie de 18.990 hectáreas lo que le convierte en el parque nacional más grande de las Islas Canarias. Es el tercero más antiguo de España tras su declaración por real decreto el 22 de enero de 1954. Con un total de 3.289.444 de visitantes en el año en 2015, es el parque nacional más visitado de España (OAPN, 2015). Dentro de sus límites se encuentra el volcán activo del Teide (3718 m.s.n.m), que le da su nombre y constituye el principal atractivo para sus visitantes. El Teide es el pico más alto de España y de cualquier tierra emergida del Atlántico. Este espacio protegido con su singular paisaje es el resultado de la actividad volcánica, en el que los conos volcánicos y flujos de lava forman un extraordinario conjunto de colores y formas y que ha resultado en la existencia de ecosistemas únicos y de gran valor. Representa una magnífica muestra del vulcanismo y de la adaptación de las formas de vida a difíciles condiciones ambientales (Hernández Álvarez, 2017).

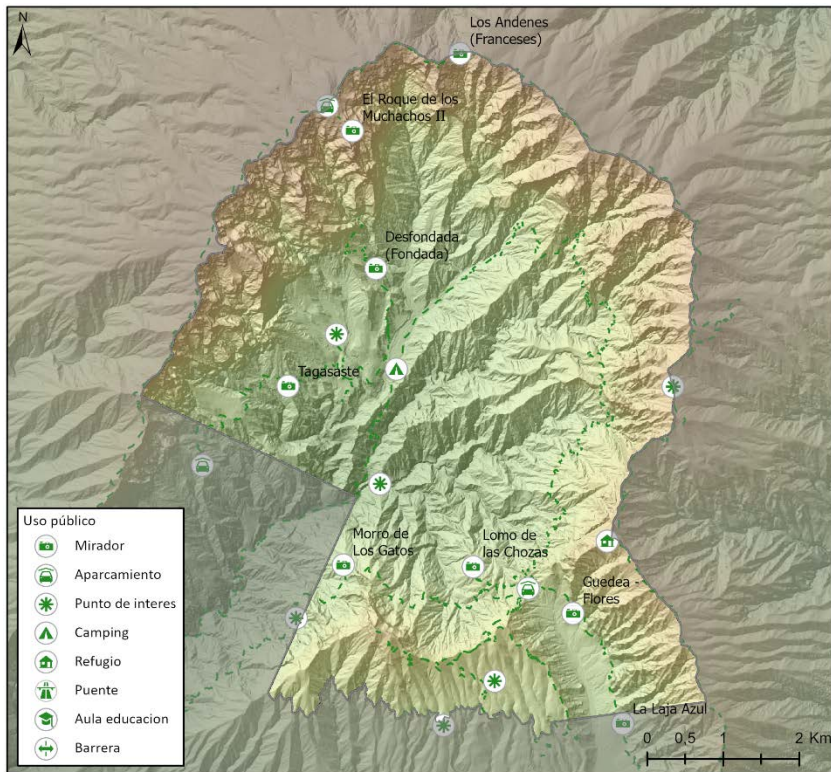
Figura 13. Mapa del parque nacional Cañadas del Teide



Caldera de Taburiente

El parque nacional Caldera de Taburiente fue establecido en 1954 y se encuentra situado en el centro de la isla de la Palma en la provincia de Santa Cruz de Tenerife de las islas Canarias. Con una superficie de 4690 hectáreas, su paisaje se halla dominado por un circo de cumbre de 8 km de diámetro con desniveles de hasta 2.000 m, con una red de arroyos y torrentes espectacular y de gran fuerza erosiva. Este paisaje escarpado se caracteriza por la presencia de crestas o “roques”, que son el resultado de la erosión de diferentes materiales volcánicos por la acción del agua. En este medio se ha desarrollado una gran y valiosa biodiversidad, en la que destaca la vegetación de cumbre que acoge un gran número de endemismos de las Canarias tales como benconia, retamón y la violeta o tajinaste rosado.

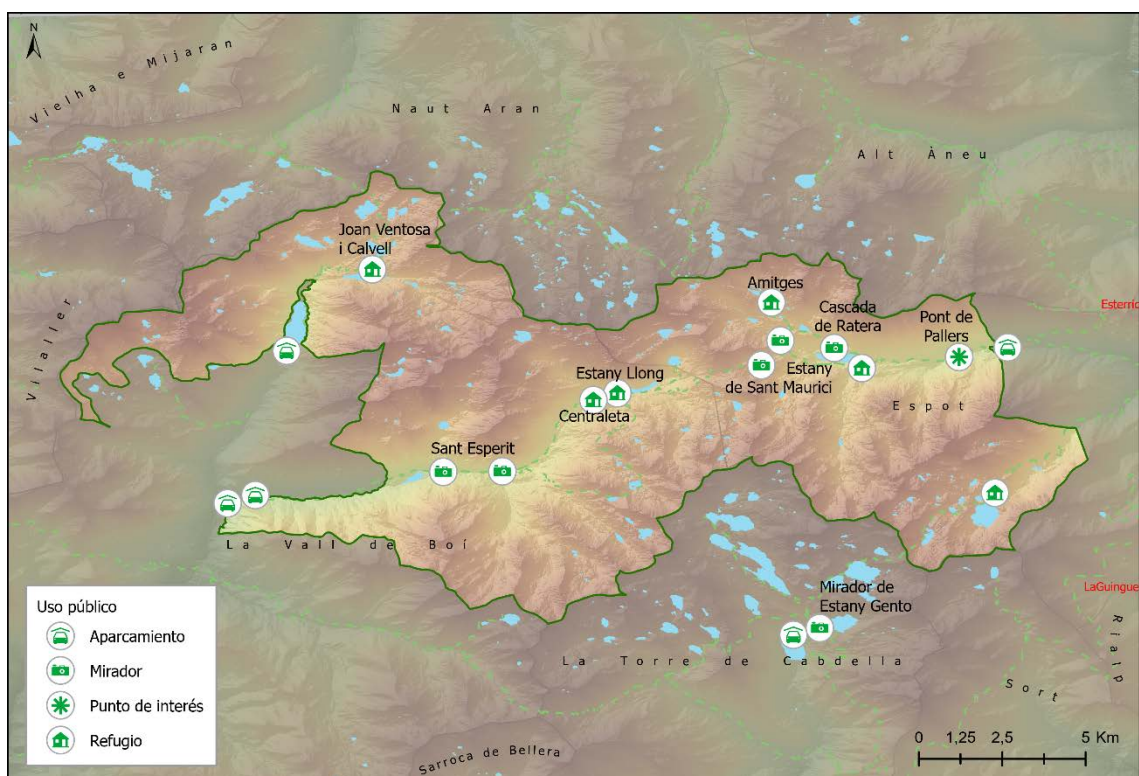
Figura 14. Mapa del parque nacional Caldera de Taburiente



Aigüestortes i Estany de Sant Maurici

El parque nacional de Aigüestortes i Estany de Sant Maurici fue creado en 1955 y se encuentra situado en el Pirineo Catalán, en la provincia Lleida. El parque cuenta con 14.119 hectáreas y constituye la zona lacustre más importante de los Pirineos. Por esta razón, el agua es la principal protagonista de este espacio y su signo de identidad, ya que cuenta con más de 200 lagos o estanys. A los numerosos lagos se suman los impresionantes riscos de “Els Encantats” y sus meandros de alta montaña conocidos como “Aigüestortes”, lo que lo convierten en un paisaje excepcional de gran atractivo para los amantes de la naturaleza y el hogar de una rica biodiversidad de origen alpino o boreal y entre la que se puede nombrar; especies vegetales como abetos y pinos negros, abedul y haya, al igual que varias especies animales entre las que destacan el urogallo y el mochuelo boreal.

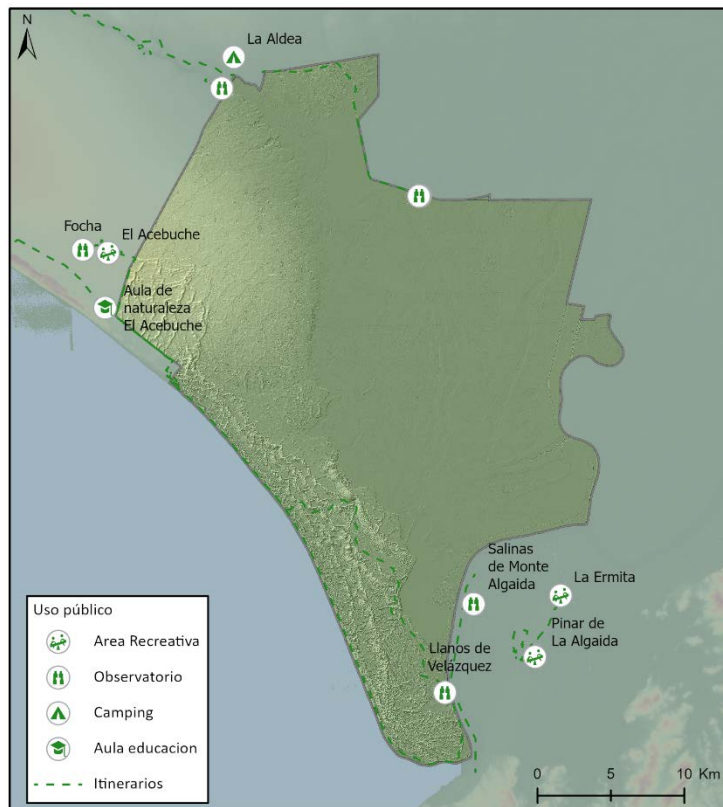
Figura 15. Mapa del parque nacional Aigüestortes i Estany de Sant Maurici



Doñana

Establecido en 1969, el parque nacional Doñana se sitúa en el extremo suroccidental de la península ibérica, en la desembocadura del río Guadalquivir en las provincias de Huelva y Sevilla. Con una superficie de 54252 hectáreas, el parque está conformado por un mosaico de playas, dunas, cotos y marismas que dan lugar a una gran variedad de biotopos de interés para varias especies vegetales y animales, algunas de ellas en peligro de extinción como el lince ibérico y el águila imperial ibérica. Entre los ecosistemas que alberga, destaca la marisma por su importancia como lugar de paso, cría e invernada para varias especies de aves europeas y africanas. Esta inmensa riqueza faunística ha sido un factor atrayente para ornitólogos y observadores de aves de todo el mundo. Además de la marisma, en el parque se encuentran otras unidades ambientales excepcionales: la playa, las dunas móviles y los corrales o cotos y numerosas lagunas dispersas entre la vegetación y una zona de transición que conecta y separa a todos estos paisajes, la vera. El parque fue declarado patrimonio de la humanidad por la UNESCO en 1994.

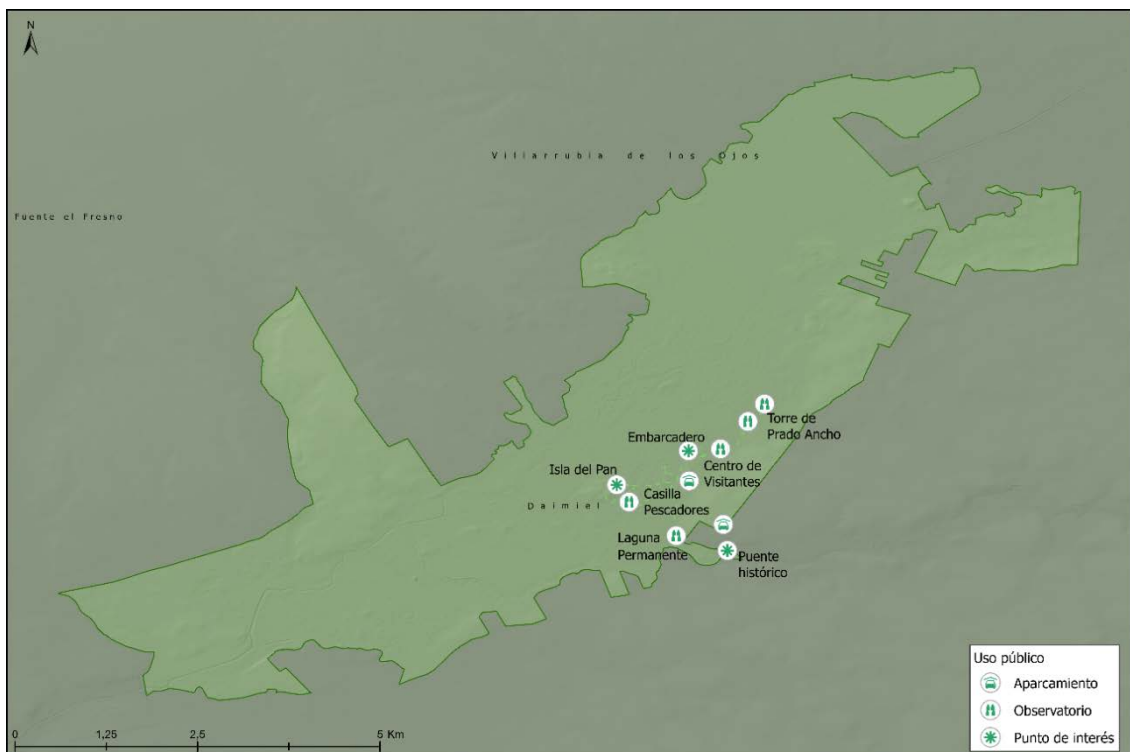
Figura 16. Mapa del parque nacional Doñana



Tablas de Daimiel

El parque nacional Tablas de Daimiel fue creado en 1973 y tiene una extensión de 3030 hectáreas. Se encuentra localizado en la provincia de Ciudad Real, dentro de la comarca de La Mancha y al sur de los montes de Toledo. Las Tablas de Daimiel son un humedal único en Europa, pues constituye el último representante del ecosistema de llanuras de inundación, localmente denominado como tablas. Este complejo ecosistema se caracteriza por las inundaciones producidas por los desbordamientos de los ríos Guadiana y Cigüela que permiten el desarrollo de una potente cubierta vegetal que, a su vez permite la supervivencia de la avifauna acuática que utiliza este espacio como zona de invernada, nidificación y paso. Entre esta gran variedad de aves destacan el somormujo lavanco, el zampullín común, garzas y todo tipo de anátidas ibéricas. En tanto que la diversidad vegetal del parque se caracteriza por la predominancia de plantas acuáticas y los árboles tarayes.

Figura 17. Mapa del parque nacional Tablas de Daimiel



Timanfaya

El parque nacional Timanfaya fue creado en 1974 y se encuentra localizado en el sector centro-occidental de la isla de Lanzarote en las Canarias. El parque cuenta con una extensión de 5107,5 hectáreas y constituye un buen exponente del vulcanismo reciente en el archipiélago canario. El proceso volcánico que dio origen al parque es el resultado de dos procesos eruptivos, el primero se inició en 1730 y estuvo activo hasta 1736 y el segundo tuvo lugar en 1824. Ambas erupciones provocaron la aparición de conos volcánicos, flujos y mares de lava y alteraciones geotérmicas que le dieron al parque su singular e inolvidable paisaje. Esto, junto con las condiciones climáticas áridas y adversas han permitido mantener intacto el paisaje desde que sucedieron las erupciones y a la vez la existencia de una diversidad animal y vegetal con gran valor adaptivo.

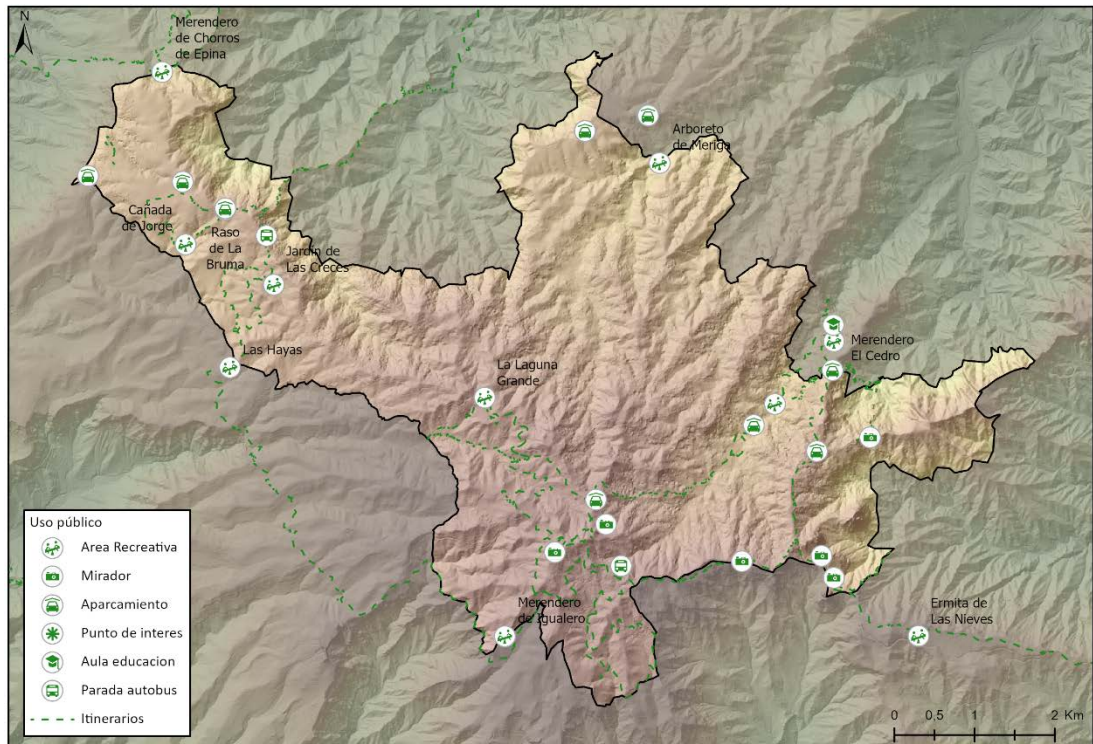
Figura 18. Mapa del parque nacional Timanfaya



Garajonay

El parque nacional Garajonay fue creado en 1981 y se sitúa en las cumbres de la isla de La Gomera en la provincia de Santa Cruz de Tenerife, en las islas Canarias. El parque tiene una extensión de 3984 hectáreas y destaca por ser el exponente del bosque laurisilva canaria, un ecosistema singular que representa lo que queda de los bosques que hace varios millones de años cubrían gran parte de Europa y África. Su vida vegetal y animal son únicas, es decir, no se pueden encontrar en ningún lugar fuera de la isla y su supervivencia está ligada a las características climáticas de este espacio que propician la humedad y exuberancia. El bosque de Garajonay fue declarado patrimonio de la humanidad de la UNESCO en 1986.

Figura 19. Mapa del parque nacional Garajonay



Archipiélago de Cabrera

Establecido en 1991, el parque nacional archipiélago de Cabrera se encuentra localizado a 10 millas náuticas al sur de la isla de Mallorca en las Baleares. El parque cuenta con 10.021 hectáreas de las cuales 1318 hectáreas son terrestres y 8703 hectáreas corresponden a la zona marina. El archipiélago comprende a isla principal de Cabrera, la isla Conejera y un conjunto de quince islotes que constituyen el más claro ejemplo de ecosistemas insulares no alterados del Mediterráneo español. El parque está dominado por arbustos leñosos formadores de la garriga, matorrales endémicos y bien adaptados a los rigores del clima mediterráneo, entre los que destacan el astrágalo de la Baleares, hipericón balear y el tragamoscas. Asimismo, el sitio es un importante punto de escala en la ruta migratoria de al menos 150 especies de aves marinas como la gaviota de Audouin. Los fondos marinos también son una parte importante del parque debido a su gran diversidad de peces e invertebrados.

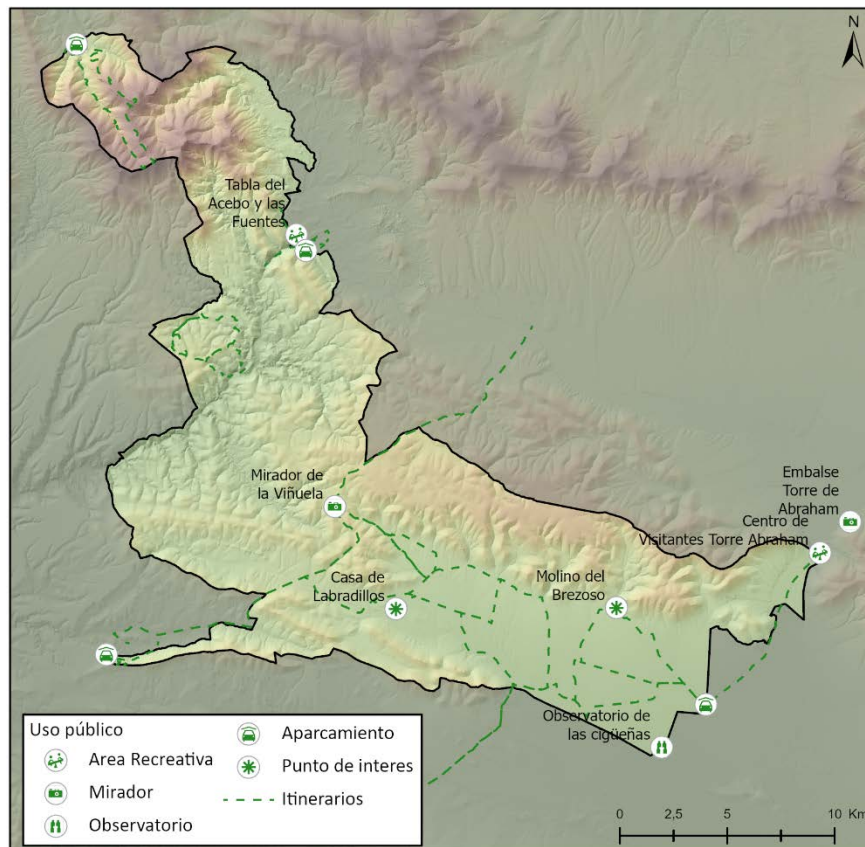
Figura 20. Mapa del parque nacional Archipiélago de Cabrera



Cabañeros

El parque nacional Cabañeros fundado en 1995, se localiza en el corazón de los montes de Toledo, entre las provincias de Toledo y Ciudad Real. Con una extensión de 40.856 hectáreas, el parque abarca los macizos del Rocigalgo y El Chorito y es considerado una de las mejores representaciones del bosque mediterráneo en España. El paisaje de Cabañeros se caracteriza por la presencia de extensas rañas con ricos pastos estacionales que ofrecen alimento a multitud de animales y una gran variedad de árboles y arbustos propios del mediterráneo. Algunas de las especies más representativas de su diversidad animal son las aves rapaces como el águila imperial Ibérica, la cigüeña y el buitre negro y los mamíferos como el ciervo y jabalí. El parque debe su nombre a las chozas tradicionales utilizadas por pastores y carboneros como refugio temporal durante sus labores en el campo.

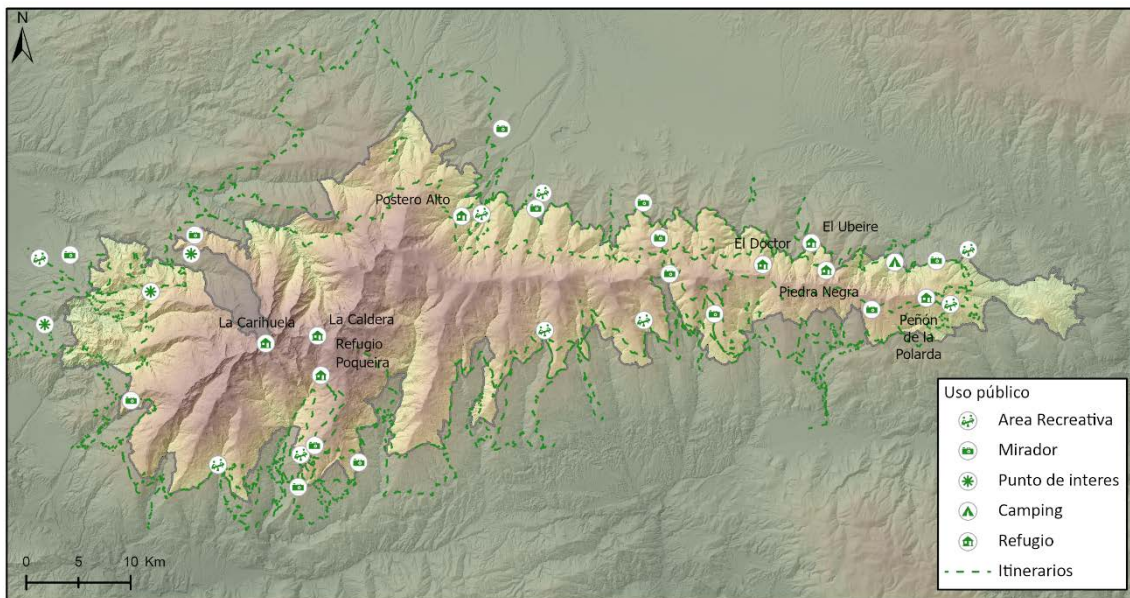
Figura 21. Mapa del parque nacional Cabañeros



Sierra Nevada

El parque nacional Sierra Nevada fue creado en 1999 y se sitúa en las provincias de Almería y Granada. Con una extensión de 85.883 hectáreas, el parque cubre 44 municipios y es el parque nacional más grande de la red. La Sierra Nevada es un extenso macizo montañoso con un relieve compacto, en el que destaca el Mulhacén, que con 3479 m.s.n.m. es la cima más alta en la península ibérica. El parque representa a los sistemas naturales ligados a la media y alta montaña mediterránea y cuenta con una rica biodiversidad entre las que cuentan unas 2100 especies vegetales, de las cuales 66 son endémicas, varias especies de anfibios, reptiles, mamíferos y aves. Entre los mamíferos, la cabra montesa (*Capra pyrenaica*) es la especie más emblemática del parque.

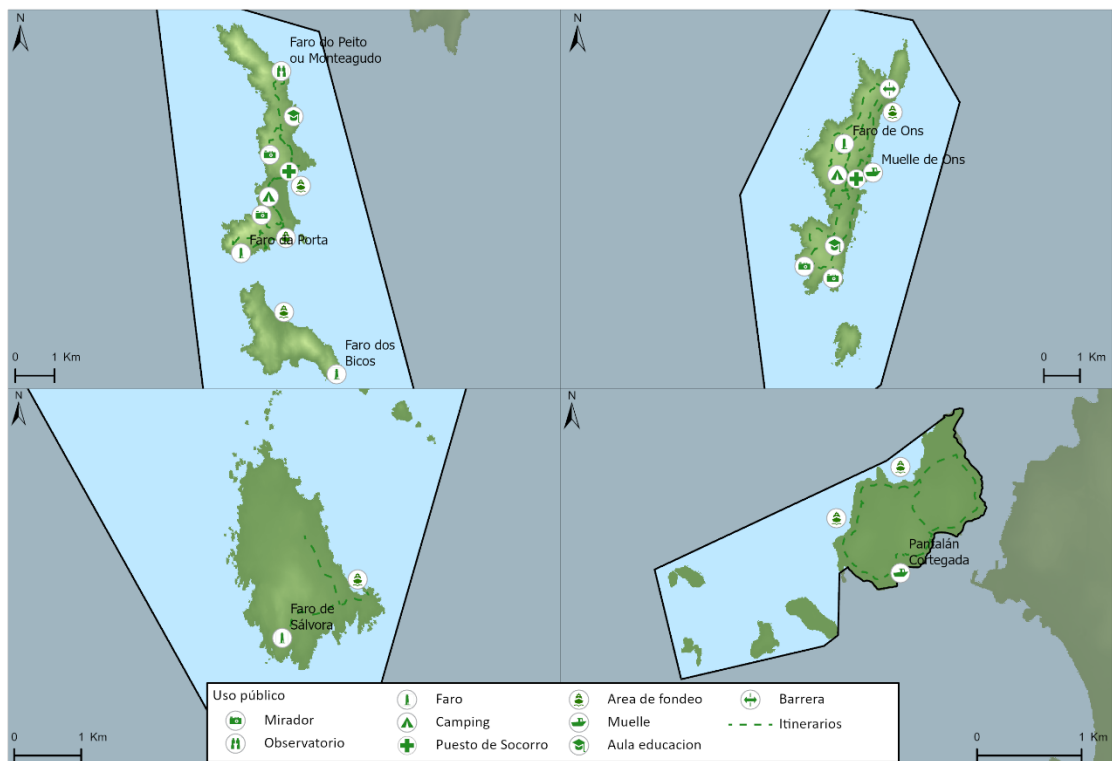
Figura 22. Mapa del parque nacional Sierra Nevada



Islas Atlánticas de Galicia

El parque nacional de Islas Atlánticas de Galicia fue establecido en 2002 y comprende las islas gallegas de Cortegada, Sálvora, Ons y Ciés, pertenecientes a las provincias de Pontevedra y A Coruña. Tiene una superficie 8480 hectáreas de las cuales; 1194,8 son terrestres y 7285,2 son marinas. Su localización frente a las Rías Baixas crea una barrera natural ante el océano que acentúa la belleza del ambiente marino con sus acantilados. Las islas ofrecen dos tipos de relieve; el occidental enfrentado al mar abierto y más abrupto y el oriental, mirando a la costa de perfil más suave con playas y dunas. Este medio ha propiciado el desarrollo de una gran biodiversidad marina expresada en las comunidades de bosques de algas pardas, gran variedad de peces y crustáceos. En tanto que en el medio terrestre ha dado hogar a las aves marinas entre las que destaca el cormorán moñudo que tiene una importante población en el parque. El parque estuvo habitado desde tiempos antiguos, sin embargo, las dificultades causadas por el aislamiento y las duras condiciones climáticas han resultado en que en la actualidad solo la isla Ons se encuentre habitada.

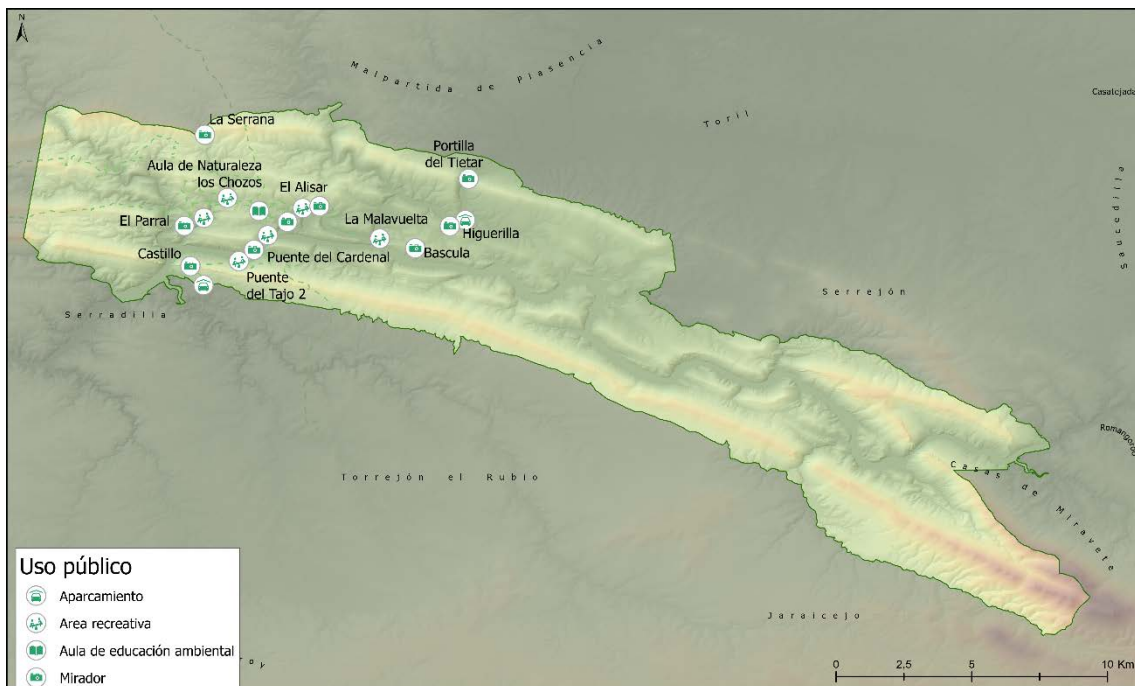
Figura 23. Mapa del parque nacional Islas Atlánticas de Galicia



Monfragüe

Establecido en 2007, el parque nacional Monfragüe se sitúa en la confluencia del río Tíetar con el Tajo, hacia el centro de la provincia de Cáceres en Extremadura. El parque cuenta con una superficie de 18.396 hectáreas y se ha consolidado como un verdadero santuario para observar aves. En Monfragüe se encuentra la más extensa y mejor conservada mancha de monte mediterráneo a nivel mundial. El parque tiene una gran variedad de biotopos; encinares, bosque y matorral mediterráneo, roquedos, embalses y arroyos, que alojan una rica diversidad vegetal y animal. Este valle ha estado habitado por los seres humanos desde hace miles de años, como lo atestiguan las pinturas rupestres que se hallan en su territorio. Entre la fauna, las especies más representativas del parque son el buitre negro, el buitre leonado, el águila imperial y la cigüeña negra.

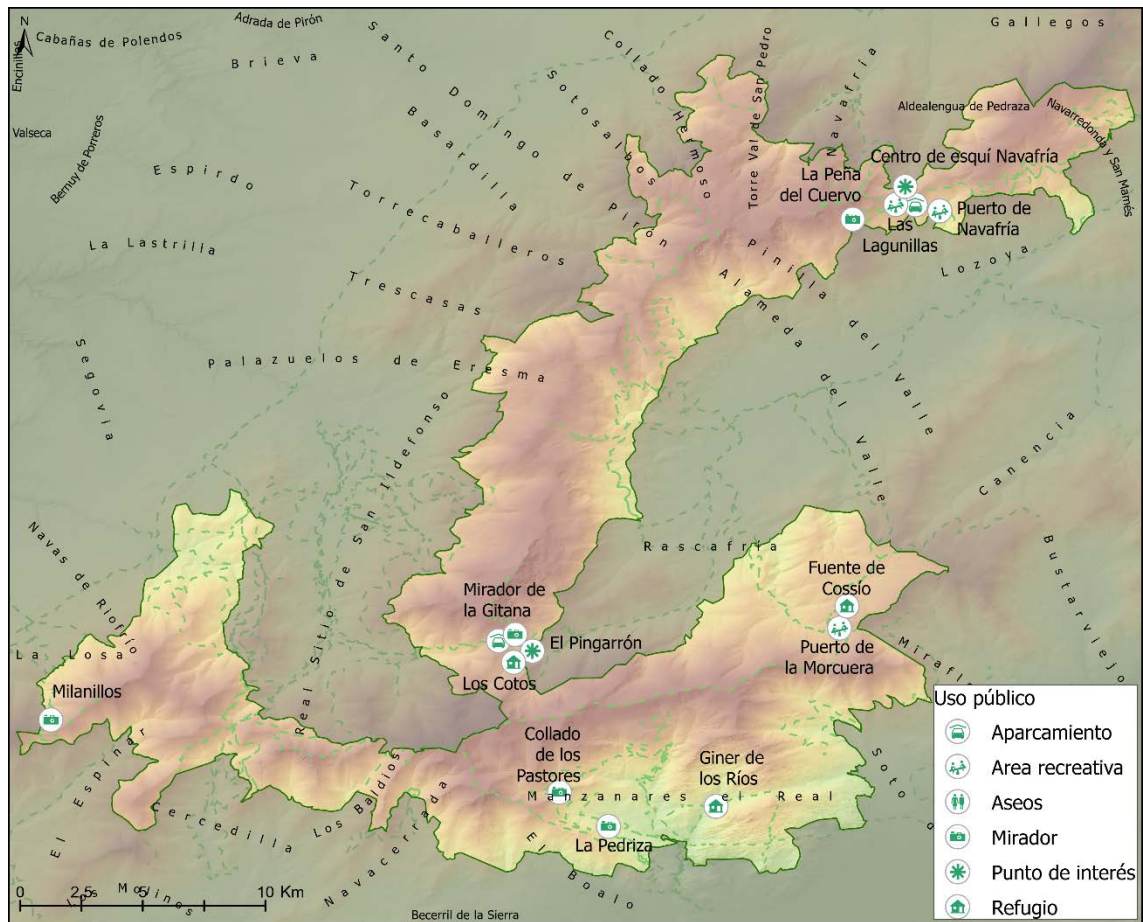
Figura 24. Mapa del parque nacional Monfragüe



Sierra de Guadarrama

El parque nacional Sierra de Guadarrama fue creado por decreto en 2013 y se encuentra localizado entre las comunidades autónomas de Castilla y León (Segovia) y Madrid. El parque ocupa 33.960 hectáreas de la Sierra de Guadarrama en el sistema central, macizo montañoso que divide en dos la meseta castellana y separa las cuencas hidrográficas del Duero y el Tajo. El pico Peñalara con 2428 m.s.n.m. es la cima de mayor altitud y junto con otras cuatro cubres aledañas forman la cresta del Macizo de Peñalara. Otra cumbre de interés son los Siete Picos que discurre entre el puerto de Navacerrada y el Puerto La Fuenfría. En su paisaje destacan elementos geológicos singulares como los circos, lagunas glaciares y roquedos graníticos. La Sierra de Guadarrama constituye una muestra representativa de los ecosistemas de alta montaña mediterránea tales como: matorrales y pastizales alpinos, pinares, melojares y turberas.

Figura 25. Mapa del parque nacional Sierra de Guadarrama



5. Datos utilizados: fuentes, procesos de descarga y tratamiento previo.

5. Datos utilizados: fuentes, procesos de descarga y tratamiento previo.

Este capítulo presenta los datos utilizados y describe la metodología implementada para descargar, depurar y preparar los datos geolocalizados provenientes de dos redes sociales (Flickr y Wikiloc), con el objetivo de extraer información relevante sobre el comportamiento de los visitantes de parques nacionales. Los resultados obtenidos presentan una idea general del volumen y distribución espacial de los datos y su potencial para ser utilizados en análisis espacial.

La primera sección contiene una breve descripción de las fuentes de datos utilizadas en esta investigación, los datos cartográficos, los datos oficiales de conteos de visitantes y los datos de redes sociales (Flickr y Wikiloc). El objetivo es tener una visión general de qué tipo de datos proporciona cada plataforma y cuáles atributos son relevantes para este estudio. Además, se presentan los datos secundarios, provenientes de fuentes oficiales, utilizados como insumos para la ejecución de diferentes análisis espaciales, por ejemplo, los datos de las redes de senderos para su uso en análisis de rutas.

Luego, en la segunda sección se explican los distintos procedimientos aplicados para eliminar el ruido e inconsistencias encontradas después de explorar los datos. Finalmente, se describe el tratamiento aplicado a los datos limpios para su uso en los análisis posteriores. Por ejemplo, se realizó la agregación espacial de las fotografías geolocalizadas de Flickr (puntos) en unidades espaciales regulares (hexágonos).

La última sección presenta los resultados obtenidos de la descarga de los datos de las dos redes sociales analizadas, haciendo uso de gráficos y estadísticas descriptivas se muestran las características principales de ambas fuentes de datos.

5.1. Fuentes de datos cartográficos y estadísticas oficiales.

Además de las fuentes empleadas para la consecución de los objetivos de esta investigación, es necesario recopilar datos secundarios que son necesarios para la realización de varios análisis dentro del proceso metodológico. De acuerdo con su tipo, los datos secundarios utilizados en esta tesis se clasifican en cartográficos y estadísticos.

Los datos cartográficos comprenden todos los datos geográficos que se utilizaron como: datos de entrada para la realización de análisis espaciales, por ejemplo, ajuste de rutas (map matching), y datos complementarios para la realización de mapas y parámetros para configurar las zonas de

descarga de datos. Mientras que los datos estadísticos se refieren a los datos sobre conteos mensuales, anuales y características de los visitantes recopilados de las memorias anuales producidas por el OPAN y que se utilizaron como contraparte para la validación de la representatividad de los datos geolocalizados.

5.1.1. Datos cartográficos.

Límites administrativos de los parques nacionales.

La cartografía vectorial de los límites administrativos de los parques nacionales se descargó de la página web del Organismo Autónomo de Parques Nacionales (OAPN) y consiste en 15 capas vectoriales de geometría de polígono en formato shapefile. La información disponible para descarga desde la página del OAPN comprende los límites de los parques nacionales establecidos de acuerdo con las leyes de declaración de cada parque y los límites de las zonas periféricas de protección (OAPN, 2018).

Infraestructura de uso público.

La cartografía vectorial de la infraestructura de uso público de la red de parques nacionales españoles consiste en 3 capas vectoriales con información sobre:

- La localización de los puntos de información y centros de visitantes en forma de geometría de puntos.
- La localización de la infraestructura de uso público por ejemplo puentes, miradores, cafeterías, paraderos, entre otros en forma de geometría de puntos.
- Los itinerarios diseñados por el OAPN para el uso público en forma de geometría de líneas. Los itinerarios consisten en recorridos sugeridos por el OAPN que además de la geometría contienen datos sobre la duración y dificultad de cada itinerario.

Estos datos se obtuvieron, tras la petición de información realizada al servicio de mapas del OAPN.

Redes de senderos de los parques nacionales españoles.

Open Street Maps (OSM) es un proyecto colaborativo para crear y editar cartografía digital libre que se actualiza constantemente por sus usuarios (OpenStreetMap, 2021). Gracias a OSM, es posible acceder a cartografía de redes de transporte y localización de infraestructura y

equipamientos de todo el mundo a través de una API. OSM representa las redes de transporte mediante tres estructuras de datos básicas (nodos, ejes y relaciones) y se organizan mediante etiquetas. Cada etiqueta describe un atributo geográfico de la característica que se muestra en ese nodo, vía o relación específica (OpenStreetMap, 2021).

Para el análisis y visualización de los resultados generados en esta investigación, fue necesario contar con cartografía actualizada de los senderos que existen dentro de los parques y de la infraestructura presente y que no se encuentran mapeados de forma oficial por el OAPN. Por lo tanto, se realizó la descarga de la información geográfica mediante el acceso a la API de Open Street Maps y utilizando un script de Python y la librería de descarga de open Street maps "osmnx". El script realiza la petición de la información sobre los senderos basándose en dos aspectos; los límites de cada parque y en las etiquetas de OSM, luego descarga los datos, crea una red de grafos, usando la librería de análisis de redes "networkx" y almacena la red en formato shapefile. A partir de esta red se construyeron , datasets de red para cada parque, utilizando las herramientas de análisis de redes de ArcGIS Pro.

Figura 26. Ejemplo de código para descarga de datos de Open Street Maps

```
In [18]: # Import libraries

import geopandas as gpd
import osmnx as ox
import networkx as nx
import datetime, itertools
from collections import deque
from networkx.algorithms.shortest_paths.generic import _build_paths_from_predecessors
import folium

%matplotlib inline
ox.__version__

Out[18]: '1.1.1'
```

```
In [25]: # Define a bounding box

north, south, east, west = 42.625081, 42.607948, -8.77772, -8.806223

39.223822, 39.107049, 2.998919, 2.890568 # Cabrera
28.764358, 28.677325, -17.832085, -17.911416 # Taburiente
39.292662, 39.092262, -3.631107, -3.789647 # Daimiel
9.919224, 39.696201, -5.723526, -6.12356 # Monfrague
42.650783, 42.480454, 1.084023, 0.778679 # Aiguastortes
29.071121, 28.933933, -13.709445, -13.848395 # Timanfaya
42.258076, 42.164867, -8.87228, -8.932743 # Cies
42.408114, 42.322995, -8.907464, -8.962255 # Ons
42.516066, 42.439426, -8.983451, -9.066339 # Salvora
42.625081, 42.607948, -8.77772, -8.806223 # Cortegada

Out[25]: (42.625081, 42.607948, -8.77772, -8.806223)
```

```
In [26]: # Download data from OSM

graph3 = ox.graph.graph_from_bbox(north,south,east,west,network_type="walk")

# Save graph into a shapefile
ox.save_graph_shapefile(graph3, './cortegada.shp',)
```

5.1.2. Estadísticas oficiales: Memorias del Organismo Autónomo de Parques Nacionales.

Los datos sobre número de visitantes anuales y mensuales para el periodo 2010-2018 se obtuvieron de las memorias anuales que produce el OAPN para cada uno de los parques que forman la red nacional. Es importante mencionar, que las metodologías de conteo de visitantes varían entre los parques, al igual que la disponibilidad de los datos y su nivel de detalle, destacando el parque Sierra Nevada como el espacio que no tenía datos sobre el número mensual de visitantes para todos los años.

Tabla 9. Disponibilidad datos visitantes por año y mes

Parque	2010	2011	2012	2013	2014	2015	2016	2017	2018
Teide	x	x	x	x	x	x	x	x	x
Timanfaya	x	x	x	x	x	x	x	x	x
Caldera de Taburiente	x	x	x	x	x	x	x	x	x
Garajonay	x	x	x	x	x	x	x	x	x
Archipiélago de Cabrera	x	x	x	x	x	x	x	x	x
Aigüestortes	x	x	x	x	x	x	x	x	x
Ordesa y monte perdido	x	x	x	x	x	x	x	x	x
Cabañeros	x	x	x	x	x	x	x	x	x
Monfragüe	x	x	x	x	x	x	x	x	x
Tablas de Daimiel	x	x	x	x	x	x	x	x	x
Sierra de Guadarrama					x	x	x	x	x
Picos de Europa	x	x	x	x	x	x	x	x	x
Doñana	x	x	x	x	x	x	x	x	x
Sierra Nevada	x	x	x	x	x	x	x	x	x
Islas Atlánticas de Galicia	x	x	x	x	x	x	x	x	x

x = disponibilidad datos anuales

Color = Disponibilidad datos mensuales

5.2. Fuentes de los datos de redes sociales.

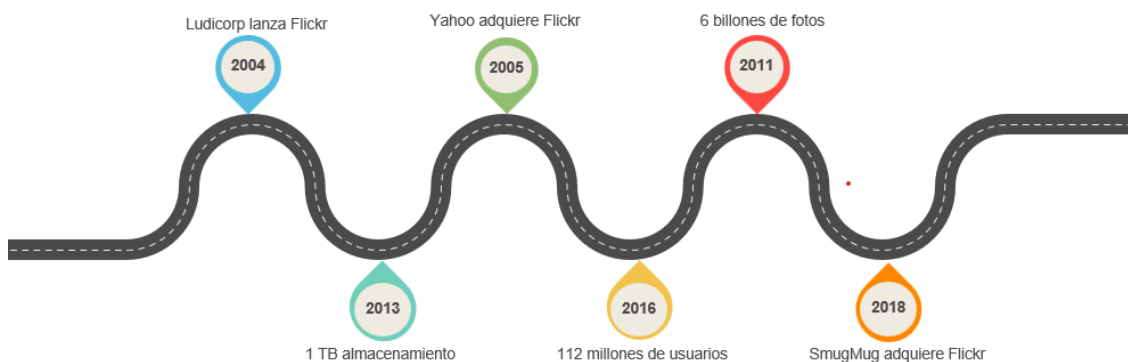
5.2.1. Flickr.

Flickr es una plataforma en línea para almacenar y compartir fotografías y videos en línea que fue lanzada en el 2004 (The Internet Archive, 2016). Esta plataforma ofrece dos tipos de cuentas; una gratuita que permite almacenar hasta un máximo de 1000 fotografías y una de pago diseñada para fotógrafos profesionales, sin límite de almacenamiento (Flickr, 2018). En el 2016, la plataforma almacenaba un total de 7000 millones de fotografías correspondientes a 112 millones de usuarios

de 63 países (Flickr, 2017a). En el 2006, la plataforma introdujo la opción de geolocalización, lo que permite a sus usuarios localizar sus fotografías mediante el GPS de la cámara o seleccionando la ubicación en un mapa mundial al cargar las fotos desde un ordenador. Flickr agrega automáticamente un valor de precisión espacial para todas las fotos geolocalizadas, basado en el nivel de zoom o escala del mapa al momento de tomar la fotografía. Los valores varían en una escala que va desde: mundo (1), país (3), región (6), ciudad (11) y nivel de calle (16).

Se ha seleccionado Flickr como fuente de datos por tres razones; en primer lugar, Flickr dispone de una API de libre acceso que permite la descarga de fotos geolocalizadas de forma masiva y gratuita. En segundo lugar, los metadatos asociados a cada fotografía geolocalizada tienen datos sobre sus usuarios que pueden ser muy útiles para el análisis del fenómeno turístico. Por ejemplo, el registro de la fecha y hora junto con la localización revelan el paradero del turista en el espacio-tiempo. Además que la toma de fotografías se hallan relacionadas estrechamente al turismo (Garrod, 2008). Finalmente, al contar con secuencias de fotografías se espera poder construir la trayectoria de los visitantes dentro de los parques y de esta manera analizar las pautas generales de comportamiento espacio temporal de los visitantes.

Figura 27. Evolución de la red social Flickr



Elaboración propia. Fuente: (Jenkins, 2017)

5.2.2. Wikiloc.

Wikiloc es un sitio web que permite a sus usuarios compartir rutas, puntos de interés GPS y fotografías de todo el mundo. La plataforma fue lanzada en 2006 por Jordi Ramot e incluye un sitio web y aplicaciones móviles de Android y iPhone. Tiene dos modalidades de acceso; la primera es creando una cuenta gratuita que permite el subir y descargar rutas GPS y una segunda opción que se denomina Wikiloc Premium y que aumenta las funcionalidades a sus usuarios permitiendo el seguimiento en vivo y la navegación sin anuncios.

En 2017, Wikiloc tenía alrededor de 4,7 millones de usuarios y más de 12,6 millones de rutas almacenadas. El senderismo y el ciclismo agrupan la mayoría de las rutas del sitio (Wikiloc, 2017). Al contrario de Flickr, esta plataforma no tiene una interfaz de programación de aplicaciones (API), que permita descargar las rutas de forma automática.

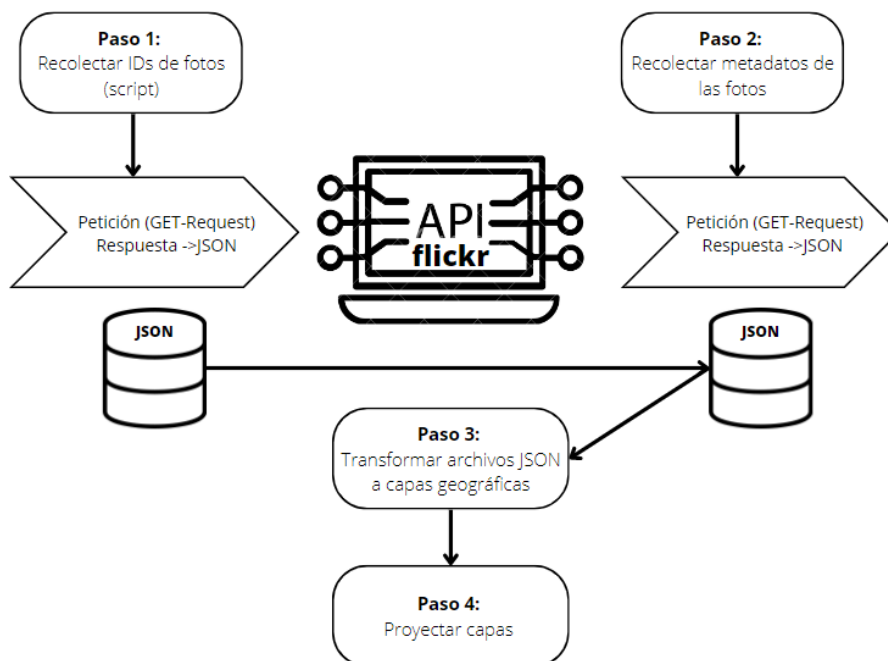
Se ha seleccionado Wikiloc como fuente de datos, porque esta red está dirigida a un grupo específico de visitantes que se enfoca en deportes al aire libre. Por esta razón, los datos provenientes de esta red social permitirían el análisis de actividades distintas a las reportadas con los datos de Flickr y que también se realizan en parques nacionales, por lo que pueden identificar a un tipo de turismo diferente. Además, los datos GPS, con mayor granularidad espacial y temporal, permiten evaluar el movimiento de sus usuarios con mayor detalle y recrear los itinerarios seguidos durante su visita.

5.3. Procesos de descarga de los datos geolocalizados de redes sociales.

Flickr.

La descarga de las fotografías se llevó a cabo a través de la interfaz de programación de aplicaciones (API) de la web de Flickr. La interfaz proporciona acceso gratuito a las fotografías para su uso no comercial, a través de la creación y envío de peticiones a la API. Para tener acceso a la API es necesario crear una cuenta de desarrollador en la plataforma y obtener las claves de acceso (API keys) (Flickr, 2017b).

Figura 28. Proceso de descarga de fotografías geolocalizadas de Flickr



El proceso de recolección y descarga de los datos geolocalizados se inicia con la recolección de todos los identificadores únicos de las fotos existentes dentro de la extensión territorial delimitada espacial y temporalmente.

La extensión del área de esta investigación está determinada por los límites de los 15 parques nacionales que forman la red española de parques nacionales, y el intervalo temporal seleccionado fue entre 2010 y 2018.

En el segundo paso, los identificadores de las fotos se utilizan para recolectar los metadatos adicionales asociados a cada foto, es decir, la información de usuario.

Ambos pasos se ejecutan mediante un script de Python y los resultados se almacenan en una base de datos no relacional de la aplicación MongoDB con formato JSON.

Finalmente, en el tercer paso, se utilizaron las coordenadas XY de cada fotografía para crear capas de puntos en un SIG. Las coordenadas de las capas de puntos se encuentran referenciadas en el sistema de coordenadas geográficas WGS84 por lo que se transformaron al sistema de coordenadas proyectadas que corresponde a cada parque nacional (Tabla 10). Todas las capas creadas se almacenaron en una base de datos espacial de formato file geodatabase para su posterior análisis.

Tabla 10. Sistema de referencia espacial para cada parque de la red nacional

Parque	Sistema de coordenadas
Teide	UTM ETRS89 Huso 28 N. EPSG:25828
Timanfaya	
Caldera de Taburiente	
Garajonay	
Archipiélago de Cabrera	UTM ETRS89 Huso 31 N. EPSG:25831
Aigüestortes	
Ordesa y monte perdido	
Cabañeros	UTM ETRS89 Huso 30 N. EPSG:
Monfragüe	
Tablas de Daimiel	
Sierra de Guadarrama	
Picos de Europa	
Doñana	
Sierra Nevada	
Islas Atlánticas de Galicia	UTM ETRS89 Huso 29 N. EPSG:

La base de datos de Flickr almacena tanto las fotografías en formato original como los metadatos asociados a las mismas. Estos metadatos incluyen la fecha y hora en que se tomó y subió la foto, un

título, la localización de la residencia del fotógrafo, una descripción y una amplia variedad de etiquetas textuales. La **Tabla 11** ofrece una descripción general y una explicación de los atributos de los metadatos relevantes para esta investigación.

Tabla 11. Descripción de los atributos de los metadatos de las fotografías de Flickr

Atributo	Descripción
<i>id</i>	Identificador de la foto
<i>owner nsid</i>	Identificador del fotógrafo
<i>owner location</i>	lugar de residencia del fotógrafo (ciudad o país)
<i>location</i>	Lugar donde se tomó la foto
<i>title</i>	Título de la foto
<i>dates taken</i>	Fecha y hora de toma de la foto
<i>posted date</i>	Fecha y hora cuando se subió la foto a Flickr
<i>description</i>	Descripción de la fotografía
<i>views</i>	Número de veces que la fotografía ha sido visualizada por usuarios de Flickr
<i>tags</i>	Conjunto de etiquetas que describen el contenido de la foto
<i>latitude</i>	Coordenadas de latitud de la foto (EPSG:4326)
<i>longitude</i>	Coordenadas de longitud de la foto (EPSG:4326)
<i>accuracy</i>	Precisión de la localización de la foto
<i>url</i>	Dirección específica para acceso a la foto

Wikiloc.

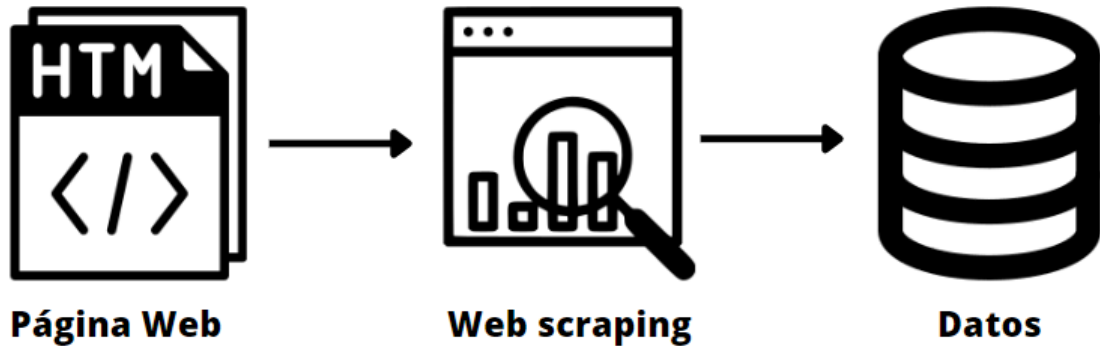
Descarga por Web scraping.

El web scraping es una técnica que consiste en recuperar el contenido de páginas web a través de software o scripts de forma automática (Batrinca & Treleaven, 2014). Esta técnica ha sido utilizada en Wikiloc, ya que la red social no cuenta con una API que permita el acceso y descarga de las rutas que almacena.

El web scraping se lleva a cabo en dos pasos secuenciales: adquirir los recursos web y extraer la información deseada de los recursos adquiridos. Ambos pasos se pueden crear y organizar mediante aplicaciones software como entornos de desarrollo integrados (IDEs) y utilizando algún lenguaje de programación. El proceso se inicia con la redacción de una solicitud HTTP para recursos de un sitio web objetivo. Esta petición se puede redactar en una URL que contenga un método de consulta GET. Luego se envía la petición al sitio web de destino. Una vez recibida la petición, el sitio

web procesa y devuelve la información solicitada. Los datos que se pueden extraer de las páginas web pueden ser de variados formatos por ejemplo XML o JSON, o datos multimedia como imágenes, archivos de audio o video (Zhao, 2020).

Figura 29. Proceso de web scraping (diagrama)

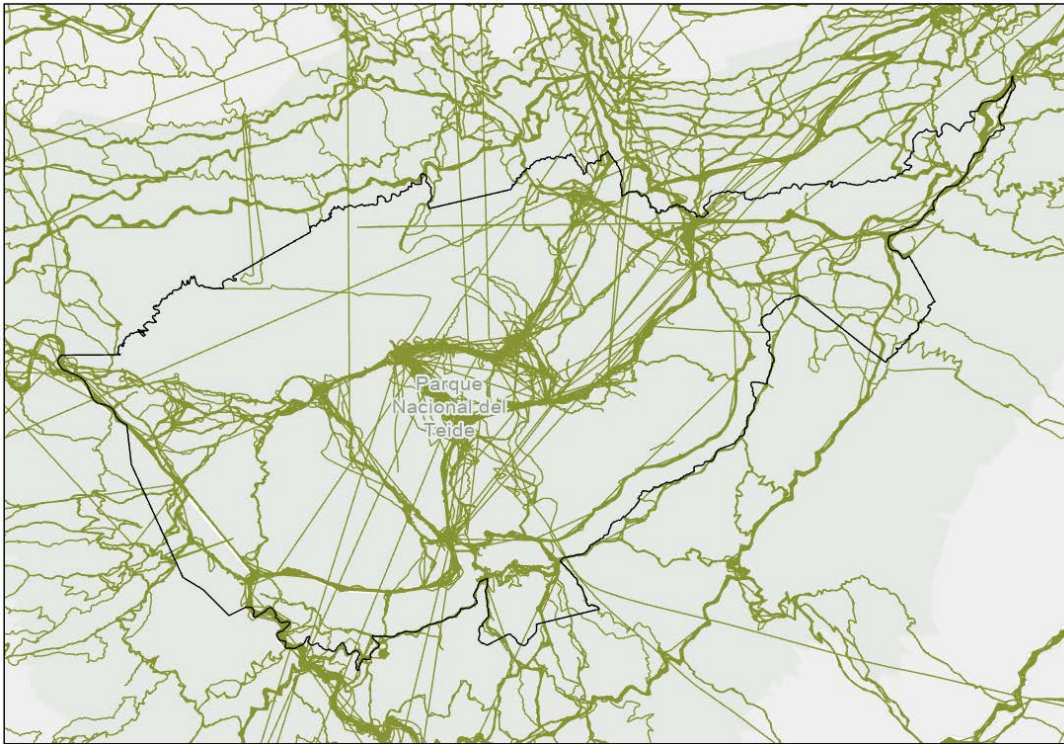


En una primera instancia se utilizó el método de web scraping para descargar todas las rutas existentes del parque nacional Teide, utilizando un script de Python. El resultado de la descarga consiste en un archivo JSON, que contiene las rutas GPS que luego se transformaron en una capa vectorial de líneas. Asociada a la información geográfica, las rutas tienen atributos que se describen en la siguiente tabla:

Tabla 12. Descripción de los atributos de las rutas de Wikiloc descargadas por web scraping.

Atributo	Descripción
<i>user_name</i>	nombre del usuario
<i>user_id</i>	identificador del usuario
<i>name</i>	nombre de la ruta
<i>date created</i>	fecha de creación de la ruta
<i>date published</i>	fecha de publicación de la ruta
<i>declared time</i>	tiempo de duración de la ruta declarado por el usuario
<i>type</i>	tipo de actividad de la ruta
<i>description</i>	descripción de la ruta
<i>trail rank</i>	clasificación de la ruta con relación a otros usuarios
<i>route url</i>	dirección url de enlace a la ruta
<i>user_url</i>	dirección url de la página de perfil del usuario

Figura 28. Rutas de Wikiloc del parque nacional Teide descargadas por web scraping



Descarga desde la página web.

Debido a las limitaciones y restricciones para poder extraer datos de forma automatizada a través de web scraping y para analizar el potencial de los datos GPS de redes sociales, se obtuvieron datos de 6 parques nacionales por medio de la descarga directa desde la página web de Wikiloc.

Se escogió trabajar con los parques que tenían menos rutas al momento de la descarga (2019), ya que Wikiloc aplica restricciones de descarga que limitan el número de rutas que se pueden descargar por día (250 rutas / día). De esta manera se optó por parques menos populares, para contar con la mayoría de las rutas disponibles al momento de la descarga y en la medida de lo posible se seleccionaron parques que representan distintos ecosistemas y paisajes. Los 6 parques escogidos son los siguientes:

- **Archipiélago de Cabrera:** islas
- **Islas Atlánticas de Galicia:** islas
- **Tablas de Daimiel:** humedal
- **Monfragüe:** bosque mediterráneo
- **Caldera de Taburiente:** volcánico
- **Aigüestortes:** montaña

Para realizar la descarga de los datos, se tuvieron en cuenta los siguientes criterios:

1. Se utilizó el buscador de rutas para introducir el nombre completo del parque nacional.

2. Solo se descargaron las rutas que caían dentro de los límites de cada parque nacional.
3. Debido a la limitación del número de descargas (máximo 200 rutas diarias), se seleccionaron únicamente 2 actividades; senderismo y ciclismo ya que son las más populares en Wikiloc.
4. Solo se descargaron las rutas subidas desde un dispositivo GPS o teléfono inteligente, por lo tanto se omitieron las rutas dibujadas en mapa.
5. Se descargaron las rutas correspondientes al mismo periodo temporal de un año (2018).

El buscador de rutas de la página de Wikiloc se utilizó para realizar la búsqueda de rutas ingresando el nombre completo del parque. Además, la página web permite establecer filtros para refinar la búsqueda. De tal manera se establecieron los siguientes filtros:

1. Las rutas circulares o de ida y vuelta fueron excluidas ya que al momento de ajustar las rutas a la red de senderos mediante map matching, el algoritmo de reconstrucción de rutas requiere que el origen y destino de la ruta sean diferentes de lo contrario no calcula ninguna ruta.
2. Las rutas dibujadas a mano sobre un mapa también se excluyeron, ya que las mismas no han sido generadas por dispositivos GPS, por lo tanto, no tienen información temporal (timestamp) que es necesaria para calcular la velocidad, distancia y duración de los recorridos.

El resultado de la descarga de rutas GPS consiste en archivos GPX, que son el formato estándar para el almacenamiento de datos de dispositivos GPS. Al igual que las rutas obtenidas por web scraping, las rutas descargadas contienen información adicional almacenada en los atributos que se describen en la Tabla 13.

Tabla 13. Descripción de los atributos de las rutas de Wikiloc descargadas de la página web

Atributo	Descripción
<i>Name</i>	nombre de la ruta
<i>Descript</i>	Descripción de la ruta
<i>Type</i>	Tipo del punto (waypoint o trackpoint)
<i>DateTimeS</i>	Fecha y hora de creación de la ruta en formato texto (string)
<i>Elevation</i>	Altitud del punto
<i>DateTime</i>	Fecha y hora de creación de la ruta en formato fecha (date)
<i>RouteID</i>	Id de la ruta

5.4. Preprocesamiento y limpieza de los datos geolocalizados de redes sociales.

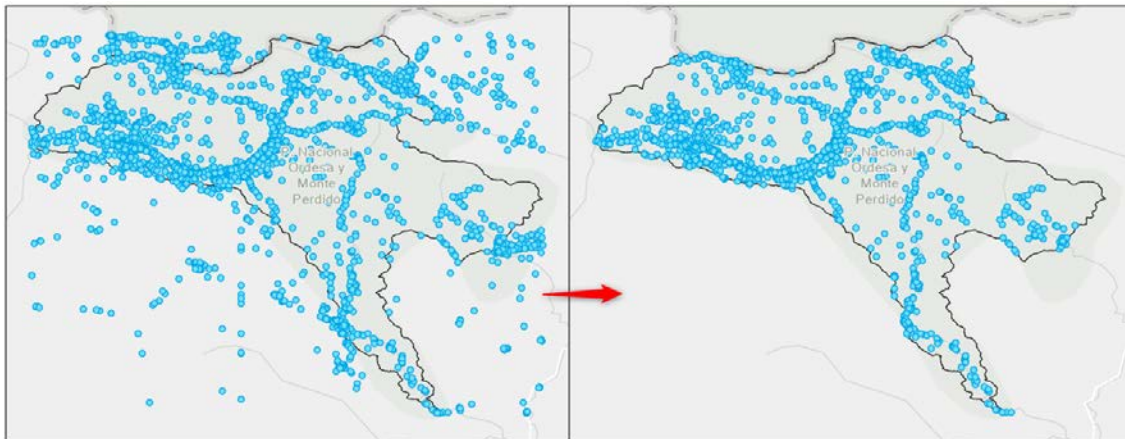
Un paso fundamental para trabajar con cualquier tipo de datos consiste en el tratamiento de los datos para filtrar y eliminar las inconsistencias y errores que tienen los datos crudos (Rahm & Do, 2000). La exploración inicial de los datos recolectados, reveló la existencia de inconsistencias en los conjuntos de datos de las dos fuentes. Por lo tanto, se desarrollaron una serie de pasos enfocados en contar y eliminar los errores de cada fuente de datos.

5.4.1. Limpieza y preparación de datos de Flickr.

Para depurar las fotografías geolocalizadas de Flickr se aplicaron los siguientes procesos:

- Identificar los registros de fotografías que fueron tomadas dentro de cada parque nacional:
Las fotografías geolocalizadas descargas corresponden con a todas las fotos que se encuentran dentro del recuadro mínimo que se utilizó como parámetro en las peticiones de descarga de la API. De esta forma el primer proceso de depuración consistió en extraer solo aquellas fotografías que se encontraban dentro de los límites administrativos de cada parque.

Figura 30. Ejemplo de recorte de fotografías geolocalizadas por límite administrativo.



- Identificar y eliminar los registros de fotografías con fecha mal asignada: Al realizar una exploración de la tabla de atributos de las fotografías geolocalizadas se observó la existencia de fotos que tenían una fecha incorrecta, por ejemplo, en algunos casos, las fechas de las fotografías correspondían con años anteriores a la creación de Flickr o posteriores al tiempo actual, por lo que se eliminaron dichos registros de la base de datos.
- Identificar y eliminar los registros de fotografías con información de fecha incompleta: Durante la exploración inicial de los datos, también se encontraron registros que no

contaban con información de hora de la toma de la fotografía en el atributo “date taken”, por lo que se eliminaron estos registros.

- *Agregar los datos de puntos a superficies (hexágonos)*: los datos de Flickr corresponden a las localizaciones de las fotografías realizados por los usuarios, por lo que es necesario agregar los datos en áreas regulares y contar el número de usuarios que se encuentran en cada área para obtener un dato de la presencia de usuarios mas no de fotos en los parques nacionales. La agregación de datos puntuales a una red de hexágonos regular se utilizó por varias razones: Primero, el objetivo de esta investigación es identificar la huella digital turística, por lo que el enfoque del análisis debe ser el usuario que crea el contenido geolocalizado. Por tanto, es necesaria una agregación espacial para contar para cada usuario en un determinado espacio y tiempo. La segunda razón está relacionada con la reducción del sesgo que producen los usuarios muy activos cuya actividad podría distorsionar los resultados. Finalmente, la agregación de datos estandariza el mapeo de los diferentes parques y ayuda a comparar los resultados entre ellos. Este proceso se describe con mayor detalle en el siguiente apartado.

5.4.2. Agregación espacial de las fotografías geolocalizadas.

La agregación de datos espaciales implica una serie de pasos a seguir. Partiendo de la selección del tamaño de la rejilla hexagonal, hasta la creación de la propia rejilla y las operaciones realizadas para resumir los datos de los puntos para unirla con los hexágonos y cuyo proceso se ilustra en la Figura 31. El tamaño óptimo de hexágono para cada parque se determinó siguiendo el enfoque propuesto por (Lee et al., 2016).

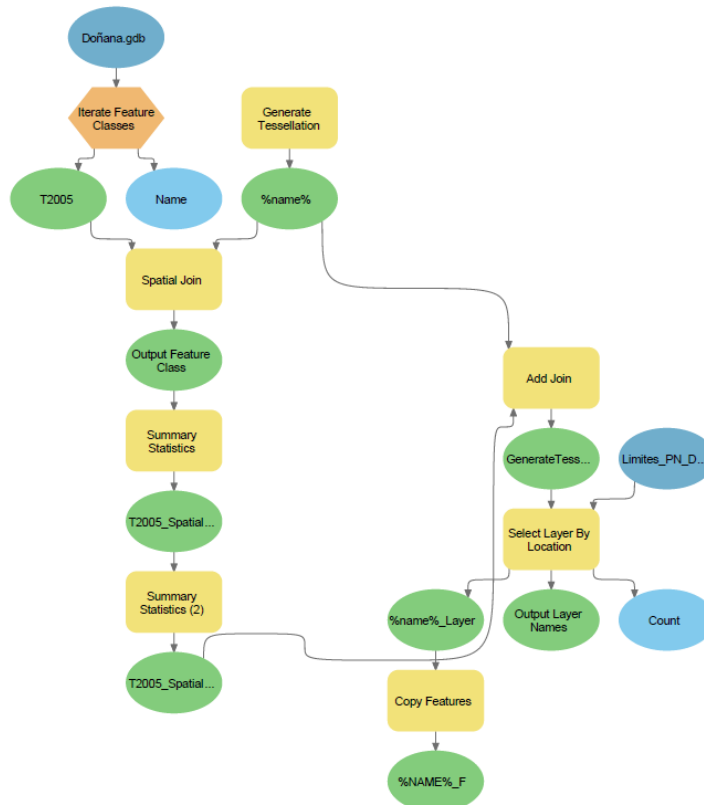
Este método utiliza el índice de Moran para evaluar qué tamaño de hexágono tiene la mayor autocorrelación según los diferentes niveles de zoom. Por tanto, los datos agregados mantienen las propiedades estadísticas de los datos puntuales y minimizan el efecto de problema de unidad de área modificable (MAUP).

Se calculó el índice de Moran para hexágonos de 100, 200, 300 y 400 metros de lado en cada parque. El tamaño óptimo del hexágono es diferente dependiendo del parque debido a la extensión de cada parque y al volumen de datos geolocalizados disponibles para cada espacio. Para los parques más pequeños, el valor de tamaño óptimo se fijó en un hexágono con lado lateral de 100 m. Los parques más grandes mostraron un tamaño óptimo de agregación entre 300 y 400 m de lado. Mientras que los parques de tamaño mediano arrojaron un valor óptimo de agregación para hexágonos con un lado de 200 m.

Tabla 14. Tamaño óptimo de lado del hexágono de agregación por parque nacional

Parque	Tamaño óptimo	Índice de Moran	Puntuación Z	Valor p
Archipiélago de Cabrera	200	0,447	23,098	0,000
Aigüestortes	200	0,418	29,221	0,000
Cabañeros	400	0,245	14,027	0,000
Caldera de Taburiente	100	0,341	10,552	0,000
Doñana	300	0,475	45,442	0,000
Garajonay	100	0,269	21,794	0,000
Islas Atlánticas de Galicia	100	0,578	21,024	0,000
Monfragüe	200	0,333	27,979	0,000
Ordesa y Monte Perdido	300	0,425	20,785	0,000
Picos de Europa	300	0,305	51,855	0,000
Sierra de Guadarrama	300	0,359	26,009	0,000
Sierra Nevada	400	0,408	35,660	0,000
Tablas de Daimiel	100	0,533	33,755	0,000
Teide	200	0,464	24,757	0,000
Timanfaya	100	0,265	25,062	0,000

Figura 31. Ejemplo de modelo de agregación de fotografías geolocalizadas



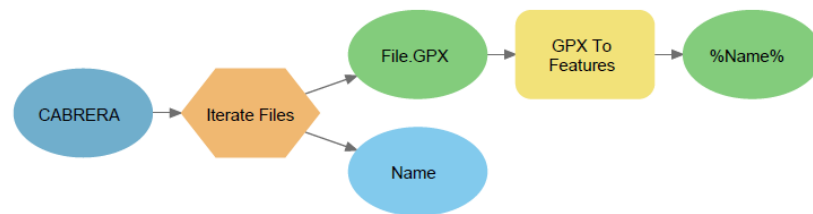
5.4.3. Limpieza y preproceso de los datos de Wikiloc.

Las rutas de senderismo y ciclismo descargadas desde Wikiloc consisten en archivos gpx, que contienen los puntos GPS con sus respectivas coordenadas (latitud y longitud) y la fecha y hora de toma de cada punto. Adicionalmente a esta información cada punto tiene asociado el nombre de la ruta y la altitud.

Para poder trabajar con las rutas en un entorno SIG, es necesario aplicar varios procesos para tener las rutas de cada parque en una sola de capa de puntos. Por lo tanto, se ha procesado los datos de acuerdo con lo siguiente:

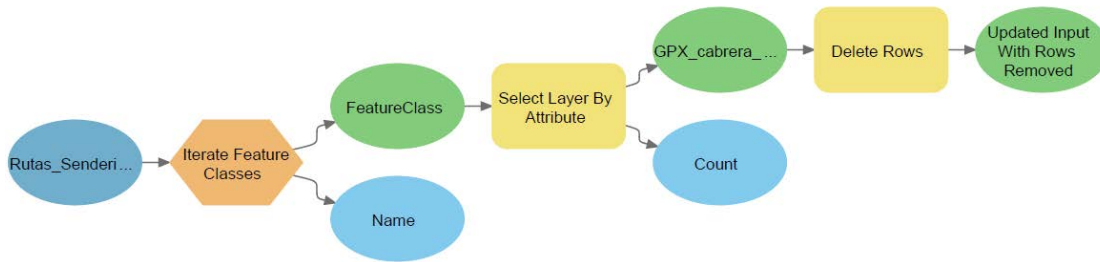
1. Conversión de GPX a capa de puntos: este proceso se realiza utilizando la herramienta de ArcGIS pro “gpx to feature class”, que transforma el archivo gpx en una capa de puntos basándose en las coordenadas almacenadas en el fichero.

Figura 32. Modelo de conversión de archivos GPX a formato SIG (ejemplo)



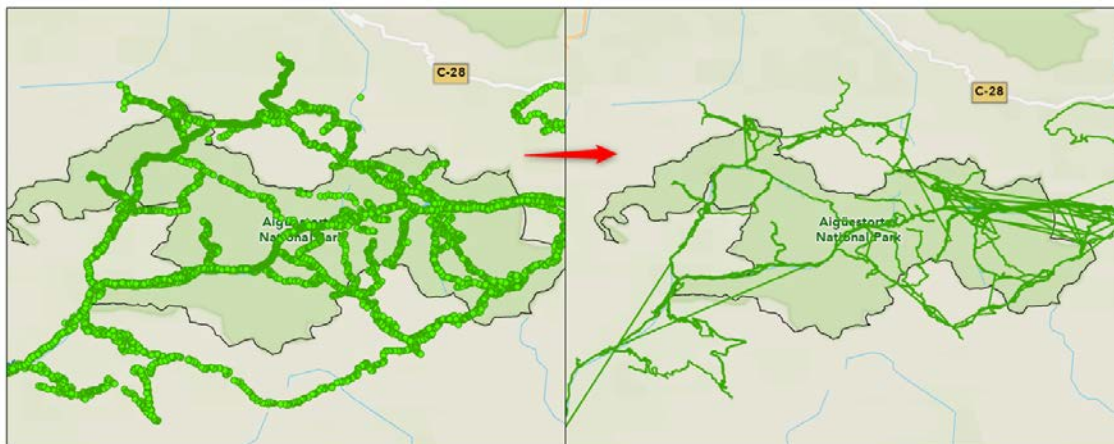
2. Cálculo de los parámetros de movimiento: usando el atributo de fecha y hora asociado a cada punto, se calcula la distancia, duración, velocidad y rumbo entre cada par de puntos GPS.
3. Detección y corrección de datos atípicos: Al realizar una exploración de los parámetros de movimiento se observa la existencia de inconsistencias como fechas incorrectas, velocidades y duración fuera del rango normal para las actividades de senderismo y ciclismo y que se producen debido a la falta de señal durante la recolección de los datos y la incertidumbre sobre la configuración de los dispositivos GPS. Por esta razón se realiza una limpieza de dichos datos usando los siguientes criterios:
 - a. Eliminar puntos que no cuentan con información de hora.
 - b. Eliminar puntos que tienen fechas fuera del rango de descarga (2018 al 2019)
 - c. Eliminar puntos en los que la velocidad calculada es superior a 5km/h para senderismo y 22 km/h para ciclismo.
 - d. Eliminar puntos en los que la duración calculada es superior a 3600 segundos.

Figura 33. Modelo para el filtrado de rutas GPS de acuerdo con criterio de duración



4. Proyección de las rutas de un sistema de coordenadas geográficas (WGS84) a un sistema de coordenadas planas.
5. Transformación de las capas de puntos en capas de líneas con las que se visualiza el recorrido exacto de cada usuario en los parques nacionales.

Figura 34. Ejemplo de transformación de puntos GPS a líneas en el parque nacional Aigüestortes



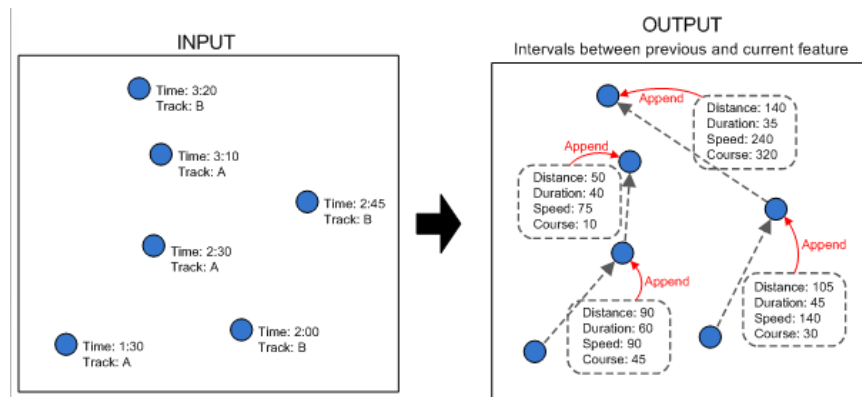
Parámetros de movimiento.

El “timestamp” o marca de tiempo de los puntos GPS recolectados de Wikiloc, registran la fecha y hora de cada punto que forman una ruta. Con esta información es posible calcular los parámetros que definen el movimiento de la ruta como la distancia, duración, velocidad y dirección en base a la diferencia entre dos puntos consecutivos. Para calcular estos parámetros, se utilizó la extensión tracking analyst de ArcMap 10.6. Esta herramienta toma cada par de puntos GPS consecutivos y calcula cuatro parámetros (velocidad, distancia, duración y dirección) para dicho par. La sumatoria de todos los pares de puntos, nos permiten conocer la duración, velocidad y distancia total de la ruta GPS. Asimismo, esta herramienta es particularmente útil para identificar errores ocasionados por la pérdida de señal de los satélites, por ejemplo, si dos observaciones consecutivas se

encuentran demasiado alejadas en términos espacial y temporal, los valores de los parámetros de movimiento resultantes serán extremos.

Una vez, que se calcularon los patrones de movimiento, se construyeron las rutas GPS, utilizando la herramienta Track intervals to line. Esta herramienta convierte los puntos en polilíneas.

Figura 35. Ilustración gráfica del cálculo de los parámetros de movimiento en ArcMap



Fuente: (ESRI, 2020).

5.5. Primera aproximación: exploración de los datos descargados.

5.5.1. Las fotografías geolocalizadas de Flickr de la red de parques nacionales.

La descarga de fotografías geolocalizadas de la red nacional de parques nacionales españoles produjo un total de 96570 fotografías entre 2010 y 2018. De este conjunto de datos, se excluyeron 34828 registros en los tres procesos de depuración de datos. La **Tabla 15** presenta el detalle de los datos excluidos debido a los procesos de limpieza de datos. Y en la **Tabla 16** se presenta un resumen general de los procesos de filtrado y el número de registros eliminados. Después de la limpieza de datos, el conjunto de datos se redujo a un total 61742 fotografías geolocalizadas distribuidas en los 15 parques nacionales y que fueron subidas por 16190 usuarios.

Tabla 15. Fotografías eliminadas durante los procesos de limpieza y depuración de datos.

Causa de depuración	# de fotografías	% del total
Recorte con el límite administrativo	31123	32%
Las fotografías no tienen una fecha correcta	2408	2%
Las fotografías no tienen información sobre la hora	1297	1%

Con respecto a cada parque se observan diferencias en el número de fotografías eliminadas durante los procesos de depuración, entre los que destacan el parque de Sierra Nevada, donde se observa que el 85% de las fotografías descargadas se excluyeron durante el proceso de depuración.

Esto se debe a que en la mayoría de las fotografías descargadas se encuentran en los núcleos urbanos alrededor del parque (Capileira, Bubión, Durcal, etc).

Tabla 16. Detalle de filtrado de fotografías geolocalizadas de Flickr por parque nacional

Parque	Datos iniciales (# fotos)	Limpieza 1*	Limpieza2**	Limpieza3** *
Aigüestortes	4466	2412	2356	2118
Archipiélago de Cabrera	780	663	653	647
Cabañeros	945	662	602	583
Caldera de Taburiente	3416	3197	3184	3177
Doñana	3556	3023	2987	2906
Garajonay	2445	2412	1965	1776
Islas Atlánticas de Galicia	3250	2635	2564	2466
Monfragüe	2073	2011	1997	1991
Ordesa y Monte Perdido	6966	5225	5036	4979
Picos de Europa	16249	12999	12678	12357
Sierra de Guadarrama	10061	6640	6389	6310
Sierra Nevada	18217	3643	2856	2767
Tablas de Daimiel	1072	997	986	979
Teide	18345	15670	15552	15482
Timanfaya	3729	3258	3234	3204

*Limpieza1: Recorte de fotos con límite

**Limpieza 2: Eliminar registros con fecha incorrecta

***Limpieza 3: Eliminar registros con fecha incompleta

El volumen de fotografías descargadas varía de acuerdo con el parque, siendo el parque nacional Teide el que reúne el mayor volumen de fotografías. Al contrario, el parque nacional Archipiélago de Cabrera fue el parque con el menor número de fotografías descargadas. La media de fotografías por usuario oscila entre 2 y 6 fotografías. El número de usuarios y fotos por km² nos puede indicar de forma general aquellos parques donde la presión turística es mayor. A modo de ejemplo, se observan parques con valores altos como Islas Atlánticas de Galicia, Ordesa, Picos de Europa y Teide donde el número de usuarios y fotografías supera la media (11 usuarios/km² y 35 fotos/km²).

En análisis posteriores, se explorará el atributo temporal de las fotos para analizar el comportamiento espacial y temporal de los visitantes e identificar con mayor fiabilidad cuáles se encuentran sometidos a una mayor presión.

Tabla 17. Detalle por parque de las fotografías geolocalizadas y usuarios de Flickr. (datos depurados)

Parques Nacionales	Fotos	Usuarios	Fotos/usuario	Fotos/km ²	Usuarios/km ²
Aigüestortes	2118	757	2.8	3.2	1.1
Archipiélago de Cabrera	647	156	4.1	4.1	1.0
Cabañeros	583	88	6.6	12.4	1.9
Caldera de Taburiente	3177	421	7.5	16.7	2.2
Doñana	2906	547	5.3	20.6	3.9
Garajonay	1776	473	3.8	3.3	0.9
Islas Atlánticas de Galicia	2466	1071	2.3	81.4	35.3
Monfragüe	1991	576	3.5	39.0	11.3
Ordesa y Monte Perdido	4979	1377	3.6	125.0	34.6
Picos de Europa	12357	3147	3.9	123.3	31.4
Sierra de Guadarrama	6310	1067	5.9	15.4	2.6
Sierra Nevada	2767	733	3.8	3.2	0.9
Tablas de Daimiel	979	364	2.7	11.5	4.3
Teide	15482	4250	3.6	84.2	23.1
Timanfaya	3204	1163	2.8	9.4	3.4
Total	61742	16190	3.8	16.1	4.2

Las figuras 36 y 37 presentan la distribución espacial de las fotografías descargadas de Flickr de los 15 parques nacionales españoles. Se ha clasificado a las fotografías en 4 colores que representan la estación en la fueron tomadas. A primera vista, se observa que el número de fotos presenta una gran variación en su distribución espacial entre cada uno de los parques nacionales, pudiendo observarse parques donde existe un gran número de fotos como los parques de Ordesa, Picos de Europa, Sierra de Guadarrama, Teide e Islas Atlánticas de Galicia. En el lado opuesto, se observan parques menos populares donde el volumen de fotos es evidentemente menor, por ejemplo, los parques de Cabañeros y Archipiélago de Cabrera. Además del volumen de fotos por espacio, estos mapas iniciales nos permiten una visualización aproximada de fenómenos espaciales de concentración y dispersión de los visitantes que dan una idea del uso del espacio por parte de los visitantes de la red.

Figura 36. Distribución espacial de las fotografías geolocalizadas de Flickr en la red de parques nacionales españoles (1 de 2)

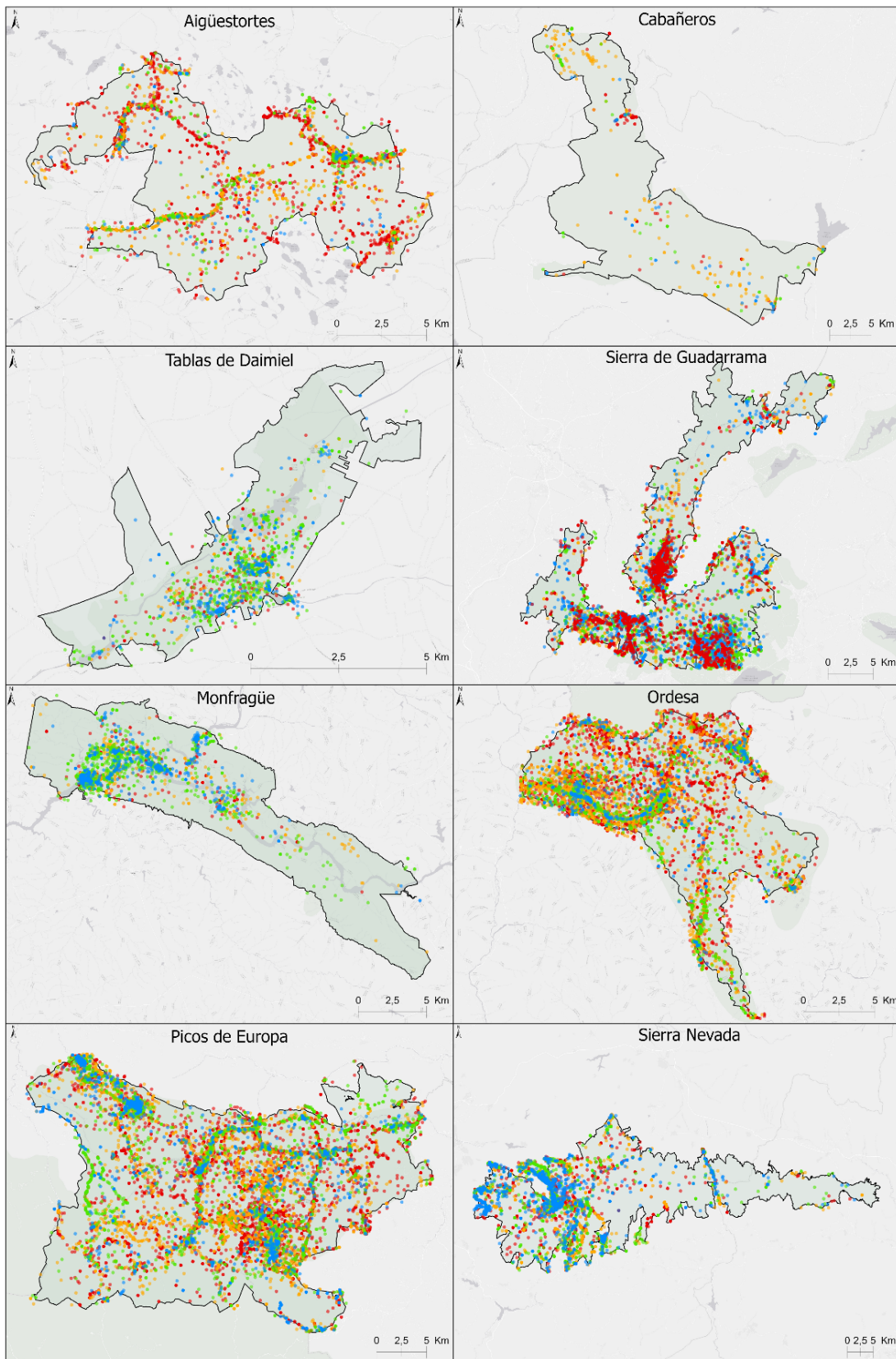
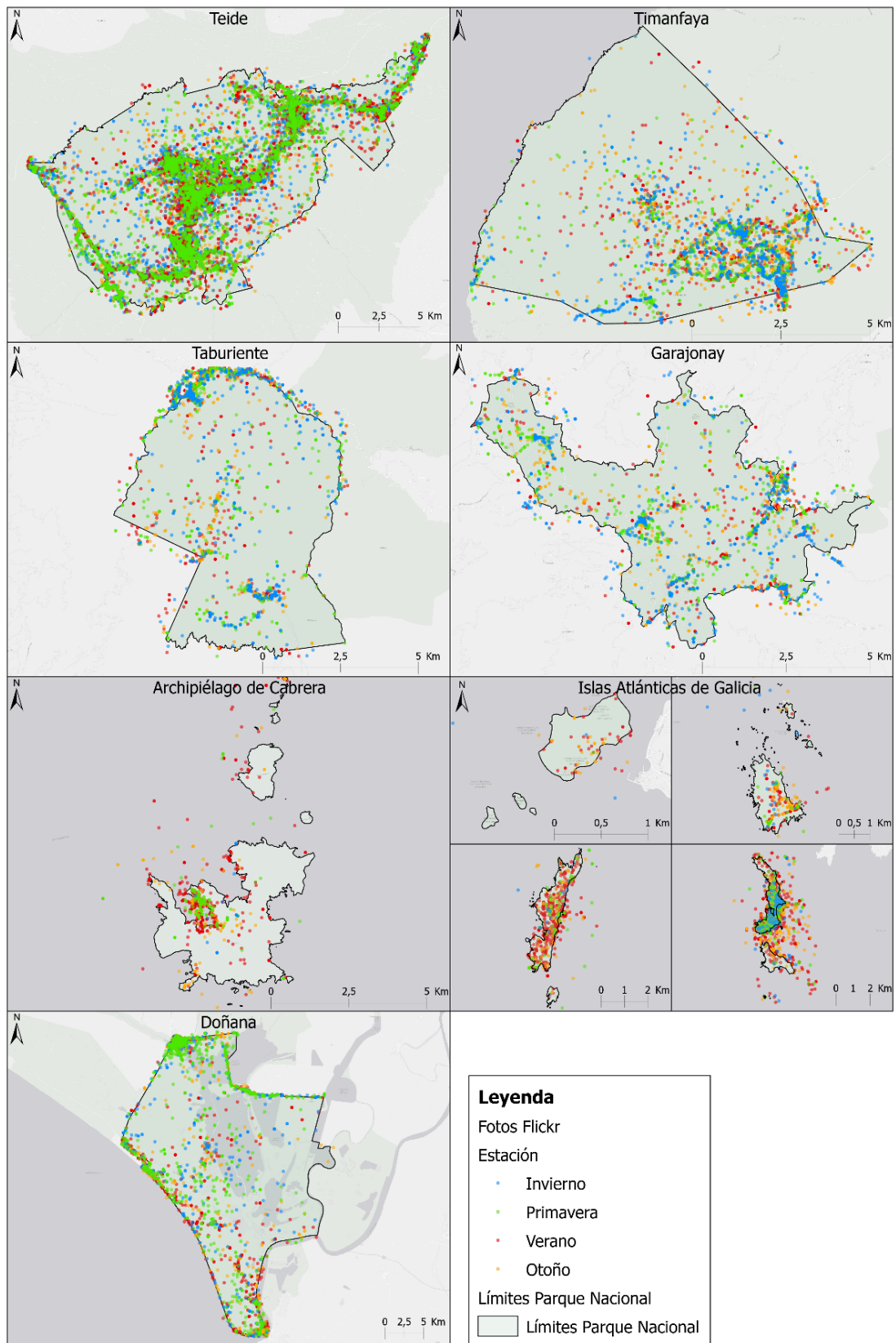


Figura 37. Distribución espacial de las fotografías geolocalizadas de Flickr en la red de parques nacionales españoles (2 de 2)



5.5.2. Las rutas GPS de Wikiloc de la red de parques nacionales.

La descarga de rutas GPS por medio de la página web de Wikiloc resultó en un total de 2741 rutas para el 2018 distribuidas en los 6 parques seleccionados para la descarga. De las rutas descargadas el 80% corresponde a rutas de la actividad de senderismo, en tanto que el 20% restante corresponde a las rutas por ciclismo. A nivel de parque, el P.N. Archipiélago de Cabrera no registró ninguna ruta de ciclismo al momento de la búsqueda en el portal de Wikiloc. El parque con el mayor número de rutas para el periodo de descarga (2018), fue el P.N. Islas Atlánticas de Galicia con 1065 rutas. El P.N. Archipiélago de Cabrera registró el menor número de rutas con 157 rutas. La Tabla 18 presenta el detalle por parque y actividad de las rutas descargadas.

Tabla 18. Rutas obtenidas de la descarga de Wikiloc por parque nacional

Parque	Senderismo	Ciclismo	Total
Archipiélago de Cabrera	157	--	157
Aigüestortes	478	250	728
Caldera de Taburiente	250	37	287
Islas Atlánticas de Galicia	1005	60	1065
Monfragüe	160	139	299
Tablas de Daimiel	169	36	205
Total	2219	522	2741

De este conjunto de datos, se excluyeron 195 rutas en un primer filtrado que consistió en eliminar las rutas que se encontraban fuera de los límites de los parques nacionales (Tabla 19).

Tabla 19. Detalle de primer proceso de depuración de rutas de Wikiloc

Parque	Datos iniciales	Limpieza 1*
Archipiélago de Cabrera	157	137
Aigüestortes	728	671
Caldera de Taburiente	287	247
Islas Atlánticas de Galicia	1065	1016
Monfragüe	299	282
Tablas de Daimiel	205	193
Total	2741	2546

*Filtrado de rutas fuera del rango de descarga

A partir del primer filtrado, se aplicaron filtros adicionales para limpiar los datos. Para ello, se calcularon los parámetros de movimiento para cada par de puntos GPS en las 2546 rutas. A partir de los valores de los parámetros de velocidad y duración se filtraron los datos de acuerdo con los siguientes criterios:

- Se eliminaron todos los puntos GPS que tenían un valor de velocidad superior a las velocidades promedio para las actividades de senderismo y ciclismo, 5km/h y 22km/h respectivamente.
- Se eliminaron todos los puntos GPS que registraron una duración superior a 3600 segundos

En la siguiente tabla, se detalla el proceso de filtrado de las rutas GPS.

Tabla 20. Procesos de limpieza y filtros aplicados a las rutas de Wikiloc

Parque	Datos iniciales	Limpieza 1*	Limpieza2**	Limpieza3***
Archipiélago de Cabrera	23999	23711	22592	22050
Aigüestortes	195152	192810	183710	179301
Caldera de Taburiente	63921	63154	60173	58729
Islas Atlánticas de Galicia	170883	168832	160864	157003
Monfragüe	64179	63409	60416	58966
Tablas de Daimiel	33758	33758	33758	33758

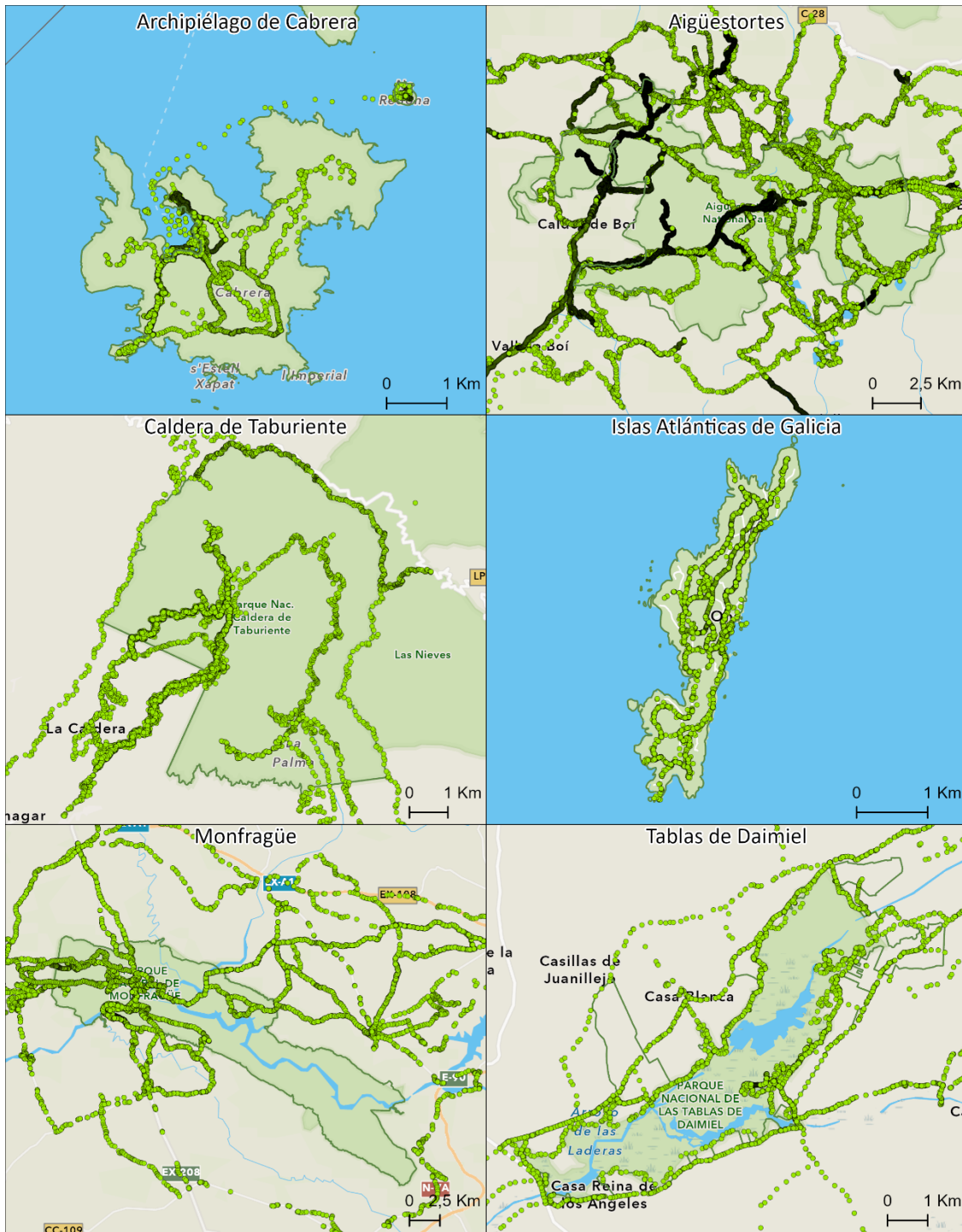
* Filtrado puntos sin información de hora

** Eliminación de puntos con velocidad superior a 5km/h (senderismo) y 22km/h (ciclismo)

*** Eliminación de puntos con duración superior a 3600 segundos

La Figura 38 presenta el despliegue y visualización de las rutas GPS, en forma de puntos, obtenidas de la descarga de Wikiloc para cada uno de los parques seleccionados. De forma general se observa que las rutas descargadas se ajustan a las redes de senderos de cada parque y que, en la mayoría de los parques, las rutas subidas por senderistas y ciclistas inician o terminan fuera de los límites administrativos de los parques. Al igual que con los datos provenientes de Flickr, las rutas de Wikiloc tienen un comportamiento similar, en el sentido de que se observan zonas de concentración de puntos, es decir con un mayor número de rutas atravesando esas localizaciones y que nos indican de forma inicial los senderos más utilizados. Destacan en el parque Archipiélago de Cabrera, la zona costera de “Sa Plageta” en el sector noroeste de la isla Cabrera y la isla Na Redona en el norte del Archipiélago. En el parque Aigüestortes se observan varias zonas de concentración de puntos GPS, entre las que sobresalen los miradores del Portarró y San Esperit que forman parte del itinerario de la ribera de Sant Nicolau. En el parque Caldera de Taburiente sobresalen las rutas que conducen a Roque de los Muchachos en el norte del parque. Con respecto al parque Islas Atlánticas de Galicia, las islas de Cíes y Ons reunieron el mayor número de rutas. En el parque Monfragüe, destacan los itinerarios de bicicleta Villareal -Serradilla ubicados en el sector occidental. En Tablas de Daimiel, se observa un mayor volumen de rutas GPS en los itinerarios hacia Isla del Pan, Isla del Descanso e Isla de Maturro que se localizan en el centro del parque.

Figura 38. Distribución espacial de las rutas GPS descargadas de Wikiloc



Las rutas de Wikiloc descargadas a través de web scraping pertenecen al parque nacional Teide y comprenden 5076 rutas de 24 actividades deportivas distintas realizadas entre 2006 y 2018.

Del total de rutas obtenidas, se excluyeron en un primer filtro 1776 rutas que se encontraban fuera de los límites administrativos del parque nacional Teide. En un segundo filtro se eliminaron todas las rutas que tenían una longitud inferior a 1 kilómetro, debido a que estas rutas correspondían con recortes de las rutas que se habían realizado fuera del parque. Finalmente, en un último proceso

de limpieza se excluyeron todas las rutas de actividades distintas al senderismo y ciclismo. El conjunto de datos final consiste en 2425 rutas, de las cuales el 90% corresponde a rutas de senderismo.

Tabla 21. Detalle de rutas por actividad y procesos de limpieza de rutas GPS del parque nacional Teide

P.N. Teide	Total	Senderismo	Ciclismo	Otras actividades
Datos iniciales	5076	3778	585	713
Filtro 1*	3300	2264	518	518
Filtro 2**	2899	2193	232	474
Filtro 3***	2425	2193	232	0

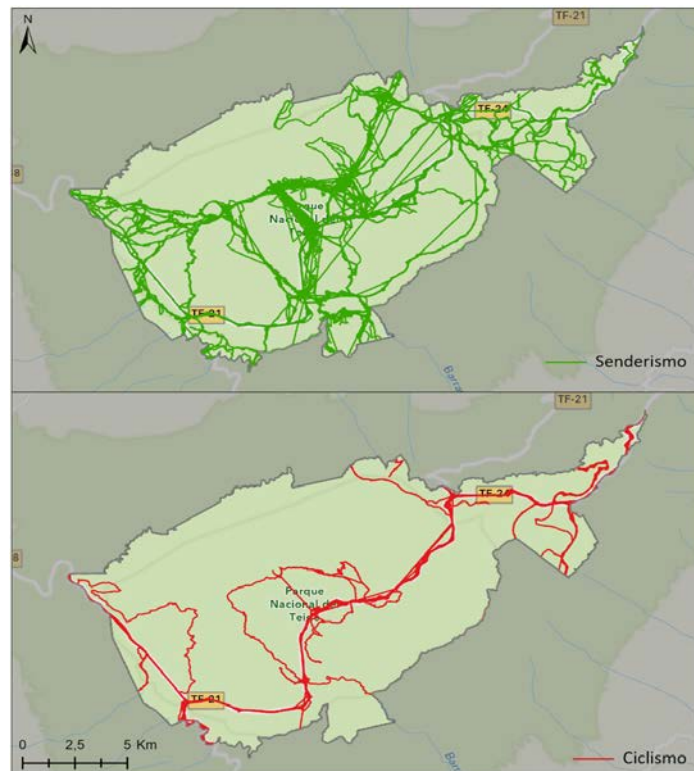
* Eliminar rutas fuera de los límites del parque

** Eliminar rutas con longitud menor de 1000 metros

*** Eliminar rutas con otras actividades

En la Figura 39 se presenta las rutas GPS de acuerdo con el tipo de actividad. Se pueden interpretar algunos patrones espaciales preliminares, por ejemplo, la concentración de rutas en los caminos que conducen al Pico del Teide para la actividad de Senderismo y en la carretera que atraviesa el parque para la actividad de ciclismo. De forma general se pueden identificar también los senderos más y menos frecuentados.

Figura 39. Distribución espacial de las rutas GPS de Wikiloc descargadas por web scraping de las actividades de senderismo y ciclismo



6. Lugar de origen de los visitantes de la red de parques nacionales españoles

6. Lugar de origen de los visitantes de la red de parques nacionales españoles.

Los datos geolocalizados de redes sociales permiten extraer el país de residencia de los turistas en base al origen de los usuarios. Este potencial se ha explorado en distintos tipos de turismo, como el urbano (Da Rugna et al., 2012), los sitios turísticos globalmente reconocidos (Bassolas et al., 2016) y los espacios naturales protegidos (Wood et al., 2013). En la literatura mencionada, el enfoque metodológico empleado se basa en la extracción de la información sobre localización del hogar o residencia del usuario, la misma que ha sido establecida por el usuario en su perfil. En esta sección se presentan los hallazgos obtenidos a partir de datos de usuarios de Flickr de la red de parques nacionales españoles.

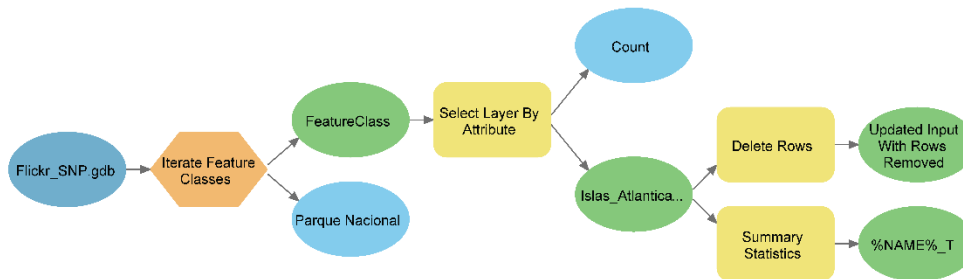
6.1. Clasificación de las nacionalidades de los usuarios.

El enfoque metodológico utilizado para extraer el país de origen de los visitantes de la red de parques nacionales hace uso del atributo “owner location”, que contiene la localización revelada por el usuario de Flickr en su perfil. Cuando un usuario de Flickr especifica la ubicación de su residencia, se adjunta como metadatos a todas sus fotos, por lo que es posible extraer este dato. Este atributo se define de manera opcional, por lo que no todos los usuarios del conjunto de datos han especificado su lugar de residencia. Para esta sección se utilizaron únicamente los datos de aquellos usuarios que había ingresado alguna información del nombre de su residencia en su perfil. De modo general, en la red de parques nacionales españoles, el 65% de los usuarios contaba con alguna información del lugar de residencia. A partir de estos datos, se aplicó un método semiautomático para la extracción y clasificación del país de residencia de cada usuario de Flickr en la red de parques nacionales españoles. Y en el caso de los usuarios españoles, se obtuvo la comunidad autónoma de residencia. La obtención de este tipo de información permite un análisis detallado del origen de los visitantes de la red de parques nacionales. El enfoque metodológico se estructura en torno a tres tareas:

1. Extracción de usuarios con información de país de residencia:

En primer lugar, se filtraron aquellos usuarios que habían ingresado información sobre su lugar de residencia usando la herramienta de selección por atributos en el campo “owner_location” y se calculó las estadísticas de resumen de la selección.

Figura 40. Modelo aplicado para la detección del país de origen de los visitantes de los parques nacionales



2. Clasificación de usuarios por país y por comunidad autónoma para los parques nacionales.

Con la base de datos filtrada, se agregaron los datos a nivel de usuario y lugar de residencia, usando la herramienta Resumir datos, de ArcGIS Pro. En un segundo paso, se clasifican por país, usando la herramienta de selección por atributo y utilizando una sentencia SQL diseñada con términos para cada país con el objetivo de agrupar a los usuarios por país. A estas selecciones se les asigna en un siguiente paso, el nombre del país correspondiente en un nuevo campo.

Figura 41. Ejemplo de sentencia SQL para selección y agrupación de usuarios por país (España)



Con los usuarios categorizados como españoles se aplica una nueva selección por atributos, utilizando una sentencia SQL elaborada para cada comunidad autónoma, y luego a cada selección se agregó el nombre de la comunidad a la que pertenecen en un nuevo campo.

Corrección manual

Luego de aplicar los dos procesos anteriores, todavía quedó un número pequeño (menos del 5%) de registros sin información del país de residencia. Esto se debe a que el atributo de "owner location" tenía algún error de tipo o ortográfico, que tuvieron que ser revisados y corregidos manualmente.

3. Validación

Se utilizó el coeficiente de correlación lineal de Pearson para validar la relación entre el lugar de origen de los usuarios de Flickr y la procedencia de los visitantes de los parques nacionales. Este coeficiente mide el grado de asociación estadística entre dos variables cuantitativas.

Por lo tanto, se calculó el coeficiente de correlación entre el origen de los usuarios de Flickr, clasificados como visitantes nacionales y visitantes extranjeros, y los datos oficiales del origen de los visitantes de los parques nacionales. Para obtener los datos oficiales del origen de los visitantes, se utilizaron los registros de la encuesta de ocupación en alojamientos de turismo rural 2010-2019 del Instituto Nacional de Estadística (INE, 2023). Dicha encuesta incluye el indicador "viajeros entrados por zonas turísticas", el cual registra el número de viajeros por mes y su procedencia (nacionales o extranjeros) en zonas turísticas rurales.

En el análisis, se consideraron ocho parques nacionales definidos en la encuesta como zonas turísticas: Sierra Nevada, Ordesa, Picos de Europa, Garajonay, Taburiente, Teide, Cabañeros y Aigüestortes. Estos parques nacionales fueron seleccionados debido a la disponibilidad de datos en la encuesta para el período analizado, ya que no se contaba con datos sobre el origen de los visitantes de los parques nacionales durante todo el periodo cubierto por los datos obtenidos de Flickr en las memorias de los parques.

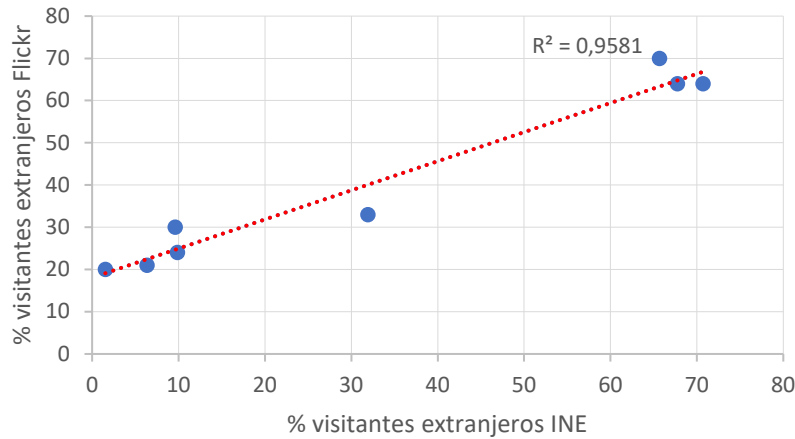
Al utilizar el coeficiente de correlación lineal de Pearson, se buscaba determinar si existía una relación significativa entre los datos de origen de los usuarios de Flickr y los datos oficiales del origen de los visitantes de los parques nacionales. Esta correlación permitiría validar el uso del lugar de origen de los usuarios de Flickr como un indicador confiable de la procedencia de los visitantes a los parques nacionales.

6.2. Validación del origen de los usuarios de Flickr como indicador del origen de los visitantes.

El origen de los visitantes, obtenido a través del atributo de ubicación de origen ("home location") de los usuarios de Flickr, se ha demostrado como un indicador confiable del origen de los visitantes de los parques nacionales. Los resultados del coeficiente de correlación de Pearson revelan una relación significativa y fuerte entre ambos indicadores ($r = 0.978$, valor $p = 0.000023$).

Estos resultados respaldan la validez y utilidad del atributo "home location" de los usuarios de Flickr para determinar la procedencia de los visitantes a los parques nacionales. La alta correlación positiva indica que los datos de origen de los usuarios de Flickr están estrechamente relacionados con los datos oficiales de origen de los visitantes de los parques nacionales. Por lo tanto, sugiere que el uso de los datos de Flickr puede proporcionar una estimación precisa y confiable del origen de los visitantes.

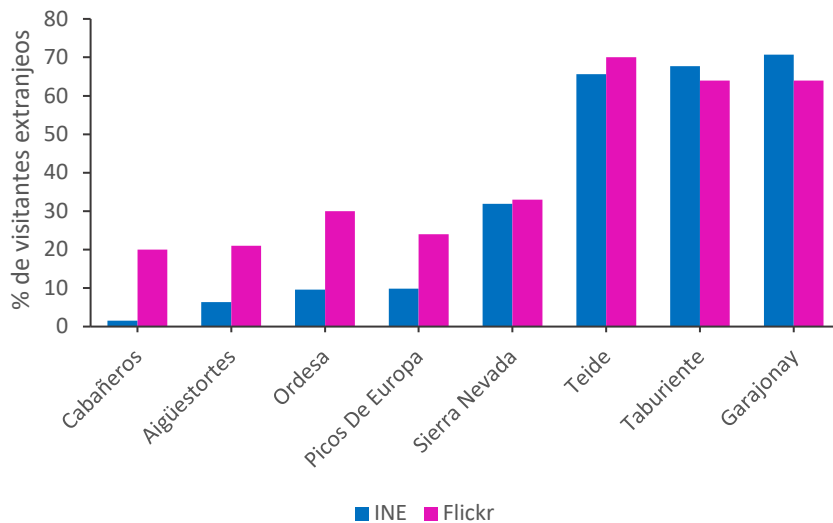
Figura 42. Gráfico de dispersión/comparación de las variables visitantes extranjeros Flickr con visitantes extranjeros de la encuesta del INE.



Al comparar los usuarios internacionales de Flickr con los visitantes extranjeros registrados en la encuesta de alojamientos rurales del INE, se observa una proporción similar entre ambos grupos. Sin embargo, se ha encontrado que el indicador de Flickr tiende a sobreestimar el número de visitantes internacionales en la mayoría de los parques nacionales (Figura 42).

Esta discrepancia puede deberse a que los visitantes extranjeros podrían estar más motivados a compartir y publicar fotos en comparación con los visitantes nacionales, ya que pueden considerar las vistas y experiencias más significativas o únicas. Es posible que los visitantes extranjeros tiendan a utilizar más activamente la plataforma de Flickr para compartir sus experiencias en los parques nacionales.

Figura 43. Comparación de % de visitantes extranjeros obtenido de Flickr y % de visitantes extranjeros de la encuesta de alojamientos rurales (INE).

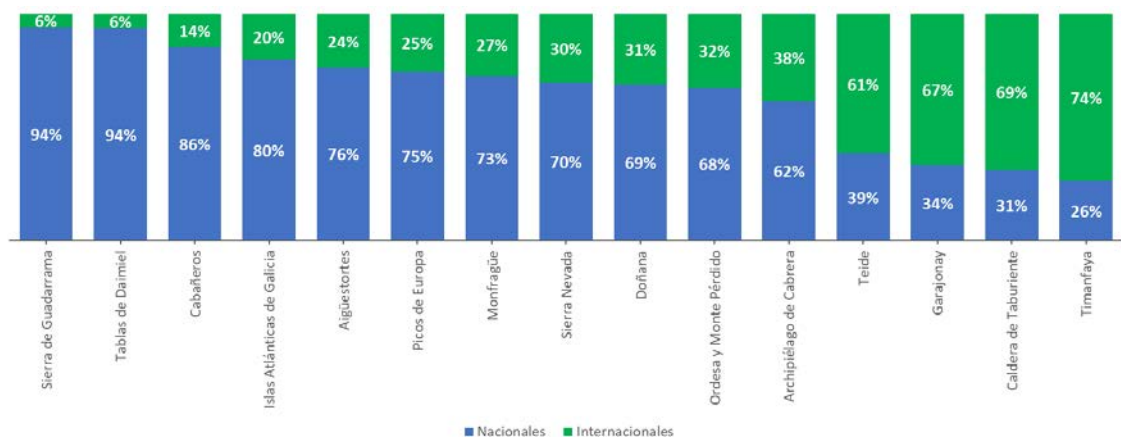


6.3. Distribución de los visitantes por origen.

La Figura 44 muestra la distribución de visitantes nacionales e internacionales en la red de parques nacionales españoles, obtenida a partir de las fotografías geolocalizadas de Flickr. En general, la demanda de visitantes nacionales es superior a la demanda de visitantes extranjeros en la mayoría de los parques, en particular en los parques peninsulares. La tendencia es inversa en los parques de las islas Canarias, donde los visitantes extranjeros representan entre el 61% y 74% de los visitantes totales.

En el caso de los parques peninsulares, destacan los parques de Tablas de Daimiel, Sierra de Guadarrama, Islas Atlánticas de Galicia y Cabañeros, como receptores de más del 80% de visitantes españoles. En el lado opuesto los parques que han recibido el menor porcentaje de visitantes españoles fueron los parques Caldera de Taburiente y Timanfaya.

Figura 44. Distribución de los visitantes por origen (Nacional / Internacional) de acuerdo con datos de los usuarios de Flickr 2010-2018



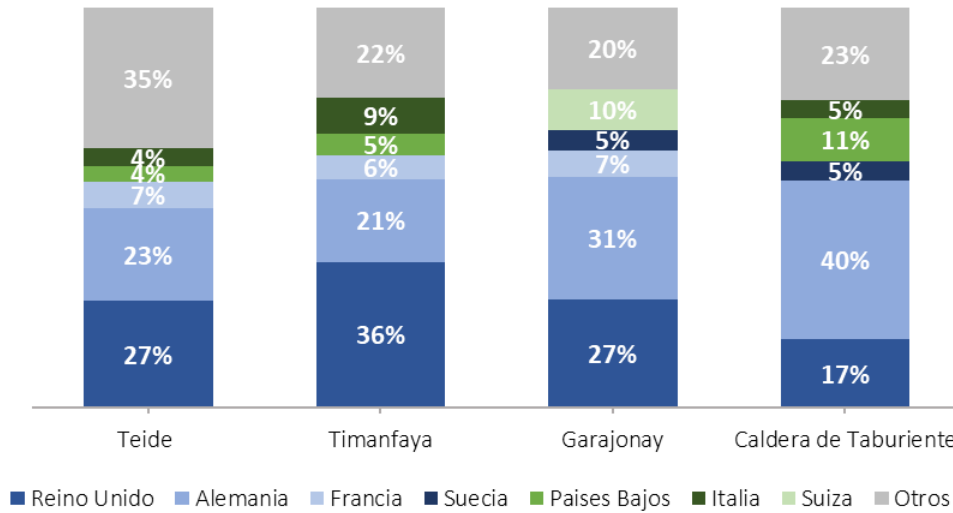
6.4. Países de origen de los visitantes extranjeros.

La distribución de visitantes extranjeros no es uniforme en la red de parques nacionales. Destaca que los principales lugares de origen de los visitantes son países de la Unión Europea. En general, Reino Unido y Alemania son los principales orígenes de los visitantes para la mayoría de los parques. Estos resultados, se encuentran en línea con los datos reportados en las estadísticas oficiales de turismo sobre el origen de los turistas extranjeros (INE, 2022). El parque que registró el mayor número de países de residencia fue el parque nacional Teide. En el lado opuesto, en el parque nacional Cabañeros se registraron solo tres países de origen.

En los parques canarios, los principales emisores de visitantes son Reino Unido y Alemania, que contribuyen entre el 50% y 58% de los visitantes internacionales (Figura 45). Otros países que

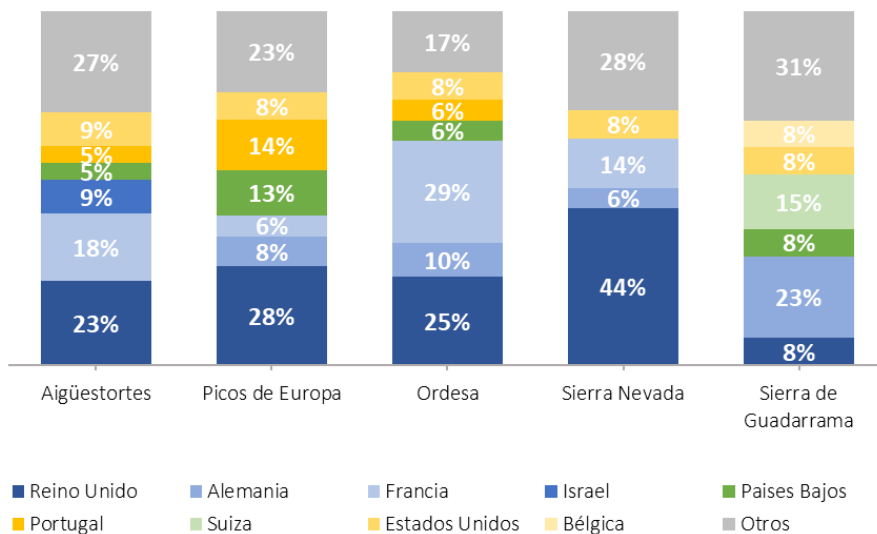
destacan por su contribución de visitantes son Francia, Italia, Suiza y Países Bajos. De forma individual, destaca el parque nacional Teide, el que se obtuvieron 27 países de residencia de los visitantes, de los cuales 20 reúnen menos del 2%.

Figura 45. Distribución por país de los visitantes internacionales en los parques nacionales canarios.



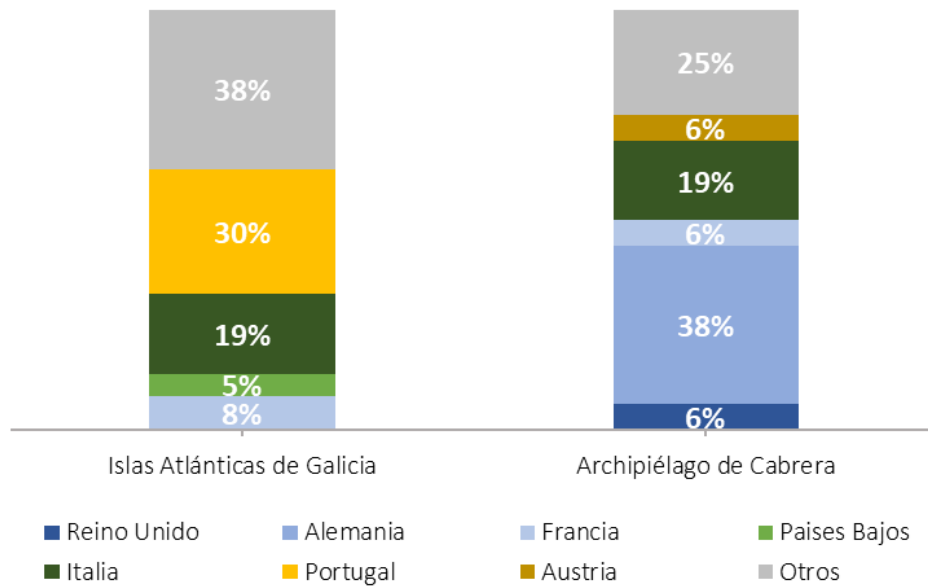
En los parques de montaña (Figura 46) se observa una tendencia similar a la de los parques canarios, siendo Reino Unido el principal emisor de visitantes para la mayoría de los parques a excepción de los parques de Ordesa y Sierra de Guadarrama. En el caso de Ordesa, la mayoría de los visitantes extranjeros provienen de Francia, condicionado por su localización en la frontera con este país. En el parque Sierra de Guadarrama, el mayor emisor de visitantes fue Alemania, que también es un importante emisor de turistas en España (INE, 2022).

Figura 46. Distribución por país de los visitantes internacionales en los parques nacionales de montaña.



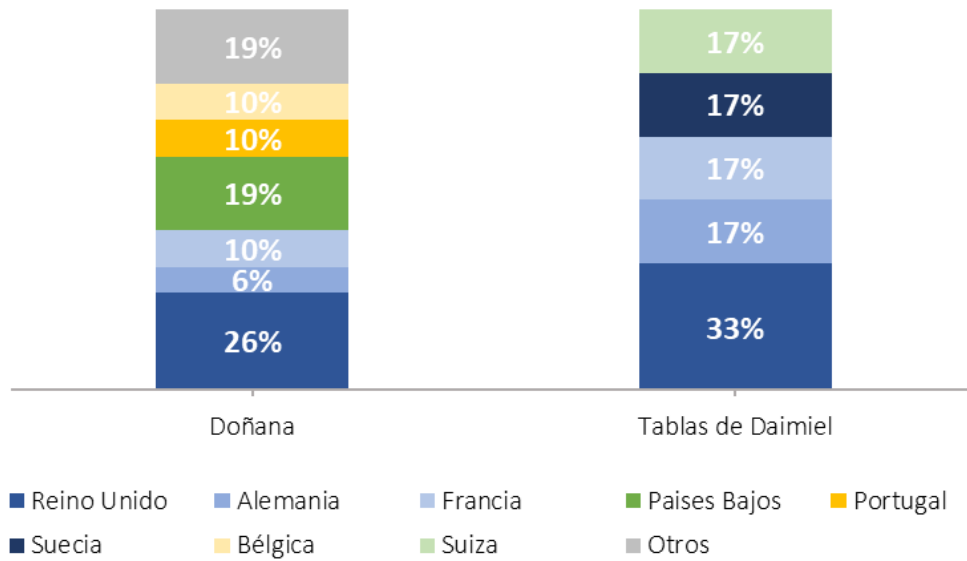
En Islas Atlánticas de Galicia, el principal emisor de visitantes es Portugal, que se relaciona con la cercanía del parque con este país y que presenta un comportamiento similar a los indicadores generales de turismo en la comunidad autónoma de Galicia (AEITG, 2019). En el Archipiélago de Cabrera, la mayoría de los visitantes vinieron de Alemania. En ambos parques destaca el elevado porcentaje de la categoría otros países que reúne a los visitantes de países que registraron un porcentaje menor al 5% y que revela una gran diversidad de visitantes en cuanto a su origen.

Figura 47. Distribución por país de los visitantes internacionales en los parques nacionales insulares



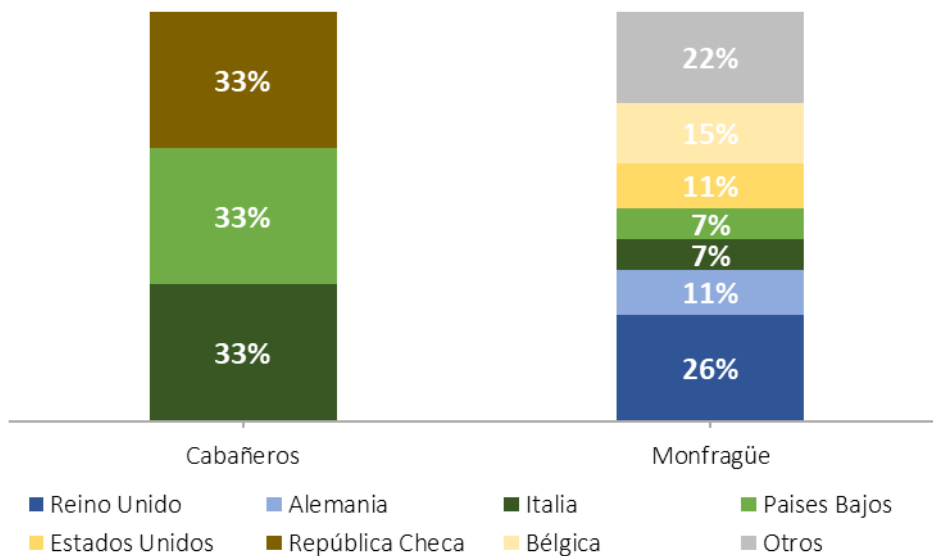
En los parques nacionales de Doñana y Tablas de Daimiel, la mayoría de los visitantes extranjeros fueron de Reino Unido, correspondiendo al 26% y 33% en Doñana y Tablas de Daimiel respectivamente. En Doñana se puede observar que los visitantes extranjeros provienen de un mayor número de países (12 países), en tanto que en el P.N. Tablas de Daimiel los visitantes reportados provienen de 5 países.

Figura 48. Distribución por país de los visitantes internacionales en los parques nacionales de humedales



En Cabañeros se registraron visitantes extranjeros provenientes de tres países distintos, sin que ninguno presente predominancia. Es importante mencionar que el P.N. Cabañeros tenía el menor volumen de datos de Flickr, por lo que los resultados son poco reseñables. Con respecto al P.N. Monfragüe, la mayoría de los visitantes extranjeros provienen de Reino Unido.

Figura 49. Distribución por país de los visitantes internacionales en los parques nacionales de bosque-paisaje mediterráneo

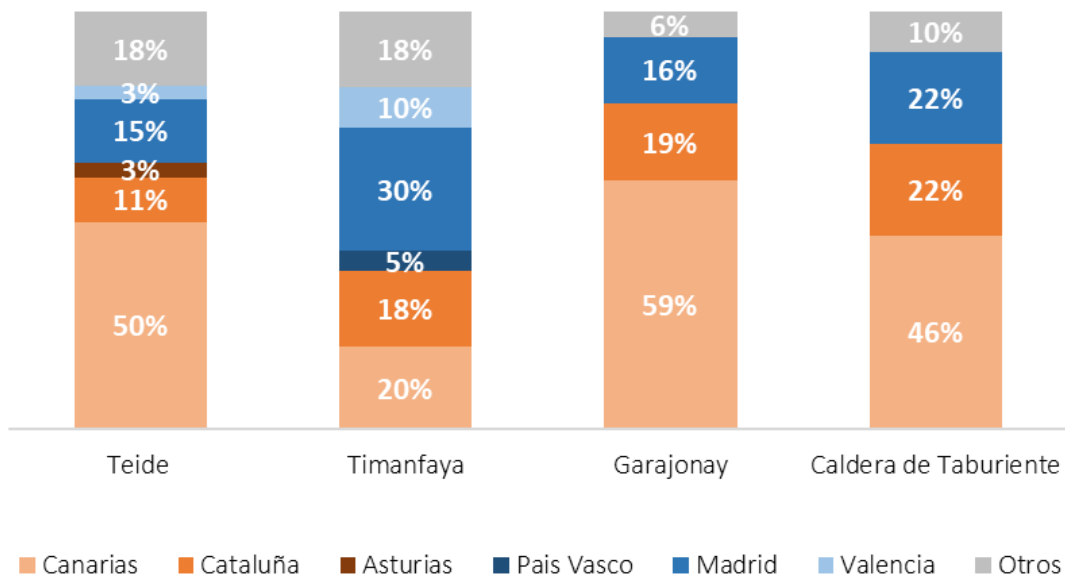


6.5. Comunidad autónoma de origen en los visitantes nacionales.

El análisis detallado de los visitantes españoles nos permite desagregar su origen a nivel de comunidad autónoma. De forma general, en todos los parques nacionales se registraron visitantes de las comunidades de Madrid y Cataluña. En tanto que, como es lógico, las comunidades de Ceuta y Melilla fueron las de menos presencia en los datos.

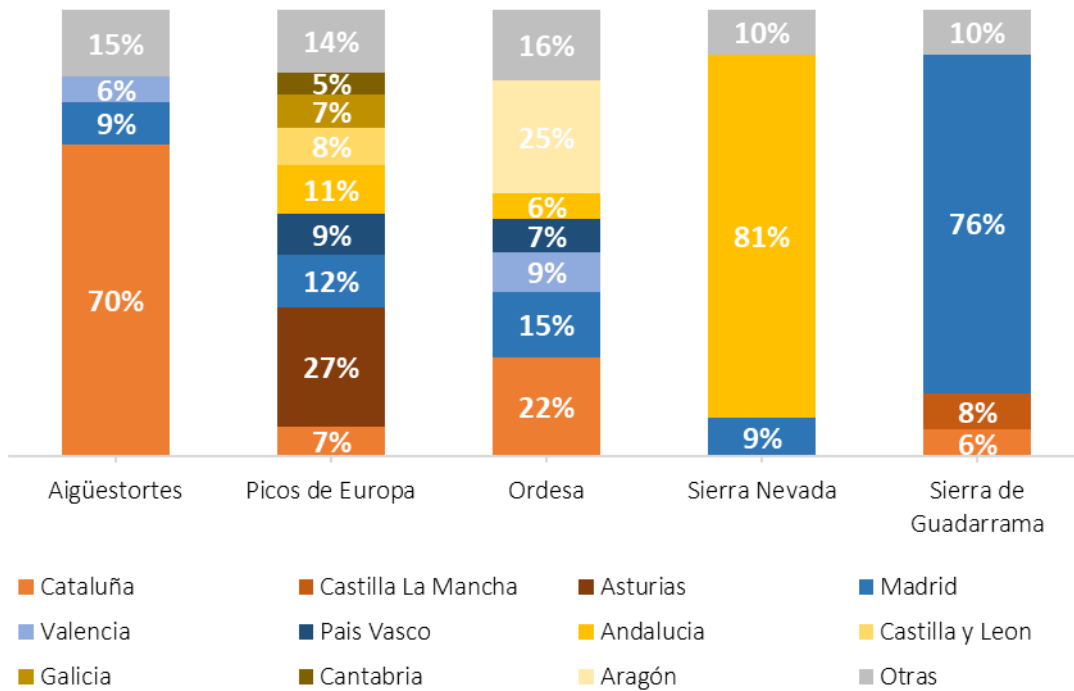
En los parques canarios, como se esperaba, la mayor parte de visitantes españoles provienen de las Islas Canarias, a excepción del parque nacional Timanfaya, en el cual la mayoría de los visitantes españoles provienen de la Comunidad de Madrid. Los visitantes de las comunidades de Cataluña y Madrid representaron una parte importante de los visitantes en todos los parques canarios.

Figura 50. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales canarios.



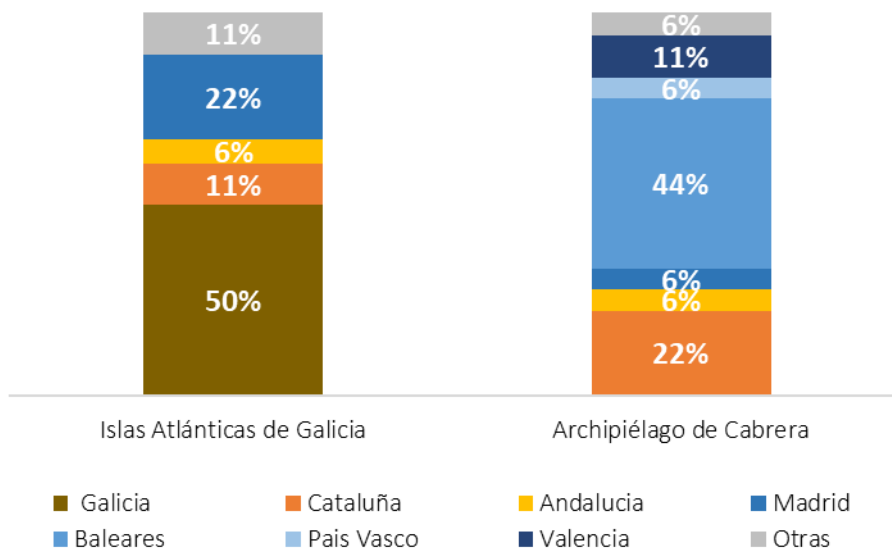
En los parques de Montaña, la comunidad autónoma con mayor porcentaje de visitantes coincide con la comunidad en la que se encuentra cada parque, por ejemplo, en el parque Aigüestortes, la mayoría de los visitantes españoles proceden de Cataluña. Destaca el parque Picos de Europa por registrar visitantes de siete comunidades autónomas distintas.

Figura 51. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales de montaña.



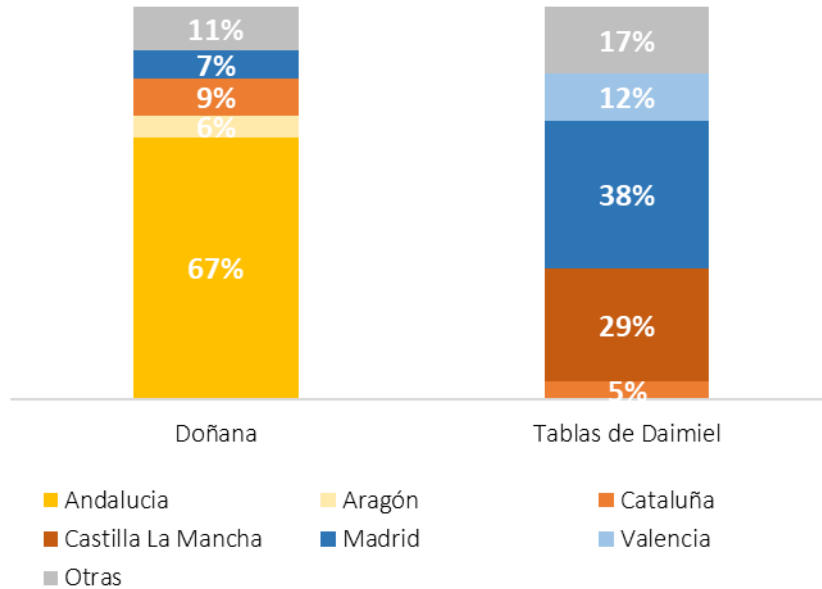
Los parques insulares muestran un comportamiento similar al de los parques de montaña, en el cuál la mayoría de los visitantes españoles provienen de la comunidad autónoma donde se localiza el parque, por ejemplo, en el parque nacional Islas Atlánticas de Galicia, el 50% registró la comunidad de Galicia como su lugar de origen y en el parque Archipiélago de Cabrera, el 44% de los visitantes proviene de las Islas Baleares.

Figura 52. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales insulares.



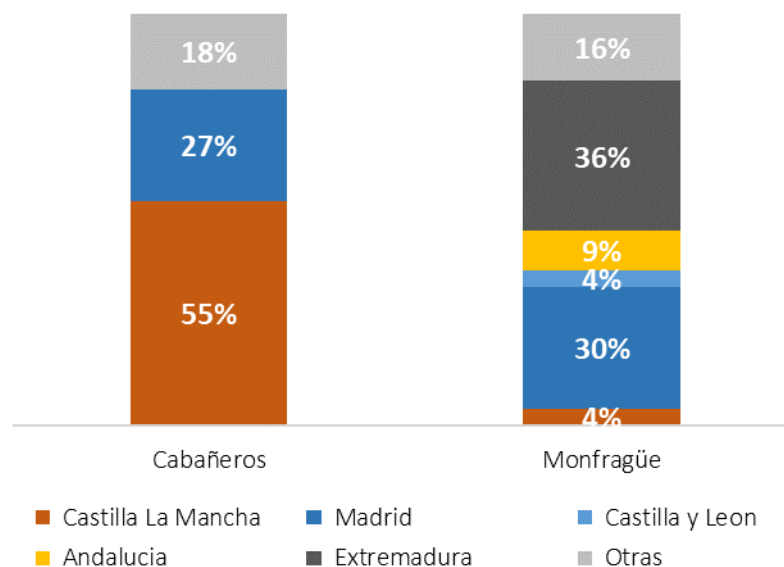
En el parque de Doñana, la mayoría de los visitantes españoles provienen de Andalucía y representaron el 67% del total. En Tablas de Daimiel, la mayoría de los visitantes españoles vinieron de la comunidad de Madrid.

Figura 53. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales de humedales



En los parques de Cabañeros y Monfragüe la mayoría de los visitantes españoles provienen de las comunidades autónomas donde se localizan los parques. En segundo lugar, en ambos parques se observó la presencia importante de visitantes de la comunidad autónoma de Madrid.

Figura 54. Distribución por comunidad autónoma de los visitantes españoles en los parques nacionales de bosque-paisaje mediterráneo



7. El ritmo temporal de los visitantes de la red de parques nacionales españoles.

7. El ritmo temporal de los visitantes de la red de parques nacionales españoles.

Contar con datos que caracterizan a los visitantes de los parques nacionales es esencial para varias tareas de gestión del uso público como la toma de decisiones, la formulación de políticas y el monitoreo de dichos espacios (Eagles & McCool, 2002).

Como se discutió en el capítulo de revisión de literatura, varios autores han demostrado la validez de las fotografías geolocalizadas como indicadores del ritmo temporal de turistas y visitantes en diferentes espacios protegidos del mundo. A su vez, conocer este ritmo, es muy importante para la gestión del uso público de estos espacios, por lo que contar con datos con detalle pueden contribuir al entendimiento del comportamiento de los visitantes en el contexto del tiempo. Por ejemplo, se pueden procesar el “timestamp” para conocer la duración de la visita, las horas pico de afluencia y a escalas mayores medir el grado de estacionalidad de la actividad turística.

En este capítulo se utiliza el “timestamp” de las fotografías geolocalizadas de Flickr para realizar un análisis exploratorio del componente temporal en tres escalas de tiempo diferentes: mensual, semanal y diario. Con el fin de determinar la validez de esta nueva fuente de datos para reflejar los ritmos temporales de los visitantes, se realizó una comparación de los datos temporales de los usuarios de Flickr con los recuentos de visitantes mensuales proporcionados por el Organismo Autónomo de Parques Nacionales (OAPN). Los resultados de esta validación mostraron la precisión de los datos geolocalizados para inferir la distribución mensual de visitantes en los 15 parques que forman la red española de parques nacionales. Luego, se desagregan los datos en tres escalas de tiempo diferentes para identificar y describir el comportamiento de los visitantes en relación con su ritmo temporal. Finalmente, se calcula un índice de estacionalidad mensual y diaria para identificar los patrones temporales que caracterizan a cada parque nacional y que luego mediante el uso de algoritmos de agrupación, como k-medias (k-means) permiten su clasificación de acuerdo con su comportamiento temporal.

7.1. Metodología para el análisis temporal.

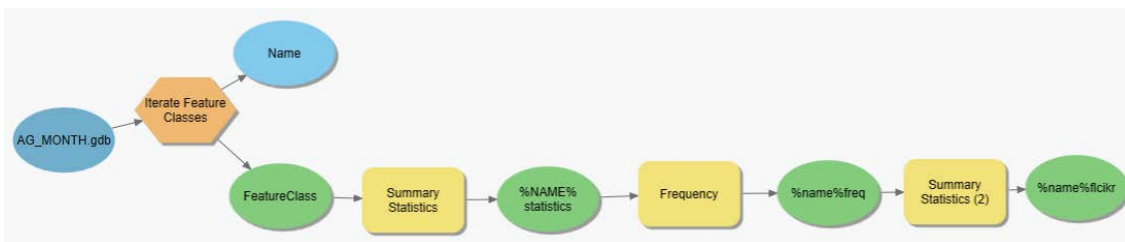
Construcción de distribuciones temporales.

A partir del conjunto de fotografías geolocalizadas de cada parque nacional, se desagregaron los datos por cada año para el periodo 2010-2018 y se contaron los usuarios únicos por día en cada mes y luego estos valores se sumaron para obtener el número mensual de usuarios únicos por año. Por usuarios únicos por día, nos referimos a contar usuarios, no el número de fotografías

realizadas, es decir, si 3 personas toman 3 fotos en el parque nacional Teide en un día, se contabilizan como 3 usuarios mientras que si una persona sube 3 fotos en un mismo día se contabiliza como 1 usuario. Una vez que se contaban con los usuarios mensuales por año se calculó el promedio para contar con la distribución promedio de usuarios de Flickr.

Para realizar la desagregación y cálculo de la distribución de forma automática de cada parque se creó un modelo de geoprocésamiento utilizando “Model Builder” de ArcGIS Pro (Figura 55). Primero, los datos se dividen en años en diferentes capas. Cada conjunto de datos anual se vuelve a dividir en meses para obtener usuarios mensuales únicos. Luego, cada capa de datos mensual se clasifica de acuerdo con el código id del usuario y el día del mes. La tabla resultante muestra cuántas fotos ha tomado cada usuario por día. A partir de esta tabla se calcula la frecuencia, que cuenta los usuarios por día en cada mes. Finalmente se suman los valores de frecuencia según el mes y se almacena los resultados en una tabla.

Figura 55. Modelo de agregación y cálculo de usuarios únicos por mes



Para obtener la distribución semanal de usuarios de Flickr, se creó en primer lugar una columna en la tabla de atributos donde se calculó el día de la semana en cada fotografía fue tomada. Para calcular el nombre del día de la semana se utilizó la función strftime de Python en la calculadora de campos de ArcGIS Pro con el código de formato “%A” que devuelve el nombre del día de la semana. El código utilizado es el siguiente: “=!Timestamp!.strftime('%A')”. Luego, se contaron los usuarios únicos por cada día de la semana utilizando el mismo razonamiento de la distribución mensual de usuarios.

La distribución diaria de visitantes se calculó a partir de la agregación del número de usuarios en franjas de una hora para las 24 horas del día, para cada mes y año. Luego estos valores se sumaron para contar con el valor total de usuarios en el periodo 2010-2018. A partir de estos valores se calculó la distribución promedio diaria de visitantes de cada parque nacional.

Validación de la distribución mensual de usuarios Flickr como indicador del ritmo mensual de visitas a la red de parques nacionales.

La validación de los datos se llevó a cabo para las distribuciones mensuales de usuarios únicos ya que los datos oficiales para todos los parques solo están disponibles a escala mensual. Se analizaron los recuentos de visitas mensuales disponibles para el periodo 2010-2018 para los 15 parques que forman la red española de parques nacionales. Estos datos se obtuvieron de las memorias anuales de cada parque y que son gestionadas por el Organismo Autónomo de Parques Nacionales (OAPN). Las memorias están disponibles para su descarga en el portal web del OAPN.

La validación se llevó a cabo mediante el cálculo del coeficiente de correlación lineal de Pearson, que permite medir la relación lineal que existe entre el número de usuarios únicos mensuales de Flickr y los recuentos mensuales de visitantes para cada parque. Este coeficiente se escogió ya que, durante la revisión del estado de arte, se observó que este indicador fue el más utilizado por estudios previos de otros autores.

Índice de estacionalidad.

Para determinar si los datos de usuarios de Flickr revelan patrones de estacionalidad de las visitas a los parques nacionales, se calculó el índice de estacionalidad de Lundtorp (ecuación **(1)**) (Lundtorp, 2001). La estacionalidad se define como la tendencia de los flujos turísticos a concentrarse en períodos relativamente cortos del año.

El índice de estacionalidad mensual (SR_m) se obtiene a partir de la división del valor más alto de visitantes registrado en un mes por el promedio de visitantes mensuales.

$$SR_m = V_{mmax} / (\bar{V}_m) \quad (1)$$

Donde: SR_m es el índice de estacionalidad mensual, V_{mmax} es el número más alto de visitantes en un mes y \bar{V}_m es el promedio de visitantes durante los meses por año. La relación aumenta con el grado de variación estacional.

Utilizando la misma ecuación de Estacionalidad de Lundtorp, se calculó un indicador para descubrir la concentración o falta de visitas en un día o días determinados, obteniendo así un indicador de estacionalidad a nivel semanal (ecuación **(2)**). Por lo tanto, la estacionalidad semanal (SR_w) se calculó como:

$$SRw = V_{dmax} / (\bar{V}_d) \quad (2)$$

Donde: SRw es el índice de estacionalidad semanal, V_{dmax} es el valor más alto de visitantes en un día y \bar{V}_d es el promedio de visitantes durante los días de la semana. La relación aumenta con el grado de variación semanal.

Tipología de parque (K-means clustering).

Los resultados obtenidos a partir del cálculo de los índices de estacionalidad revelaron la existencia de distintos patrones temporales de los visitantes de la red de parques nacionales. Utilizando el análisis de conglomerados de K-medias (K-means clustering) que agrupará los parques en grupos similares de acuerdo con las variables de estacionalidad se pretende encontrar y entender las causas de la estacionalidad en las visitas.

El algoritmo de K-medias tiene como objetivo la partición de un conjunto de n observaciones en k grupos en el que cada observación pertenece al grupo cuyo valor medio es más cercano (Jain, 2010). De esta forma cada observación de un mismo grupo será similar mientras que las diferencias entre los grupos serán amplias.

Los índices de estacionalidad mensual y semanal se utilizaron como variables para el análisis y clasificación de grupos, junto con la proporción de fotos tomadas por usuarios internacionales que se obtuvo del perfil de los usuarios de Flickr.

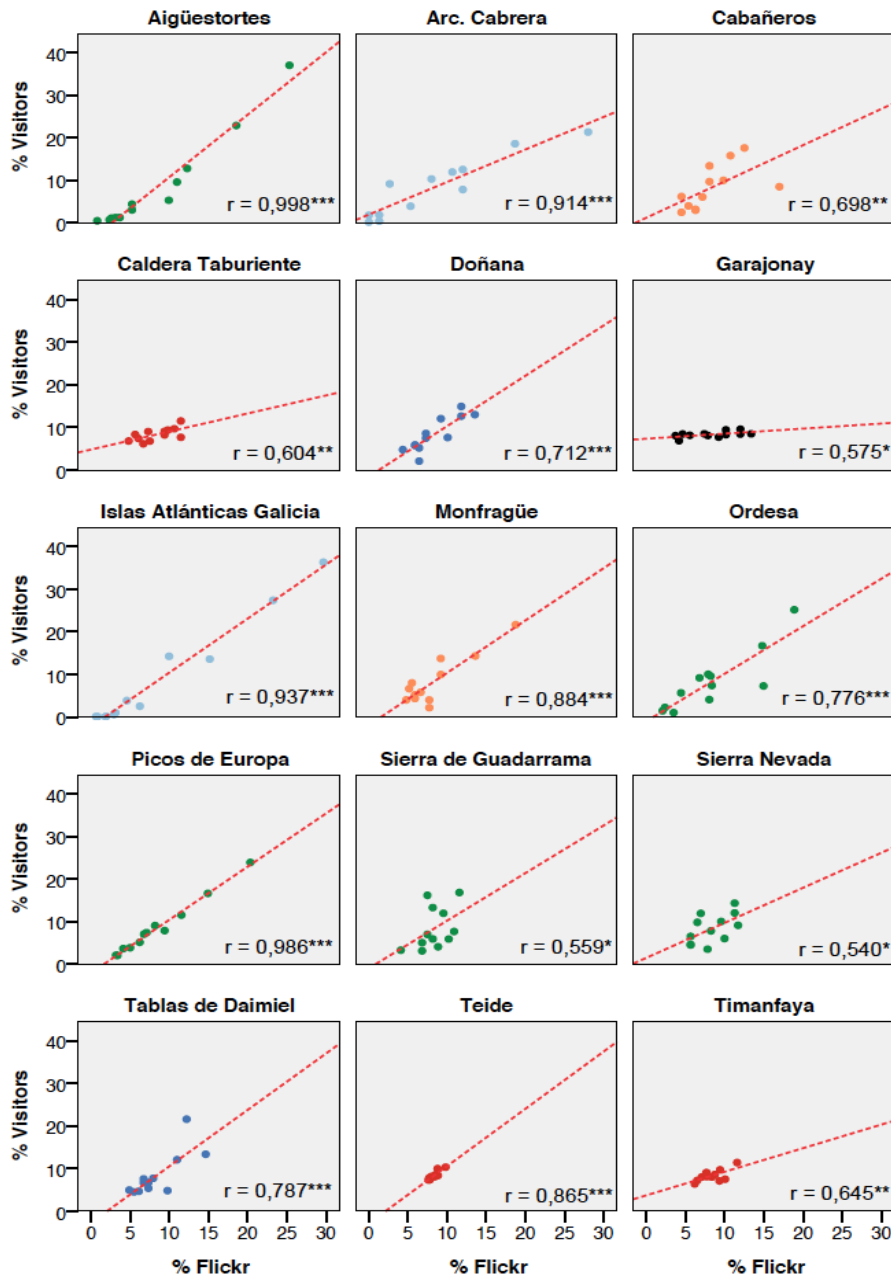
La selección del número de grupos “k” es un parámetro que se establece antes de aplicar el algoritmo. Por esta razón, se aplicó la prueba post hoc ANOVA de una vía para comparar las medias de las variables de 2, 3 y 4 grupos con el fin de seleccionar el número óptimo de conglomerados (Pham et al., 2005). Los resultados del análisis de ANOVA mostraron que el número óptimo de grupos era 4.

7.2. ¿Los usuarios de Flickr reflejan la distribución mensual de visitantes?

Los gráficos de dispersión muestran que la mayoría de los parques tienen una fuerte relación lineal entre la media de usuarios únicos por mes de Flickr y el recuento promedio mensual de visitantes (Figura 56). La correlación de Pearson confirma que todos los parques tienen una correlación positiva. Cuatro parques tienen una correlación muy alta ($r > 0,91$, $p > 0,001$). Cinco parques tienen

una correlación alta ($r > 0,71$) y seis parques tienen una correlación moderada, entre 0,7 y 0,54. De los últimos parques, tres muestran una correlación moderada, pero no significativa ($p > 0,05$), lo que podría estar relacionado con la falta de datos sobre el número de visitantes mensuales para todos los años. Dos de estos parques son Sierra de Guadarrama y Sierra Nevada. Ambos parques se encuentran muy próximos a dos grandes ciudades (Madrid y Granada, respectivamente), con una alta presencia de visitantes recurrentes (PNSG, 2016). Es de esperar que los visitantes recurrentes tomen menos fotografías a medida que aumentan sus visitas.

Figura 56. Gráficos de dispersión de la distribución mensual de % de visitantes y % de usuarios únicos de Flickr.

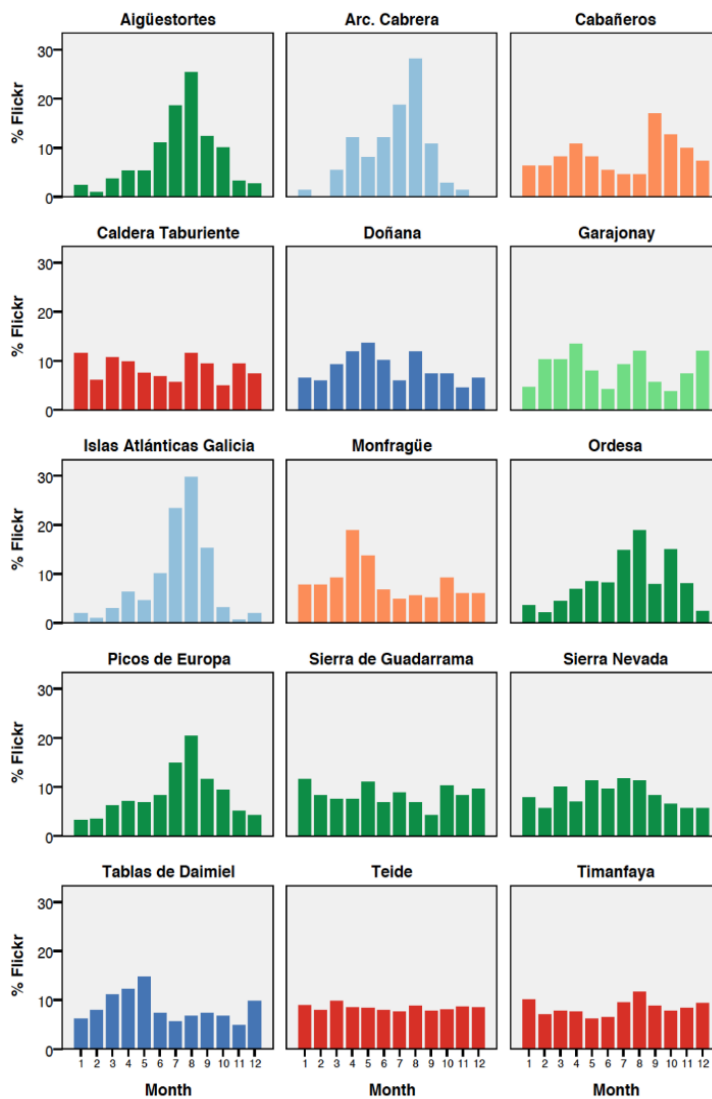


*** $P < 0,01$, ** $P < 0,05$ y * $P < 0,1$

7.3. El ritmo mensual, semanal y diario de las visitas.

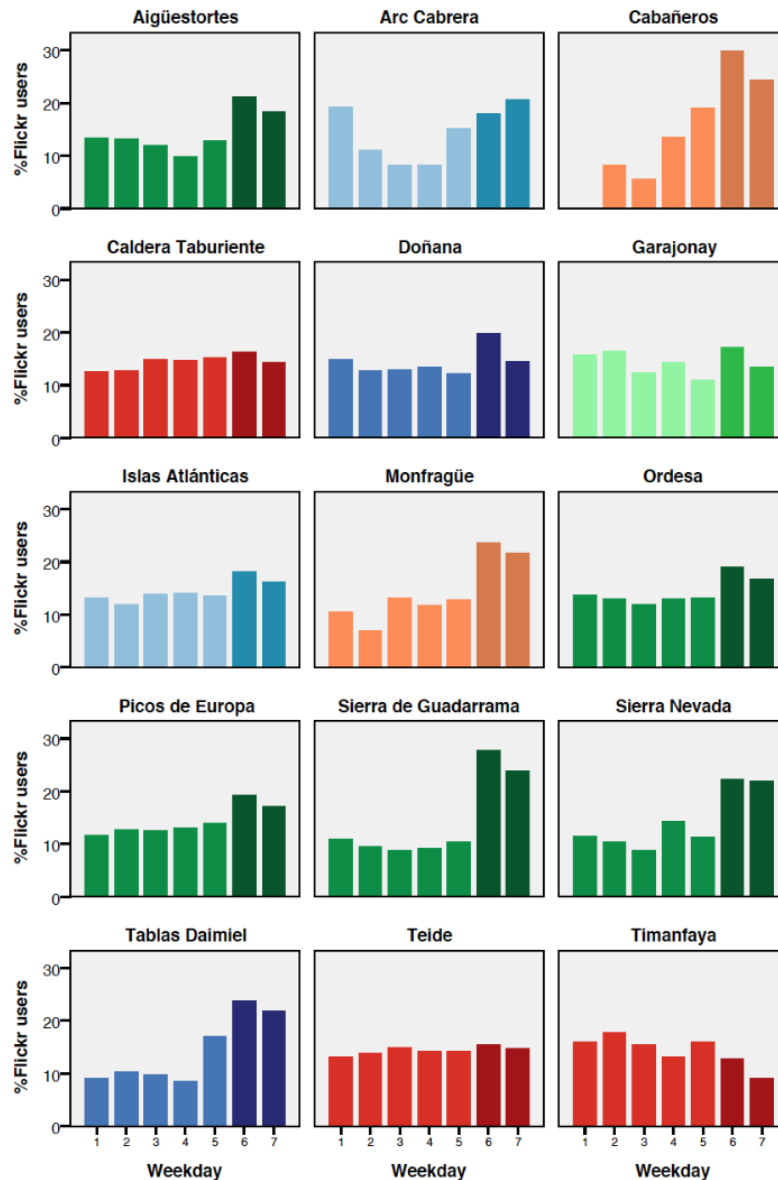
Las visitas mensuales varían significativamente entre los parques. Sin embargo, las visitas tienden a ser más altas durante los meses de verano (julio y agosto) para la mayoría de ellos. Los picos suelen coincidir con festivos y vacaciones para la población española, tales como las vacaciones de Semana Santa y las vacaciones escolares de verano. La inspección visual de las distribuciones mensuales de los parques (Figura 57) destaca claras diferencias entre los parques. Por ejemplo, nueve parques muestran cierto grado de estacionalidad, por lo que los picos en un mes se detectan fácilmente. El patrón de pico estacional de verano es mucho más pronunciado en cuatro parques: Islas Atlánticas, Archipiélago de Cabrera, Aigüestortes y Picos de Europa, mientras que otros parques muestran un patrón de visitas más regular durante todo el año por ejemplo los parques ubicados en las islas Canarias.

Figura 57. Distribución mensual de usuarios únicos de Flickr por parque nacional



Al visualizar las distribuciones semanales de los parques nacionales (Figura 58) se distinguen dos tipos de patrones. En primer lugar; parques en los que la mayoría de las vistas se concentra en torno a los fines de semana y en segundo lugar, parques en los que la distribución de visitas es más o menos regular para todos los días de la semana.

Figura 58. Distribución por días de la semana de usuarios únicos de Flickr por parque nacional



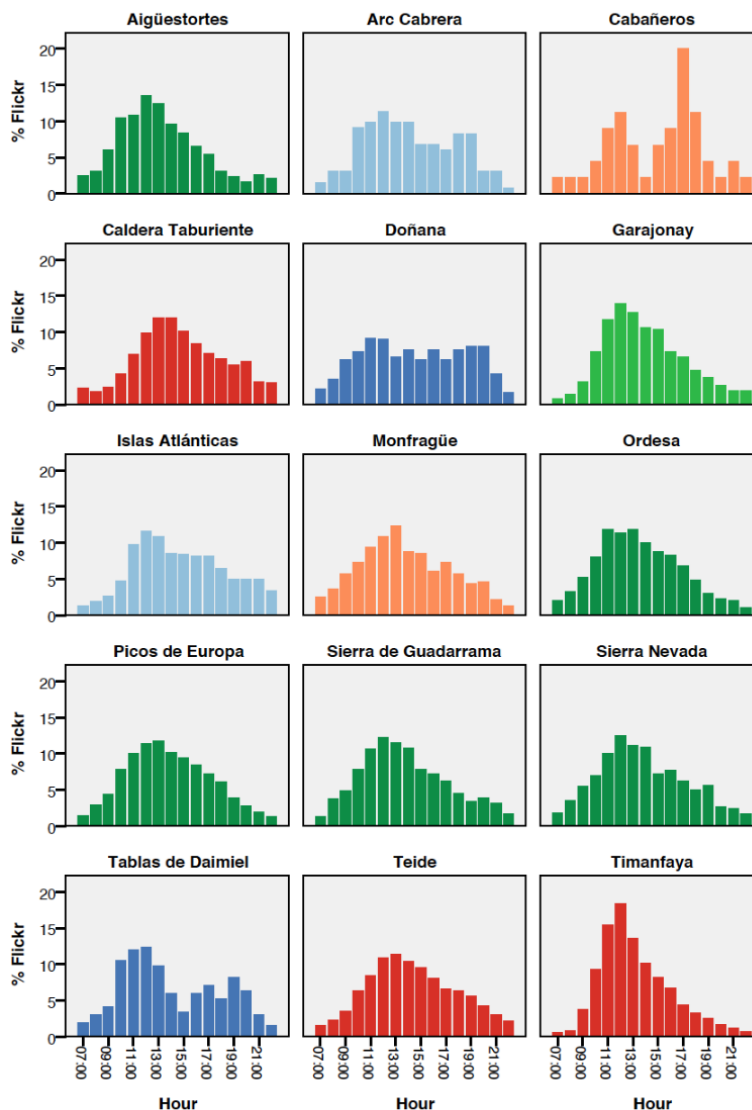
El primer grupo de parques que corresponde a la mayoría con once parques, las visitas ocurren durante los fines de semana, llegando a concentrar entre el 35% y más del 50% de la demanda en tan solo esos dos días. Entre los 11 parques, cuatro reúnen el 50% o más de visitantes en los fines de semana. Entre estos parques destaca el parque nacional Sierra de Guadarrama, cuya principal demanda podrían ser los residentes de ciudades cercanas, principalmente de Madrid. Siete parques reunieron entre el 35% y el 40% de los visitantes los fines de semana. Estos parques se

caracterizan por que la mayor proporción de su demanda son visitantes locales y reciben una pequeña proporción de visitantes extranjeros.

Los cuatro parques restantes que corresponden a los parques canarios presentan un flujo de visitantes regular a lo largo de la semana que se relaciona justamente con la vocación turística de las Islas. Los turistas al encontrarse disfrutando de un viaje de descanso, pueden visitar los parques cualquier día de la semana por lo que la demanda se distribuye de forma constante alrededor de la semana.

La distribución de visitantes por hora para cada parque se muestra en la **Figura 59**. La afluencia a lo largo del día en la red de parques nacionales presenta algunas pautas comunes, por ejemplo, la hora de mayor afluencia en la mayoría de los parques se encuentra en torno al mediodía entre las 11:00 a 13:00, siendo las 12:00 la hora de mayor concurrencia para 10 de los 15 parques

Figura 59. Distribución diaria de usuarios únicos de Flickr por parque nacional



Hay un mayor número de visitantes en la mañana que en la tarde, aunque estos datos también podrían deberse al “efecto retorno”, cuando los usuarios toman menos fotografías al volver sobre su ruta por el parque para regresar. Durante las horas de la noche, la presencia de visitantes disminuye a menos del 3% en todos los parques.

Entre las diferencias se destacan los patrones de distribución de visitas a lo largo del día, donde se pueden diferenciar que hay parques que concentran la afluencia en periodos cortos de tiempo (2-4 horas) por ejemplo los parques de Timanfaya y Cabañero. Esta concentración de visitas en torno a pocas horas puede indicar que existen momentos de saturación del parque, más aún si se analizan la ubicación de los visitantes en torno a estos momentos. En tanto que existen parques con un patrón de afluencia bastante regular donde el flujo de visitantes a lo largo del día es continuo, por ejemplo, el Teide y Archipiélago de Cabrera.

Las primeras fotografías en todos los parques se toman normalmente entre las 10:00 a.m. y las 11:00 a.m., cuando comienzan la mayoría de las visitas y que coincide con los horarios de acceso a estos espacios. Siete parques concentraron más del 45% de las visitas solo en un intervalo de tiempo de tres horas. El parque nacional de Timanfaya presenta la mayor proporción de visitantes en una hora (18%), mientras que el parque nacional de Doñana presenta el porcentaje más bajo, con un 9% en la hora punta.

7.4. Patrones de estacionalidad en las visitas a la red de parques españoles.

La mayoría de los parques, nueve en total, muestran un grado alto de estacionalidad (>1.8), por lo que los picos de visitantes en un mes se detectan fácilmente. El patrón de pico estacional es muy pronunciado en cuatro parques: Islas Atlánticas, Archipiélago de Cabrera, Aigüestortes y Picos de Europa. Según lo previsto, el pico estacional se observa en los meses de verano para todos los parques. El índice de estacionalidad mensual SRm confirma estas diferencias, resultando en valores superiores a 3 en los parques mencionados. Mientras que seis parques muestran un patrón de visitas más regular durante todo el año. Los parques ubicados en las Islas Canarias son un ejemplo. Por ejemplo, el parque Nacional del Teide que se ubica en el extremo bajo con un valor de 1,2 (Tabla 22).

El índice de estacionalidad semanal presenta menos variación entre parques, con valores que oscilan entre 1.1 y 2.1. Los parques con un SRw mayor a 1,5 reflejan que en estos espacios las visitas tienden a concentrarse en pocos días.

Tabla 22. Índice de estacionalidad mensual y semanal de los 15 parques nacionales

Parque	SRm*	% Vmax/mes	SRw**	% Vmax/día
Aigüestortes	3	25.0	1.5	21.0
Picos de Europa	2.4	20.0	1.3	19.1
Islas Atlánticas de Galicia	4	33.3	1.3	18.0
Archipiélago de Cabrera	3.4	28.3	1.4	20.6
Monfragüe	2.3	19.2	1.6	23.4
Teide	1.2	10.0	1.1	15.4
Tablas de Daimiel	1.8	15.0	1.7	23.6
Ordesa	2.3	19.2	1.3	19.0
Doñana	1.8	15.0	1.4	19.7
Cabañeros	2.2	18.3	2.1	29.7
Timanfaya	1.4	11.7	1.2	17.6
Taburiente	1.4	11.7	1.1	16.1
Garajonay	1.6	13.3	1.2	17.1
Guadarrama	1.4	11.7	1.9	27.7
Sierra Nevada	1.4	11.7	1.6	22.1

*SRm: Índice de estacionalidad mensual

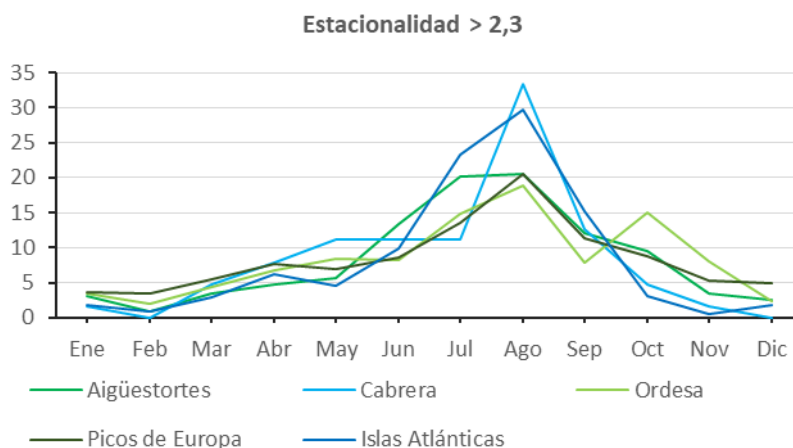
**SRw: Índice de estacionalidad diaria

% Vmax/mes: porcentaje máximo de visitas en un mes

% Vmax/día: porcentaje máximo de visitas en un día de la semana

En base al rango de valores del índice de estacionalidad mensual (Tabla 22), se pueden distinguir tres grupos que reflejan distintos grados de estacionalidad; alta, media y baja. El primer grupo reúne a los parques con un índice superior a 2.3, es decir estacionalidad alta. En estos parques se observa una marcada estacionalidad de las visitas, la mayor proporción de visitas ocurren en un periodo de tres meses, que coincide con la estación de verano (Figura 60).

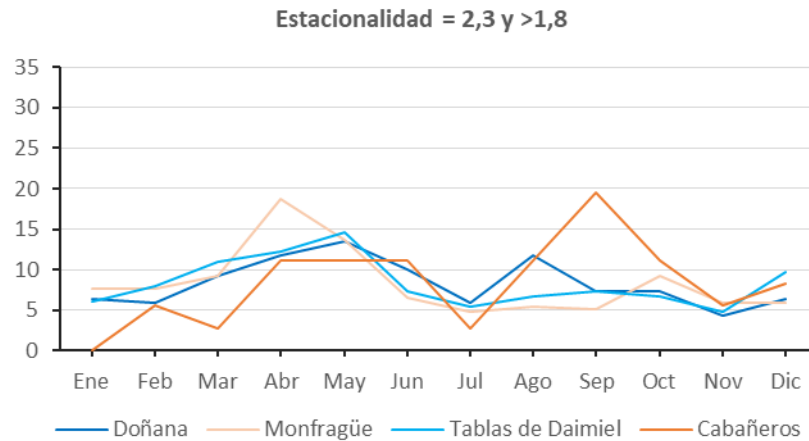
Figura 60. Distribución mensual de visitantes en parques con estacionalidad alta.



El segundo grupo se compone de los parques con un índice de estacionalidad entre 2.3 y 1.8 (Figura 61). En este grupo se observa que la mayor parte de visitas se concentran en torno a la

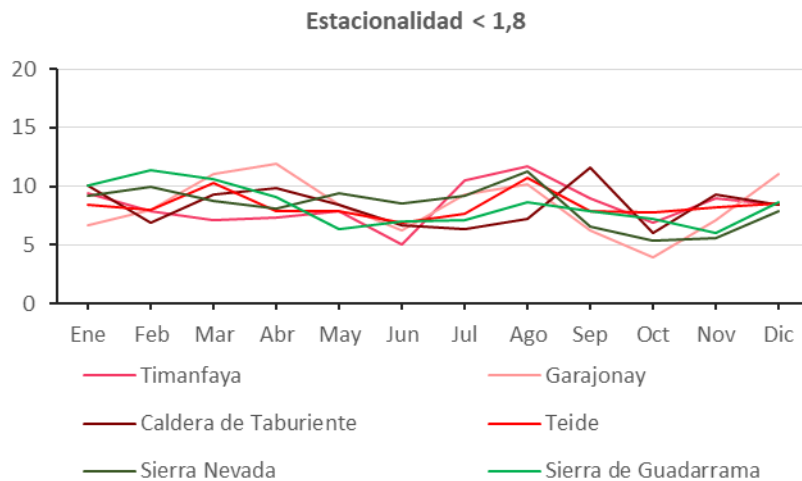
primavera (cerca del 50%), lo que puede estar ocasionado por el atractivo que ofrecen flora y fauna en esta estación, en especial por las aves que están en su periodo de anidación por ejemplo en los parques Doñana, Monfragüe y Tablas de Daimiel que son espacios conocidos como sitios de interés para la observación de aves. También durante esta estación se encuentran las festividades de Semana Santa y el puente de mayo que facilitan la presencia de visitantes sobre todo de origen nacional.

Figura 61. Distribución mensual de visitantes en parques con estacionalidad media



El tercer grupo reúne a los parques con un SRm menor a 1.8, lo que indica una afluencia regular de visitantes a lo largo del año. En este grupo se encuentran todos los parques canarios y los parques de Sierra Nevada y Guadarrama. Los parques canarios, al encontrarse en las islas Canarias cuya principal actividad económica es el turismo, tienen un flujo de turistas constante a lo largo del año que contribuyen al número de visitantes de los parques nacionales. Incluso, la visita a los parques nacionales de Timanfaya, El Teide y Caldera de Taburiente se incluyen como actividades en los paquetes turísticos para los viajeros internacionales. En tanto que parques como el de Sierra de Guadarrama y Sierra Nevada, resultan parques atractivos para los residentes de las ciudades cercanas a los mismos, los mismos que visitan estos espacios en festivos, puentes y fines de semana. Además, estos parques ofrecen una variedad de actividades durante todo el año, entre las que destacan esquí en invierno, senderismo y ciclismo de montaña.

Figura 62. Distribución mensual de visitantes en parques con estacionalidad baja



7.5. Tipología de parques según la temporalidad de sus visitas.

Se identificó un conjunto de cuatro tipos de grupos de parques a partir del análisis de K-medias. En el análisis k-medias se utilizaron los valores de los indicadores de estacionalidad por mes (SRm) y por semana (SRw) y el porcentaje usuarios internacionales como variables para la clasificación en grupos (Tabla 23).

Tabla 23. Valor de las variables utilizadas para el análisis k-medias para cada parque

Parque	SRm	SRw	% UsInt*
Aigüestortes	3	1.5	23.6
Picos de Europa	2.4	1.3	25.4
Islas Atlánticas de Galicia	4	1.3	20.1
Archipiélago de Cabrera	3.4	1.4	38.2
Monfragüe	2.3	1.6	27.2
Teide	1.2	1.1	61.4
Tablas de Daimiel	1.8	1.7	6.1
Ordesa	2.3	1.3	32.3
Doñana	1.8	1.4	30.9
Cabañeros	2.2	2.1	14.3
Timanfaya	1.4	1.2	73.6
Taburiente	1.4	1.1	68.7
Garajonay	1.6	1.2	66.5
Guadarrama	1.4	1.9	5.9
Sierra Nevada	1.4	1.6	30.1
*% UsInt: porcentaje de usuarios internacionales			

La comparación post-hoc de ANOVA unidireccional entre los grupos obtenidos mostró que las diferencias de los grupos son estadísticamente significativas para todas las variables (Tabla 24)

Tabla 24. Comparación post-hoc de ANOVA unidireccional entre grupos

Variable		Suma de cuadrados	gl*	Media cuadrática	F	Sig.
SRm	Entre Grupos	8,657	3	2,886	38,153	0
	Dentro de los grupos	0,832	11	0,076		
	Total	9,489	14			
SRw	Entre Grupos	0,56	3	0,187	3,537	0,052
	Dentro de los grupos	0,581	11	0,053		
	Total	1,141	14			
Internacionales	Entre Grupos	0,598	3	0,199	21,396	0
	Dentro de los grupos	0,102	11	0,009		
	Total	0,701	14			

gl: grados de libertad

El grupo 1 reúne a los parques con los valores más bajos para ambos índices de estacionalidad (Tabla 25). Este clúster está formado por parques con una distribución de visitantes más homogénea a lo largo del año y semanas. Aunque hay un mes pico de visitantes, la variación de visitantes se mueve entre el 5% y el 10%, aproximadamente. Estos parques se caracterizan por un flujo regular de visitantes durante el año, principalmente visitantes internacionales. Tal es el caso de los parques ubicados en Canarias, que reciben una importante afluencia de turismo internacional durante todo el año.

Tabla 25. Detalle de los grupos obtenidos a partir de K-medias

Clúster (Grupo)	N	Variabes	Media	Des. Est.	Error Est.	Centro Final
1	4	SRm	1.40	0.163	0.082	1.40
		SRw	1.16	0.068	0.034	1.16
		% UsInt	0.68	0.054	0.027	0.68
2	5	SRm	2.44	0.321	0.144	2.44
		SRw	1.57	0.310	0.139	1.57
		% UsInt	0.24	0.066	0.029	0.24
3	2	SRm	3.70	0.424	0.300	3.70
		SRw	1.35	0.127	0.090	1.35
		% UsInt	0.29	0.127	0.090	0.29
4	4	SRm	1.60	0.231	0.115	1.60
		SRw	1.63	0.235	0.117	1.63
		% UsInt	0.18	0.142	0.071	0.18

El grupo 2 muestra una alta tasa de estacionalidad por mes y por semana. En este grupo, un promedio del 20% de visitantes totales se concentran en solo un mes, principalmente en agosto. Durante la temporada de invierno, el número de visitantes del Clúster 2 caen a valores por debajo

del 10%, debido al clima severo. De los parques que forman el clúster, tres se caracterizan por ser zona de montaña (Aigüestortes, Ordesa y Picos de Europa), mientras que los dos restantes corresponden a bosques mediterráneos (Cabañeros y Monfragüe).

El grupo 3 presenta valores de índice de estacionalidad semanal ligeramente más bajos durante la semana, pero la estacionalidad más alta durante todo el año y los valores de visitantes internacionales aumentan ligeramente. Dos parques forman este grupo; Islas atlánticas de Galicia y el Archipiélago de Cabrera que protegen ecosistemas marinos. La estacionalidad se concentra durante la temporada de verano, alcanzado el valor más alto de visitantes en el agosto.

El grupo 4 presenta el índice de estacionalidad semanal más alto y una baja estacionalidad mensual. De acuerdo con el tipo de paisaje que protegen, en este grupo se distinguen dos tipos: los humedales (Doñana y Tablas de Daimiel) y parques de alta montaña (Sierra Nevada y Sierra de Guadarrama). En el caso de los humedales, se observa que el pico de estacionalidad coincide con los meses de primavera en tanto que los parques de montaña, el pico se observa en julio. La mayoría de los visitantes de este grupo corresponde a visitantes locales y ambos parques se encuentran próximos a ciudades grandes. Por ejemplo, Sierra Nevada y Sierra de Guadarrama tienen visitas regulares durante todo el año, lo que podría deberse a su proximidad a ciudades como Granada y Madrid, respectivamente).

Esta tipología está íntimamente relacionada con la clasificación basada en las características del paisaje, establecida por Somoza Medina (Somoza Medina, 2009).

Figura 63. Mapa de ubicación de los parques y asignación de los clústeres.



8. Patrones espaciales del uso turístico de la red de parques nacionales españoles.

8. Patrones espaciales del uso turístico de la red de parques nacionales españoles.

El potencial de los datos geolocalizados de redes sociales radica justamente en que son registros de eventos con coordenadas x y y , lo que permite ubicarlos en el espacio. En el contexto del turismo de naturaleza, contar con la geolocalización de los visitantes de parques nacionales presenta la oportunidad de identificar y analizar los patrones espaciales de la actividad turística dentro del parque, que son importantes para determinar entre otras cosas la capacidad de carga turística de los espacios protegidos y el desempeño de un espacio protegido como destino turístico.

En base a los datos recolectados de Flickr se aplicaron una serie de herramientas de sistemas de información geográfica para explorar, caracterizar y modelar el comportamiento espacial de los visitantes de parques nacionales.

El objetivo de este capítulo es demostrar el potencial de los datos geolocalizados para explorar la actividad turística en los parques nacionales y examinar sus patrones espaciales mediante herramientas de estadística espacial.

La primera interrogante planteada es: ¿A dónde van los visitantes dentro de los parques nacionales? Esta pregunta analiza la distribución espacial de los visitantes, a la vez que busca identificar los patrones de concentración y dispersión. Utilizando el indicador de densidad Kernel sobre las fotografías geolocalizadas de Flickr, se crearon mapas de densidad que permiten visualizar y clasificar a cada espacio protegido en base a la distribución de las fotografías.

La siguiente pregunta que se planteó es: ¿Cuáles de los lugares donde las concentraciones de visitantes son más representativas? Con esta pregunta se pretende identificar los puntos de interés turístico dentro de los parques nacionales que atraen a la mayoría de los visitantes, en base al número de usuarios de Flickr. Para obtener las localizaciones representativas, se utilizaron los índices de autocorrelación global y local de Moran.

Una vez que se ha obtenido y validado la distribución espacial de los visitantes en los parques nacionales, surge una pregunta final: ¿Podemos utilizar esta distribución como un proxy de la demanda en un modelo de localización-asignación que nos ayude a identificar las ubicaciones óptimas para nuevos puntos de información? Esta interrogante busca resaltar una aplicación potencial de los datos geolocalizados en la gestión del uso público en los parques nacionales. Para ilustrar esta aplicación, se ha seleccionado el parque nacional Teide como nuestra área de estudio.

8.1. Metodología para el análisis espacial y la localización óptima de instalaciones de servicios.

Densidad Kernel.

Los análisis de densidad son una herramienta fundamental para examinar la distribución espacial y la intensidad de un fenómeno en un área determinada, como el turismo en un parque nacional. Un enfoque particularmente útil es la densidad Kernel (KDE por sus siglas en inglés). Este método no paramétrico permite estimar la densidad de un conjunto de puntos o líneas de manera precisa y detallada (Grekousis, 2020). El objetivo de este método es producir una visualización de la distribución espacial de un fenómeno, basándose en un mapa de calor entre áreas centrales (kernels) y sus vecinos circundantes. Al utilizar la densidad de Kernel, podemos obtener una representación visual de cómo se concentra el turismo en diferentes áreas del parque nacional, identificando las zonas de mayor actividad y las que presentan menor afluencia.

En el campo de la estadística espacial, el término "kernel" se utiliza para describir una ventana centrada en cada una de las celdas del ráster para el que se calcula, que se mueve sistemáticamente a cada ubicación mientras calcula la respectiva función. Las funciones Kernel ponderan los eventos en el entorno de la celda en función de su distancia, de manera que los eventos cercanos reciben una mayor ponderación que aquellos más distantes (Tran et al., 2006). En este trabajo de investigación se utilizó la función cuadrática de Kernel (Silverman, 1986) que sigue la siguiente fórmula:

$$D = \frac{1}{nh^2} \sum_{i=1}^n \left[\frac{3}{\pi} \left(1 - \left(\frac{d_i}{h} \right)^2 \right)^2 \right]$$

Donde:

h: radio de búsqueda (tamaño de la ventana)

n: número de eventos (puntos) dentro del área de búsqueda

d_i: distancia entre la celda para la que se calcula la densidad y el evento dentro del radio de búsqueda.

El radio de búsqueda define el tamaño de la ventana sobre la que se va a calcular la densidad, por lo que la selección de este parámetro es crucial para obtener un resultado óptimo (Jiménez, 1991). Por ejemplo, un radio más grande mostrará patrones más generales en la distribución de la densidad ("puntos de calor"), mientras que radios pequeños reflejan los patrones locales en la distribución de la densidad.

En esta investigación, se empleó la técnica de densidad de Kernel para estudiar las pautas de concentración espacial de los visitantes de parques nacionales, utilizando las localizaciones de las fotografías geolocalizadas de Flickr como fuente de datos y un radio de búsqueda de 500 metros. La elección de este radio en particular se basó en la comparación de los efectos de diferentes radios y en una revisión exhaustiva de los valores utilizados en estudios de turismo presentes en la literatura (Gutiérrez-Puebla & García-Palomares, 2016; Lyon et al., 2011).

La utilización de un radio de 500 metros permitió capturar de manera precisa y representativa la distribución espacial de los visitantes, brindando una visión detallada de las áreas de mayor concentración y actividad turística en los parques nacionales españoles.

Índice Global de Moran.

La autocorrelación espacial es el grado de dependencia espacial, asociación o correlación entre el valor de una observación de una entidad espacial y los valores de observaciones vecinas de la misma variable (Grekousis, 2020). La autocorrelación mide cuánto se relaciona el valor de una variable en una ubicación específica con los valores de la misma variable en ubicaciones vecinas.

Una de las estadísticas que miden la autocorrelación espacial, es el índice de Moran. El índice global de Moran mide la autocorrelación espacial de un conjunto de datos basada en las ubicaciones y valores de los datos (Moran, 1948). Dado un conjunto de entidades geográficas y un atributo asociado, este índice evalúa si el patrón de distribución de ese atributo es agrupado, disperso o aleatorio (Getis, 2008). Los valores P y la puntuación z que calcula este índice nos indican la significación estadística del patrón espacial que forman los datos.

La ecuación para calcular el índice Global Moran es la siguiente:

$$I = \frac{n}{\sum_i^n \sum_j^n w_{ij}} \frac{\sum_i^n \sum_j^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i^n (x_i - \bar{x})^2}$$

Donde:

n : número de entidades espaciales

x_i : valor del atributo en la entidad i

x_j : valor del atributo en la entidad j

\bar{x} : media del atributo

w_{ij} : peso según el grado de vecindad entre las entidades i y j .

$\sum_i^n \sum_j^n w_{ij}$: agregación de todos los pesos espaciales.

En esta investigación, el índice global de Moran, se utilizó para comprobar y medir el grado de concentración espacial de los datos agregados de usuarios de Flickr y a partir de esta medida, realizar análisis locales con el Índice local de Moran.

Índice Local de Moran.

La autocorrelación espacial global identifica si existe una pauta general de agrupamiento o dispersión en los valores de una variable, pero no indica dónde se encuentran dichos agrupamientos. Además, la magnitud de la autocorrelación espacial puede variar dentro del espacio analizado, debido a la heterogeneidad espacial (Grekousis, 2020). Para estimar la autocorrelación a nivel local y localizar las zonas donde se presentan los agrupamientos, se utilizan estadísticos locales como el índice local de Moran.

El índice local de Moran es una estadística espacial inferencial que se utiliza para calcular la autocorrelación espacial a nivel local. Para los objetos espaciales en un vecindario “n” (Anselin, 1995), el I local de Moran de una entidad espacial se expresa como:

$$I_i = \frac{x_i - \bar{X}}{m_2} \sum_j w_{ij} (x_j - \bar{X})$$

$$m_2 = \frac{\sum_i (x_i - \bar{X})^2}{n}$$

Donde:

n : número de entidades espaciales

x_i : valor del atributo en la entidad i

x_j : valor del atributo en la entidad j

\bar{X} : media del atributo

w_{ij} : peso espacial según el grado de vecindad entre las entidades i y j.

m_2 : constante para todas las localizaciones.

El resultado consiste en la representación de las entidades espaciales (ubicaciones) con valores significativos del índice de Moran de acuerdo con los valores p y z, identificando así la presencia de cuatro tipos de agrupamientos: “High-high clusters” (zonas de agrupamiento de valores altos), “Low-Low clústers” (zonas de agrupamiento de valores bajos), “High-Low outliers” (valores altos atípicos rodeados de valores bajos) y “Low-High outliers” (valores bajos atípicos rodeados de valores altos).

En esta investigación, se aplicó el índice local de Moran para identificar los puntos de interés turístico (POIs, por sus siglas en inglés: Points of Interest), utilizando como variable la cantidad de usuarios de Flickr por hexágono. Los POIs se definieron como los clústeres HH (High-High), que son las áreas donde se observa una concentración significativamente alta de usuarios. También se analizaron los clústeres de valores atípicos altos HL (High-Low outliers), porque pueden indicar la presencia de puntos de interés o atracciones turísticas específicas que se destacan en medio de áreas menos activas o la presencia de nuevos puntos de interés descubiertos por los turistas.

El análisis del índice local de Moran permitió detectar de manera precisa los lugares que generan un gran interés turístico.

Modelos de localización-asignación.

La distribución espacial de la demanda, el atractivo de la localización y la proximidad de las instalaciones son las variables necesarias para aplicar un modelo de ubicación y asignación. Los modelos de ubicación-asignación se utilizan para seleccionar las ubicaciones óptimas de un conjunto de instalaciones propuestas y asignar la demanda respectiva a estos sitios. En este sentido, para la identificación de localizaciones potenciales para nuevos puestos de información en el parque nacional del Teide, se propone utilizar la distribución espacial de los usuarios de Flickr como un indicador de la distribución espacial de la demanda potencial de puestos de información. El fundamento de esta aplicación es que cuanto mayor sea el número de visitantes junto a los quioscos, mayor será la probabilidad de que sean utilizados por los visitantes.

Se utilizó el módulo de análisis de red de ArcGIS para modelar la demanda potencial a partir de los usuarios de Flickr y utilizar el solucionador de localización-asignación para encontrar las ubicaciones óptimas de los centros de información. De las seis soluciones propuestas en ArcGIS, optamos por el enfoque de maximizar la asistencia. Esta solución tiene como objetivo maximizar la atención de la demanda que la instalación puede cubrir dentro de una distancia específica (Holmes et al., 1972). En este modelo, las instalaciones se eligen para asignar la mayor ponderación posible de la demanda, asumiendo que la ponderación de la demanda disminuye con respecto a la distancia entre la instalación y el punto de demanda (ESRI, 2018).

Los modelos de ubicación y asignación requieren tres entradas principales; una red vial, un conjunto de instalaciones candidatas (y, en su caso, existentes) y los puntos de demanda (Mitchell, 2012). Construimos una capa de red a partir de las capas de caminos y senderos proporcionadas por OAPN. Se consideraron 557 ubicaciones candidatas para el establecimiento de los puestos de

información. Estos se ubican a lo largo de la vía principal que cruza el parque a un intervalo fijo de 100 metros debido a que la vía conecta los cuatro puntos de acceso al parque y es el punto de partida de la mayoría de los senderos, por lo que todos los visitantes deben transitar por esta vía. La demanda potencial de los visitantes se modeló a partir de las fotografías geolocalizadas de Flickr, haciendo un recuento de usuarios únicos agregados en la capa de hexágonos que luego se transformó en una capa de puntos. Cada punto de demanda representaba el número de usuarios dentro de un solo hexágono. Asimismo, en el modelo se consideraron dos instalaciones requeridas que corresponden a los dos centros de visitantes existentes.

Se consideraron cuatro escenarios basados en el número de instalaciones adicionales a las dos existentes: 1, 3, 5 y 10. También se establecieron varias distancias de corte (1000, 1500, 2000 y 2500 m) para realizar un análisis de sensibilidad.

8.2. La distribución espacial de los visitantes de la red de parques nacionales.

Los resultados del análisis de densidad proporcionan una representación visual clara de las áreas con mayor presencia de usuarios dentro de los parques nacionales. Los mapas de densidad nos permiten tener una primera aproximación a la distribución espacial del turismo en cada espacio protegido (Figuras 64, 65, 66). Estas visualizaciones son especialmente útiles para identificar tres elementos clave: los valores máximos de densidad, el tamaño de las áreas de máxima densidad y el número de zonas con densidad máxima.

Al examinar los valores medios y máximos de densidad (Tabla 26), podemos identificar el grado de presión que experimentan los parques nacionales en función de la concentración de turistas. Su localización a partir de la cartografía puede indicar puntos de interés turístico, atracciones populares o áreas de mayor afluencia. Las estadísticas descriptivas calculadas a partir de la densidad Kernel nos permiten diferenciar tres grupos de parques.

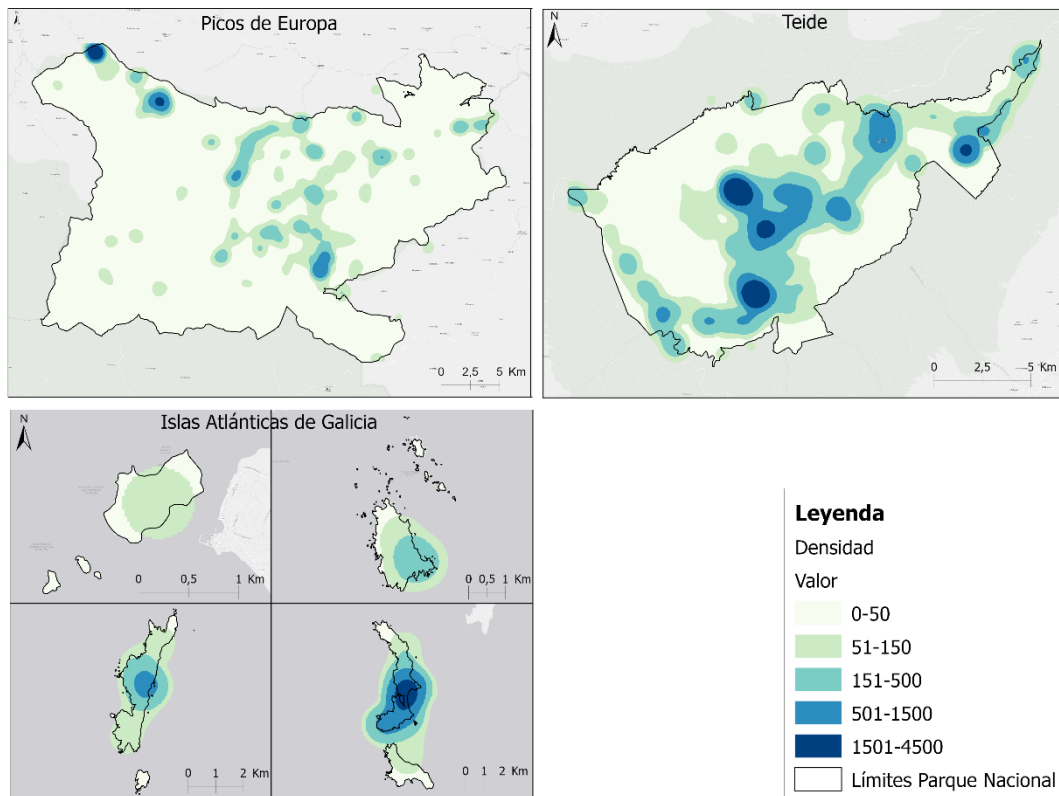
En el primer grupo se encuentran los parques donde los valores de densidad máxima son los más altos y que comprenden a los parques de Teide, Picos de Europa y las Islas Atlánticas de Galicia. Sobresalen los parques de Teide y Picos de Europa que presentaron valores de 4670 y 4252 fotos/km² respectivamente, dos de los parques más visitados. Los parques representan distintos paisajes; alta montaña, volcánico e insular. En general, los visitantes están distribuidos en toda el área de los parques, observándose áreas de densidad media-alta en torno a los equipamientos de uso públicos y lugares de interés natural.

El parque formado por las islas Atlánticas de Galicia revela que las islas de Ons y Cíes son las islas preferidas por los visitantes, siendo la isla de Monteagudo la única que presenta un área de densidad máxima. Mientras que la isla Cortegada, la más pequeña, presenta una densidad muy baja.

En el parque nacional Picos de Europa, la zona de Covadonga (Iglesia y alrededores) presenta la densidad máxima, y en segundo lugar los lagos Enol y Ercina. Otras zonas que presentan una densidad media son Fuente De donde se encuentran equipamientos de uso público (funicular, parqueadero, mirador y camping).

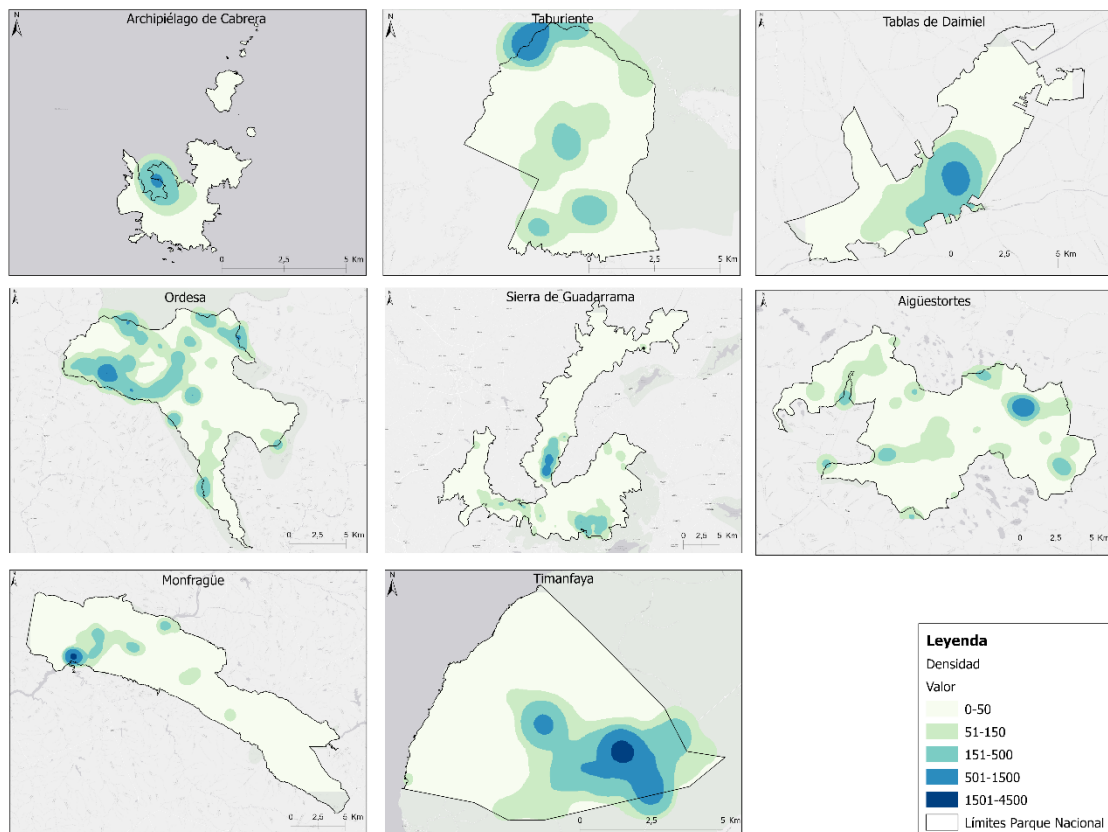
Las localizaciones con densidades máximas en el parque Teide corresponden con dos puntos importantes; el pico del Teide, que constituye el principal atractivo de este espacio y el centro de visitantes Cañada Blanca, donde se cuenta con varios equipamientos de uso público (información, restaurante, aparcamientos). Se observa un tercer foco de densidad máxima que no necesariamente corresponden con sitios de atractivo turístico, sino que son los lugares donde se puede acceder o tomar fotografías de los atractivos paisajísticos que ofrece el PN Teide. Este es el caso de la estación base del teleférico para la subida a la cima del Teide y el mirador de Minas de San José.

Figura 64. .Densidad de fotografías de Flickr por Km2 en el grupo 1 (parques con valores altos de densidad máxima)



En un segundo grupo se encuentran la mayoría de los parques nacionales y presentan densidades máximas de fotografías que oscilan entre 500 y 2000 fotografías por km². Entre los parques que se encuentran en este rango se incluyen el Archipiélago de Cabrera, la Caldera de Taburiente, las Tablas de Daimiel, Ordesa, Sierra de Guadarrama, Aigüestortes, Monfragüe y Timanfaya. Los parques de este grupo representan diferentes los tipos de paisaje, así tenemos tres parques de alta montaña, dos volcánicos, un archipiélago, un humedal y un bosque mediterráneo

Figura 65. Densidad de fotografías de Flickr por Km² en el grupo 2 (parques con densidad máxima entre 500 y 2000 fotos/km²)



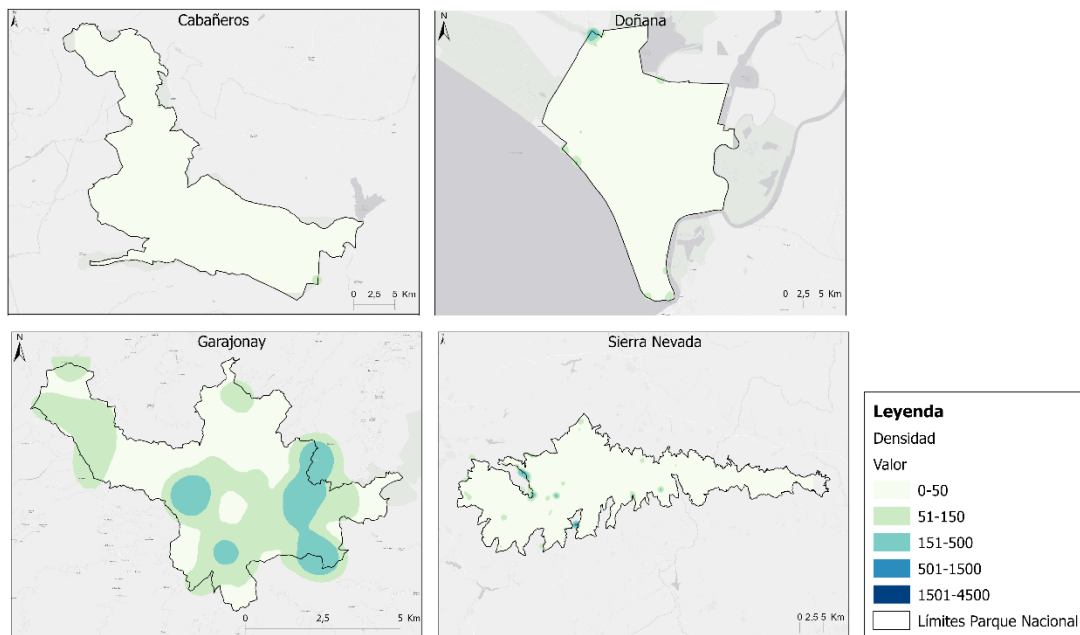
Como era de esperar, las zonas de mayor densidad coinciden con los equipamientos de uso público y áreas de gran atractivo natural en los distintos parques. En el caso del Parque Caldera de Taburiente, las áreas de mayor concentración son el punto de información de Roque de los Muchachos y la zona de acampada de Roque de la Brevera Mecha. En los parques de Tablas de Daimiel y Sierra de Guadarrama, los centros de visitantes son los puntos de máxima densidad. En Sierra de Guadarrama, además, se encuentra un segundo punto de alta densidad en el centro de información de Puerto de los Cotos. Por otro lado, en el Parque de Ordesa y Monte Perdido, los visitantes se concentran principalmente en el centro de visitantes Casa Oliván. En el Parque Nacional Timanfaya, la zona de mayor densidad se encuentra en el área de ingreso al parque conocida como "Echadero de camellos". En cambio, en los Parques de Monfragüe y Aigüestortes,

los puntos de máxima densidad se relacionan con las ubicaciones de los atractivos naturales de los parques. Por ejemplo, en Monfragüe, la densidad máxima se encuentra en torno al castillo, mientras que en el Parque Aigüestortes, la zona de mayor densidad coincide con el Estany Sant Maurici. Estos resultados confirman la relación entre los puntos de interés, los servicios de uso público y la concentración de visitantes en los diferentes parques nacionales.

En contraste en el tercer grupo tenemos los parques cuyas densidades máximas no sobrepasan las 500 fotografías por Km² y que incluyen a los parques Cabañeros, Doñana, Garajonay y Sierra Nevada. Sobresalte el parque nacional Cabañeros que registró el valor más bajo de densidad máxima con 119 fotos por km².

En este grupo se encuentran los parques con mayor superficie, como es el caso de los parques Sierra Nevada y Doñana. Además, es importante señalar que el parque Doñana cuenta con restricciones de acceso y recorrido dentro del parque, por lo que se evidencia que la única zona de concentración de visitantes corresponde con el acceso al parque en la localidad de El Rocío. En el parque Sierra Nevada en cambio, son pocas las áreas accesibles debido al relieve del parque.

Figura 66. Densidad de fotografías de Flickr por Km² en el grupo 3 (parques con densidad máxima baja menor a 500 fotos/km²)



A nivel de parque se puede señalar lo siguiente; los puntos de mayor densidad en el parque nacional Cabañeros coinciden con el centro de visitantes y área de recreación Casa Palillos, y el área de recreación de Acebo y las fuentes. Zonas con valores de densidad media (250) se encuentran distribuidas en la zona de rutas 4x 4 y el punto de información de la encina en el norte del parque. En cuanto al parque nacional Sierra Nevada, los puntos de mayor densidad se localizan

en dos refugios de alta montaña La Carihuela y La Caldera. Otros puntos de densidad media coinciden con refugios y miradores ubicados en el occidente del parque. En el parque nacional Garajonay, se observa que el área de densidad máxima es mayor y se encuentran múltiples puntos con densidad máxima, como es el caso de los miradores, punto de información de laguna grande, El Contadero y Los Roques. El parque nacional Doñana, presenta solo una localización de densidad máxima que coincide con el ingreso al parque en la aldea de El Rocío. Finalmente, el parque nacional del Archipiélago de Cabrera, la zona con densidad máxima coincide con el área recreativa, La playa S'Espalmador, el centro de información y el punto de partida a la ruta hacia el Castell.

Tabla 26. Tabla resumen de densidad de fotos de Flickr por parque nacional

Parque	Densidad media	Densidad Máxima	DT	CV
Aigüestortes	48,34	1378,46	105,98	2,19
Cabañeros	1,28	119,76	5,06	3,95
Tablas de Daimiel	86,81	753,46	140,61	1,62
Sierra de Guadarrama	37,35	1123,37	93,21	2,50
Monfragüe	29,37	1782,50	116,11	3,95
Ordesa	93,84	1503,49	141,96	1,51
Picos de Europa	44,99	4252,44	166,23	3,69
Sierra Nevada	7,21	452,89	19,98	2,77
Teide	209,96	4670,04	454,55	2,16
Timanfaya	143,31	1932,40	274,17	1,91
Caldera de Taburiente	73,64	1234,42	126,70	1,72
Garajonay	79,03	341,99	67,22	0,85
Archipiélago de Cabrera	48,96	554,45	90,89	1,86
Islas Atlánticas de Galicia	316,41	2363,15	469,81	1,48
Doñana	5,36	497,83	19,50	3,64

DT: Desviación Típica, CV: Coeficiente de variación

Al evaluar el tamaño de las zonas de máxima densidad, se obtiene información importante sobre la extensión y la magnitud de las zonas turísticas más concurridas. Esto puede ser crucial para la gestión de visitantes y la implementación de medidas adecuadas para preservar el equilibrio entre el turismo y la conservación de los espacios naturales. Para analizar el tamaño de estas zonas, se extrajeron las zonas con una densidad igual o superior a 300 fotos por km², en lugar de 500 fotos por km², con el fin de incluir en el análisis el mayor número posible de parques (Tabla 27). El único parque que no presenta zonas con densidades superiores a las 300 fotos por km² es el Parque Cabañeros.

En cuanto al tamaño de las zonas de máxima densidad, se observa que los Parques Teide, Picos de Europa y Ordesa presentan las zonas más extensas de alta densidad, registrando valores superiores

a 100 hectáreas. En contraste los parques de Archipiélago de Cabrera, Sierra Nevada y Garajonay presentan las zonas más pequeñas con valores que varía por debajo de las 50 hectáreas.

Es aún más interesante examinar la superficie ocupada en relación con la superficie total del parque, ya que esto proporciona una mejor idea de la ocupación del suelo para uso público. En este sentido, se observa que, en los Parques Islas Atlánticas de Galicia, Teide y Timanfaya, las áreas de máxima densidad ocupan el mayor porcentaje del área total del parque, con valores superiores al 15% de la superficie total. Esto podría indicar que los visitantes se concentran en áreas específicas de gran atractivo, como las zonas con servicios turísticos

Por otro lado, es importante destacar que, en Doñana y Sierra Nevada, las áreas de máxima densidad ocupan una proporción mínima de la superficie total de estos espacios, representando menos del 1% del área total. Esto indica que la mayor parte del parque se mantiene en un estado natural preservado, con una ocupación humana relativamente baja en relación con su extensión.

Tabla 27. Superficie y número de zonas con densidades superiores a 300 fotografías/Km2.

Parque	Área (ha)	Porcentaje*	# zonas
Aigüestortes	234,4	1,7	2
Tablas de Daimiel	231,1	7,6	1
Sierra de Guadarrama	847,5	2,5	2
Monfragüe	224,5	1,2	1
Ordesa	1085,4	6,9	4
Picos de Europa	1388,1	2,1	8
Sierra Nevada	48,7	0,06	2
Teide	3337,8	17,6	10
Timanfaya	797,8	15,6	2
Caldera de Taburiente	130,6	2,8	2
Garajonay	40,6	1	5
Archipiélago de Cabrera	50,5	3,8	1
Islas Atlánticas de Galicia	309,8	25,9	2
Doñana	64,5	0,1	1

*Con respecto al área total del parque

Finalmente, al considerar el número de zonas con densidad máxima, podemos obtener información relevante sobre la cantidad de zonas de alta concentración turística en los parques nacionales. Este análisis nos ayuda a comprender la distribución espacial general del uso turístico y ofrece datos valiosos para la planificación de rutas, la asignación de recursos y la toma de decisiones estratégicas en la gestión de los parques.

Además, la identificación de zonas de densidad alta puede indicar una presión turística significativa debido a la concentración de visitantes en esas áreas específicas. Estas zonas pueden requerir una atención especial en términos de gestión y medidas para garantizar la sostenibilidad ambiental y la calidad de la experiencia turística. En la **Tabla 27**, se presenta un desglose detallado del número de zonas que superan el umbral de 300 fotos por km² en cada uno de los parques.

Al analizar los resultados, se observa que los Parques Teide y Picos de Europa presentan el mayor número de zonas con densidades superiores a 300 fotos por km² con 10 y 8 zonas respectivamente. Esto indica la presencia de múltiples áreas que podrían estar experimentando una mayor presión turística debido a su atractivo y popularidad.

En contraste, en la mayoría de los otros parques, la densidad máxima se concentra en alrededor de una o dos zonas específicas. Esto indica que la presión turística está más focalizada y se puede abordar de manera más efectiva mediante la implementación de medidas de gestión adaptadas a esas áreas particulares.

En conjunto, los resultados del análisis de densidad de fotografías, incluyendo los mapas de densidad y la información resumida en las tablas, ofrecen una valiosa perspectiva sobre la distribución espacial del turismo en los parques nacionales y permiten identificar patrones clave para la gestión y planificación efectiva de estas áreas protegidas. Además, los resultados resaltan la importancia de identificar y monitorear las zonas de alta densidad en los parques nacionales y revelan que en los parques nacionales españoles los visitantes ocupan áreas pequeñas del territorio de los parques.

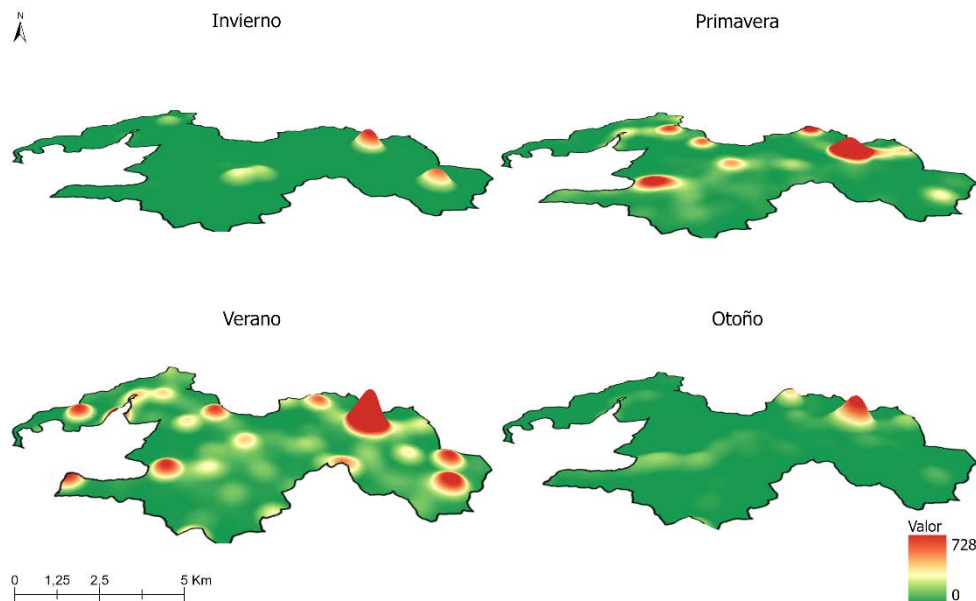
8.3. Análisis espacio - temporal de los visitantes de la red de parques nacionales.

Para confirmar y analizar si la concentración de los visitantes es un fenómeno permanente o estacional, se ha calculado la densidad Kernel para cada estación en cinco parques, que representan un parque de cada tipo de la clasificación de Somoza (Somoza Medina, 2009).

Como representante de los parques de montaña, se ha seleccionado al parque Aigüestortes (**Figura 67**), que, como se discutió previamente en el capítulo anterior, exhibe un alto nivel de estacionalidad. Se observa que la concentración y distribución espacial de visitantes se encuentra muy influenciada por el momento de la visita. Se puede apreciar que las densidades máximas ocurren durante la estación de verano, mientras que en las demás estaciones la densidad disminuye, lo que indica una reducción en la presión sobre el parque. Además, se evidencia que, en primavera y verano, estaciones con condiciones más favorables para la visita, hay un mayor

número de zonas con alta densidad. Por el contrario, en otoño e invierno se identifica una alta densidad concentrada en una sola ubicación, correspondiente al lago Sant Maurici. Esta tendencia se podrá profundizar a través del análisis de autocorrelación espacial local en las secciones siguientes, lo que permitirá identificar con mayor precisión los puntos de interés con el mayor número de usuarios y, por lo tanto, una mayor presión sobre el parque.

Figura 67. Densidad de fotografías geolocalizadas por estación del parque nacional Aigüestortes

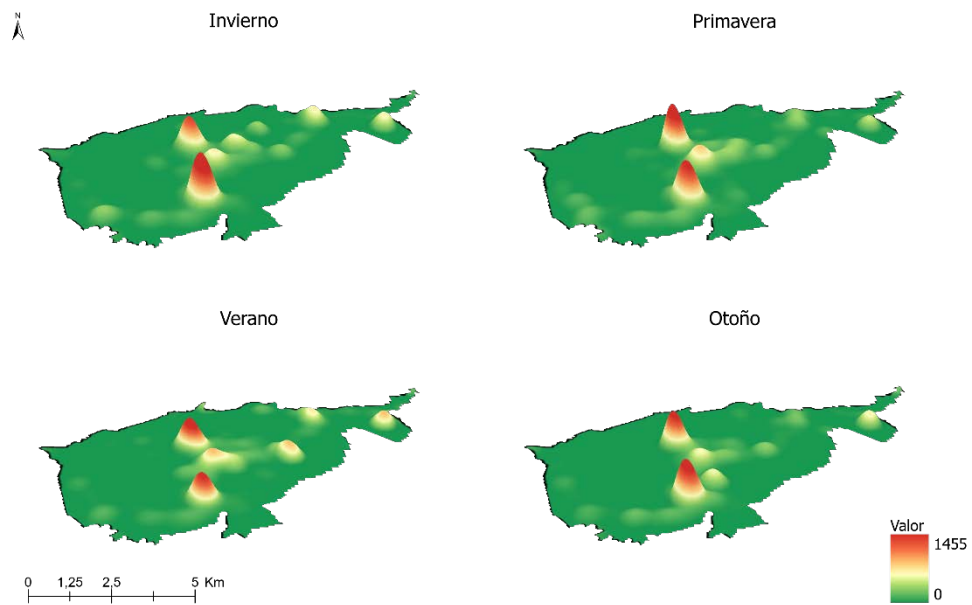


En el caso de los parques canarios, el parque del Teide (Figura 68), presentó el índice de estacionalidad más bajo, lo que indica que las densidades altas en los puntos más fotografiados se mantienen durante todas las estaciones. Esto revela que la zona del teleférico en el pico del Teide y la zona del centro de visitantes Cañada Blanca son las áreas más susceptibles a la presión turística y soportan la mayor concentración de visitantes a lo largo de todo el año.

En la primavera, se observa que los picos de densidad en el pico del Teide y en Cañada Blanca alcanzan sus valores más elevados. Estos puntos destacados atraen a un gran número de visitantes durante esta estación, probablemente debido a las condiciones climáticas favorables.

En el verano, se observa una disminución en los picos de densidad en el pico del Teide y Cañada Blanca. Sin embargo, aparecen nuevos focos de intensidad media alrededor de la estación base del teleférico, el mirador de Minas San José y el centro de visitantes Portillo. Estos lugares también experimentan una afluencia considerable de visitantes durante la temporada, lo que indica la existencia de otros puntos de interés populares que pueden aliviar parcialmente la presión sobre las áreas principales.

Figura 68. Densidad de fotografías geolocalizadas por estación del parque nacional Teide

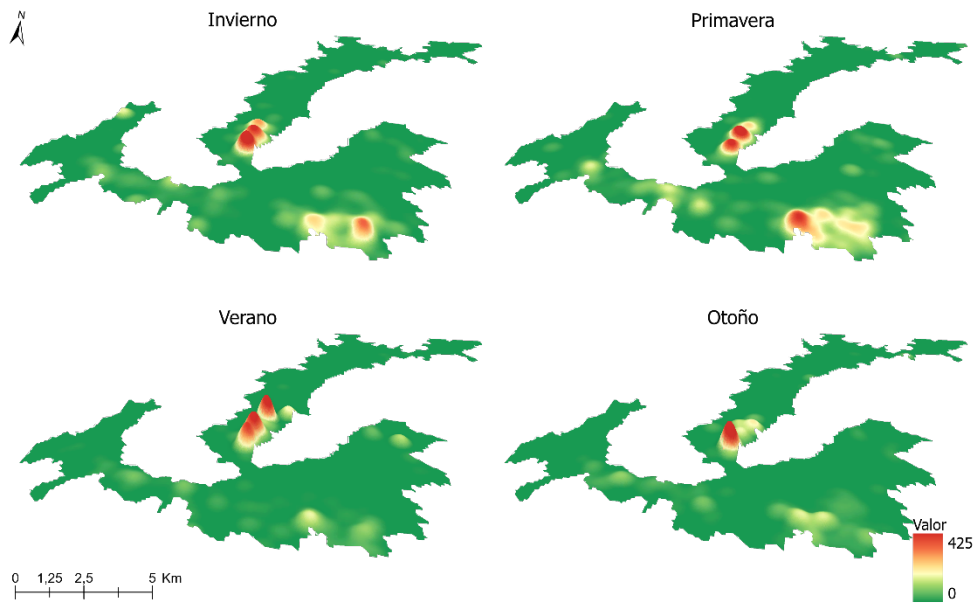


En el parque sierra de Guadarrama (Figura 69), la densidad de fotos está distribuida de forma regular en todas las estaciones. Esto indica que la captura de fotografías se produce de manera constante en diferentes áreas del parque a lo largo del tiempo. El puerto de los Cotos, ubicado en el centro del parque, se destaca como una zona con densidad máxima en todas las estaciones. Este punto específico atrae a un gran número de visitantes durante todo el año, posiblemente debido a su fácil acceso y su oferta de actividades al aire libre.

En invierno y primavera, se observan también focos de alta densidad en la zona de la Pedriza, lo que indica que esta área en particular es popular entre los visitantes durante esas estaciones. Estos momentos del año pueden ofrecer condiciones favorables para la práctica de actividades al aire libre. Adicionalmente, en invierno y primavera se observan también focos de densidad media distribuidos en el sector sur del parque. Esto indica la presencia de otras áreas con un nivel moderado de afluencia de visitantes durante esas estaciones.

En otoño, se observa solo un foco de alta densidad de fotografías localizado en el puerto de los Cotos, lo que indica que la concentración de visitantes se centra en esta área central del parque.

Figura 69. Densidad de fotografías geolocalizadas por estación del parque nacional Sierra de Guadarrama

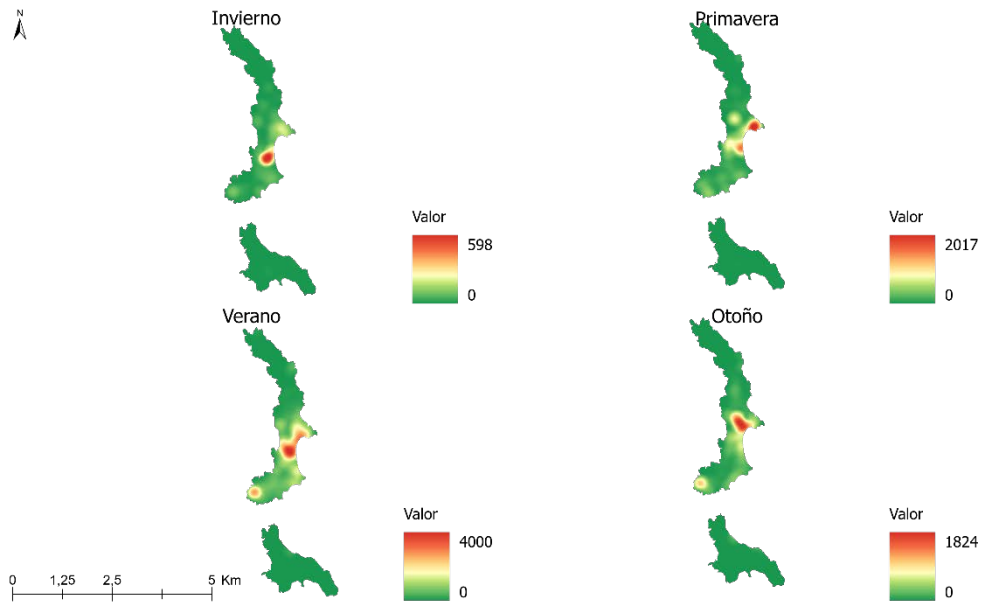


En el parque nacional Islas Atlánticas de Galicia, se observa una variación en la densidad de fotos según la estación y la isla, siendo las islas menores de Cortegada y Sálvora las que presentan una actividad relativamente baja en comparación.

En particular, la isla Ciés (Figura 70) se destaca como el principal atractivo para los visitantes, ya que agrupa la mayor cantidad de fotos. Tanto en primavera como en verano, se registra una mayor densidad de fotos en esta isla. Durante el verano, se alcanzan valores máximos de hasta 4000 fotos por km², indicando una gran afluencia de visitantes y una alta actividad turística en ese período.

Es importante destacar que, en comparación con los demás parques analizados, las Islas Atlánticas presentan la mayor densidad de fotos. Este hecho puede sugerir una mayor presión turística en el parque, lo que resalta la importancia de implementar medidas adecuadas para la gestión del turismo y la conservación del entorno natural en este delicado ecosistema insular.

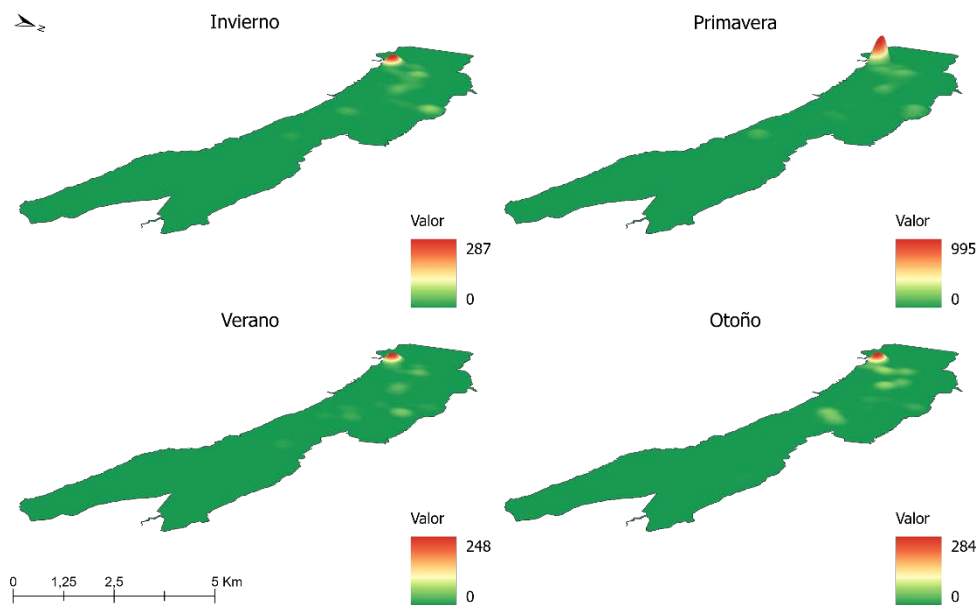
Figura 70. Densidad de fotografías geolocalizadas por estación del parque nacional Islas Atlánticas de Galicia



En el parque de Monfragüe (Figura 71), se destaca que la mayoría de las fotos se tomaron durante la primavera, la cual coincide con la época óptima para la observación de aves. Durante esta estación, se registran las densidades máximas, alcanzando un valor de 995 fotos por km².

Es interesante destacar que el foco de intensidad alta se mantiene en todas las estaciones y está asociado a los observatorios de aves de Villareal de San Carlos. Estos observatorios son puntos estratégicos dentro del parque donde los visitantes tienen la oportunidad de avistar y fotografiar diversas especies de aves, lo cual explica la concentración de fotos en esa área específica a lo largo del año. Asimismo, la presencia de un foco de alta intensidad durante todas las estaciones en los observatorios de aves de Villareal de San Carlos indica su relevancia como puntos de interés destacados dentro del parque.

Figura 71. Densidad de fotografías geolocalizadas por estación del parque nacional Monfragüe



El análisis de densidades proporcionó una aproximación inicial de la distribución de las fotografías en la red de parques nacionales, que reveló los lugares con mayor interés fotográfico dentro de cada parque. Esta aproximación permite visualizar la distribución de los visitantes en el territorio de cada parque e identificar las tendencias de distribución ya sea de concentración en torno a un solo punto o una agrupación en torno a varias localizaciones.

No obstante, para comprender mejor las pautas del comportamiento espacial de los visitantes y determinar las localizaciones específicas que constituyen puntos de interés turístico, se llevó a cabo un análisis de autocorrelación espacial global y local. Utilizando el número de usuarios por unidad de área (hexágono) como unidad de estudio (en este caso hexágonos con un tamaño de 300 m), se identificaron y delimitaron las zonas que concentran a la mayoría de los usuarios o fotógrafos. Al comparar estos resultados con las localizaciones de los puntos de interés natural y los equipamientos de uso público designados por el OAPN, fue posible establecer la importancia de cada sitio como atractivo y verificar su popularidad.

Este análisis ofrece una visión más detallada de las preferencias y patrones de comportamiento de los visitantes en cada parque nacional. Nos permite identificar las zonas de mayor afluencia y popularidad, así como evaluar la correlación entre los puntos de interés designados y la concentración de usuarios o fotógrafos en esas áreas.

8.4. Análisis de autocorrelación espacial: identificando puntos calientes.

El índice global de autocorrelación espacial calculado para los usuarios de Flickr muestra valores positivos y estadísticamente significativos en todos los parques nacionales. Esto indica claramente la presencia de autocorrelación espacial positiva en cada uno de los parques (Tabla 28). Estos resultados demuestran una clara tendencia de agrupamiento espacial de los visitantes especialmente pronunciada en los parques de Doñana y en los parques insulares (Islas Atlánticas de Galicia y Archipiélago de Cabrera), donde los valores de autocorrelación son los más elevados. La autocorrelación positiva sugiere que ciertas áreas han captado un mayor interés y afluencia de visitantes, lo que ha dado lugar a la formación de clústeres de alta concentración de usuarios.

Tabla 28. Índice global de autocorrelación espacial por parque nacional

Parque	índice	z-score	valor p
Aigüestortes	0,43	31,62	0,0000
Cabañeros	0,27	15,07	0,0000
Tablas de Daimiel	0,63	37,36	0,0000
Sierra de Guadarrama	0,48	35,39	0,0000
Monfragüe	0,44	35,58	0,0000
Ordesa	0,40	30,88	0,0000
Picos de Europa	0,15	21,24	0,0000
Sierra Nevada	0,30	27,84	0,0000
Teide	0,30	24,09	0,0000
Timanfaya	0,32	30,36	0,0000
Caldera de Taburiente	0,29	29,87	0,0000
Garajonay	0,23	17,06	0,0000
Archipiélago de Cabrera	0,44	37,31	0,0000
Islas Atlánticas de Galicia	0,60	60,52	0,0000
Doñana	0,50	54,5	0,0000

Una vez, confirmada la existencia de un patrón de agrupación espacial, se calculó el índice local de autocorrelación espacial de Moran (LISA) para cada uno de los parques nacionales. El análisis LISA permite mapear la localización de los clústeres de valores altos y valores bajos que se relacionan con áreas específicas donde se concentran zonas de alta o baja afluencia de usuarios (figuras 72 y 73).

Los mapas resultantes presentan la clasificación de los clústeres en 4 categorías; Estas categorías incluyen "HH" (High-High), "LL" (Low-Low), "HL" (High-Low) y "LH" (Low-High). Los clústeres "HH" se pueden describir como "Clústeres de valores altos rodeados por valores altos", mientras que los

clústeres "LL" representan "Clústeres de valores bajos rodeados por valores bajos". Las categorías HL y LH representan los hexágonos con valores atípicos altos y bajos.

Los mapas resultantes muestran que en todos los parques se han detectado hexágonos que forman clústeres de valores altos (HH). Estos clústeres muestran ubicaciones específicas dentro de los parques que atraen a un número significativo de visitantes, creando puntos críticos de actividad turística. El número y extensión de los clústeres HH varían de un parque a otro. Los parques en los que se observa un mayor número de hexágonos HH son Ordesa, Teide, Timanfaya e Islas Atlánticas de Galicia. En contraste en los parques que presentan un número menor de hexágonos HH son Cabañeros, Archipiélago de Cabrera y Doñana, que también se encuentran entre los menos visitados. De forma general, los clústeres HH identificados están asociados a puntos de interés naturales, como lagos, miradores y atractivos culturales, como centros de visitantes y ermitas en los parques nacionales.

En segundo lugar, el análisis LISA también identificó hexágonos que forman clústeres de valores bajos (LL) y que se asocian con una actividad baja, es decir donde la presencia de usuarios fue menor de lo esperado. Los clústeres de hexágonos LL fueron identificados en 10 de los 15 parques y presentan un menor número de hexágonos en comparación con los clústeres HH. Destacan los parques de Teide, Sierra de Guadarrama, Ordesa y Picos de Europa por presentar un mayor número de hexágonos LL.

Además, se han identificado hexágonos con valores atípicos altos y bajos (HL, LH), y que corresponden a áreas donde se registraron valores anómalos en la distribución de usuarios. Por ejemplo, los hexágonos con valores atípicos altos (HL) indican ubicaciones donde se registraron un número inusualmente alto de usuarios y que posiblemente pueden estar asociados a nuevos puntos de atracción turística. Se encontraron hexágonos HL en 12 parques nacionales. Los hexágonos HL fueron más frecuentes y abundantes en parques con un menor número de usuarios como Cabañeros y Tablas de Daimiel. En general los hexágonos HL, están asociados a puntos de partida de rutas de senderismo y otros atractivos naturales como ríos y cuevas.

En contraste los hexágonos con valores atípicos bajos (LH) indican ubicaciones donde se registraron un número bajo de usuarios cuando se esperaba un número elevado de usuarios. Estos hexágonos se encuentran en todos los parques y en su mayoría se encuentran en la vecindad de los hexágonos HH. Por otro lado, los hexágonos LH que se encuentran lejos de los clústeres HH, podrían estar asociados con lugares con accesibilidad limitada o con características menos atractivas para los visitantes. Los parques de Picos de Europa y Garajonay registraron el mayor número de hexágonos LH.

Figura 72. Mapas del índice local de autocorrelación espacial LISA (1 de 2).

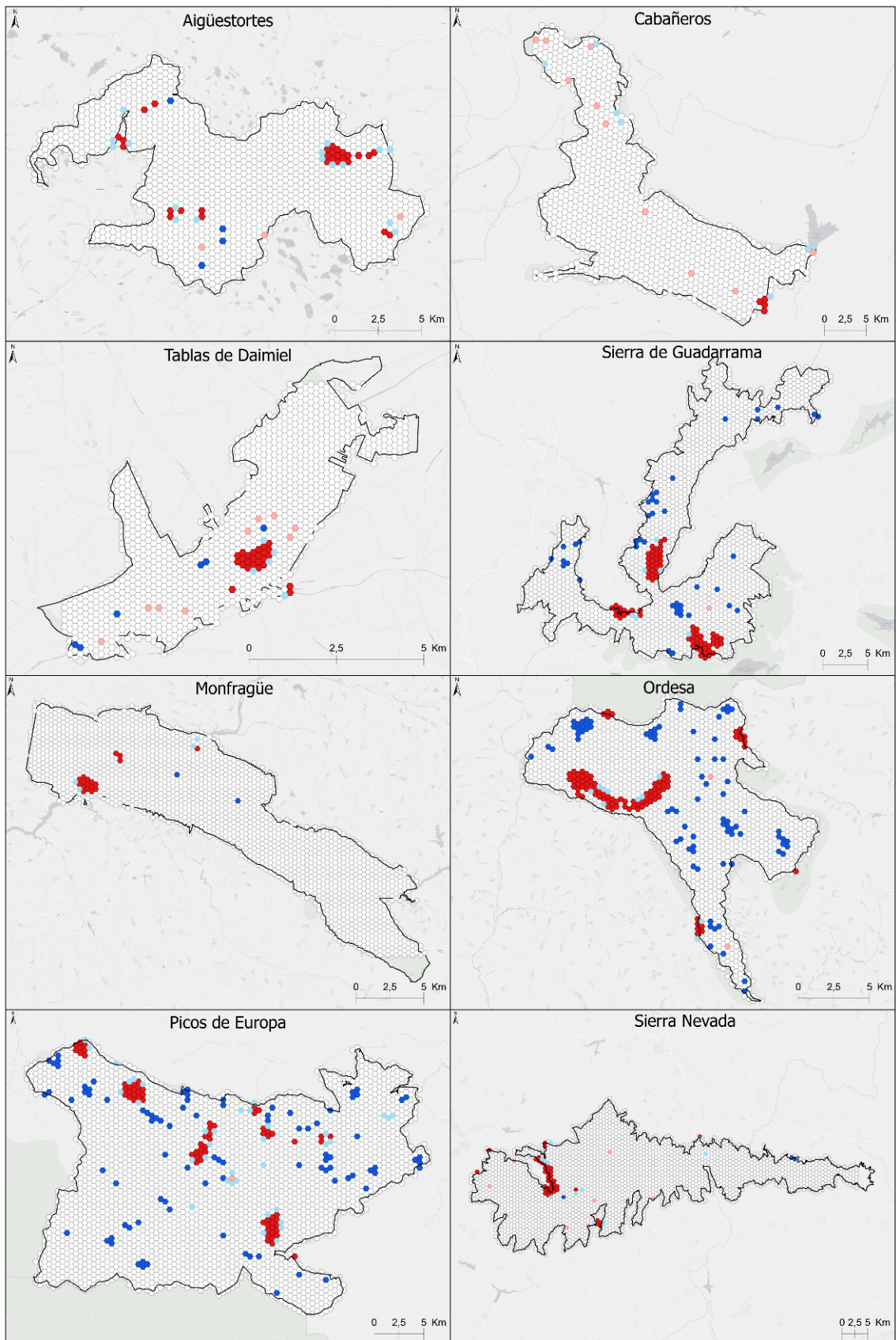
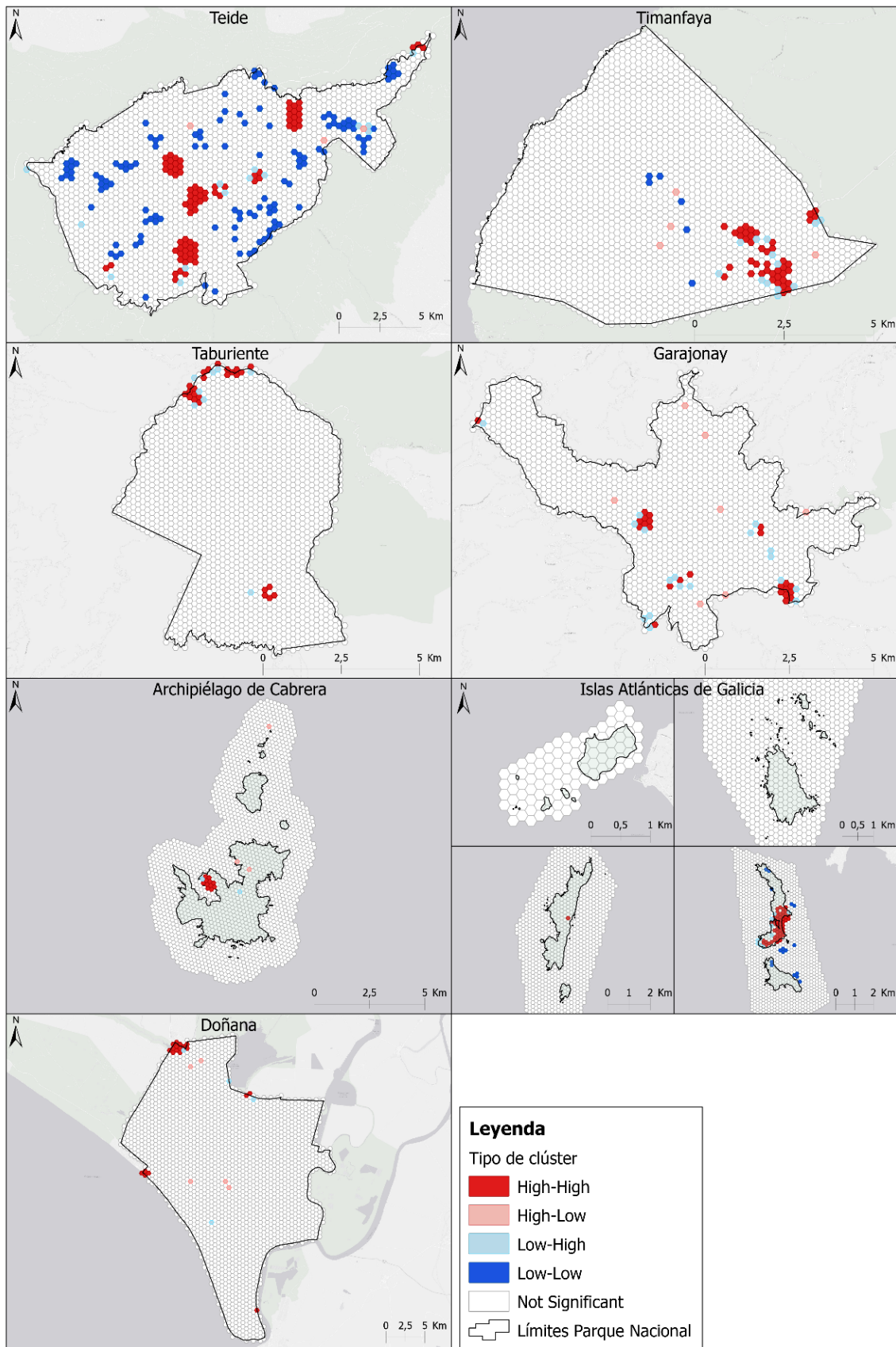


Figura 73. Mapas del índice local de autocorrelación espacial LISA (2 de 2).



Al analizar la relación entre la superficie ocupada por los clústeres del análisis LISA y la superficie total de cada parque (Tabla 29), se evidencia que la ocupación del espacio por parte de los visitantes es reducida. Menos del 12% del territorio de los parques es utilizado por los visitantes. Esta concentración en áreas pequeñas, que favorece la conservación general del parque, podría generar presiones sobre dichos espacios y ocasionar impactos negativos, sobre todo en los clústeres HH. Sobresalen los parques de Teide, Ordesa, Sierra de Guadarrama y Picos de Europa como los parques en los cuales los clústeres ocuparon los porcentajes más altos de superficie (entre 7% y 11%), los mismos que se encuentran entre los parques más visitados en la red.

Tabla 29. Detalle del número de hexágonos por clústeres HH, LL, HL y LH y superficie que ocupan en los parques.

Parque	HH	LL	HL	LH	Total	Sup. (ha)	% sup.
Aigüestortes	28	4	3	16	51	476,87	3,38
Cabañeros	4	0	10	7	21	763,32	1,87
Tablas de Daimiel	25	6	9	5	45	108,91	3,59
Sierra de Guadarrama	73	42	1	6	122	2375,64	7,00
Monfragüe	18	2	0	4	24	236,48	1,29
Ordesa	88	81	2	10	181	1701,20	10,84
Picos de Europa	71	92	1	36	200	4422,11	6,59
Sierra Nevada	41	3	8	11	63	2275,90	2,65
Teide	80	123	3	14	220	2122,87	11,18
Timanfaya	42	6	4	13	65	157,94	3,09
Caldera de Taburiente	25	0	0	8	33	85,27	1,82
Garajonay	24	0	7	18	49	121,40	3,05
Archipiélago de Cabrera	11	0	3	2	16	9,82	0,74
Islas Atlánticas de Galicia	51	16	0	10	77	26,74	2,24
Doñana	20	0	5	4	29	647,40	1,19

Para identificar y asociar los clústeres de valores altos hallados con los atractivos turísticos y equipamientos de uso público dentro de los parques nacionales se realizó un análisis de superposición mediante el cual se combinaron los hexágonos HH con las localizaciones de los atractivos turísticos y equipamientos de uso público designados por el OAPN. A través de este análisis fue posible asociar cada clúster de hexágonos HH a un punto de atracción turística o de equipamiento de uso público.

Por ello, en primer lugar, debido al número elevado de tipos de equipamiento y puntos de interés definidos por el OAPN, estos se clasificaron en 5 categorías que identifican su función para el visitante. Las categorías establecidas son las siguientes:

- Equipamiento de acogida: agrupa los equipamientos que sirven para dar información a los visitantes sobre las características de los parques nacionales y sus servicios.
- Equipamiento de alojamiento: comprende las instalaciones para el alojamiento de visitantes.
- Equipamiento de recreación: Espacios y facilidades para actividades de recreación.
- Equipamiento de servicios: Equipamientos y servicios básicos para los visitantes, como baños, estacionamientos, áreas de descanso, etc.
- Equipamiento Educativo: Equipamientos destinados a la educación y concienciación de los visitantes sobre los valores naturales y culturales de los parques nacionales.

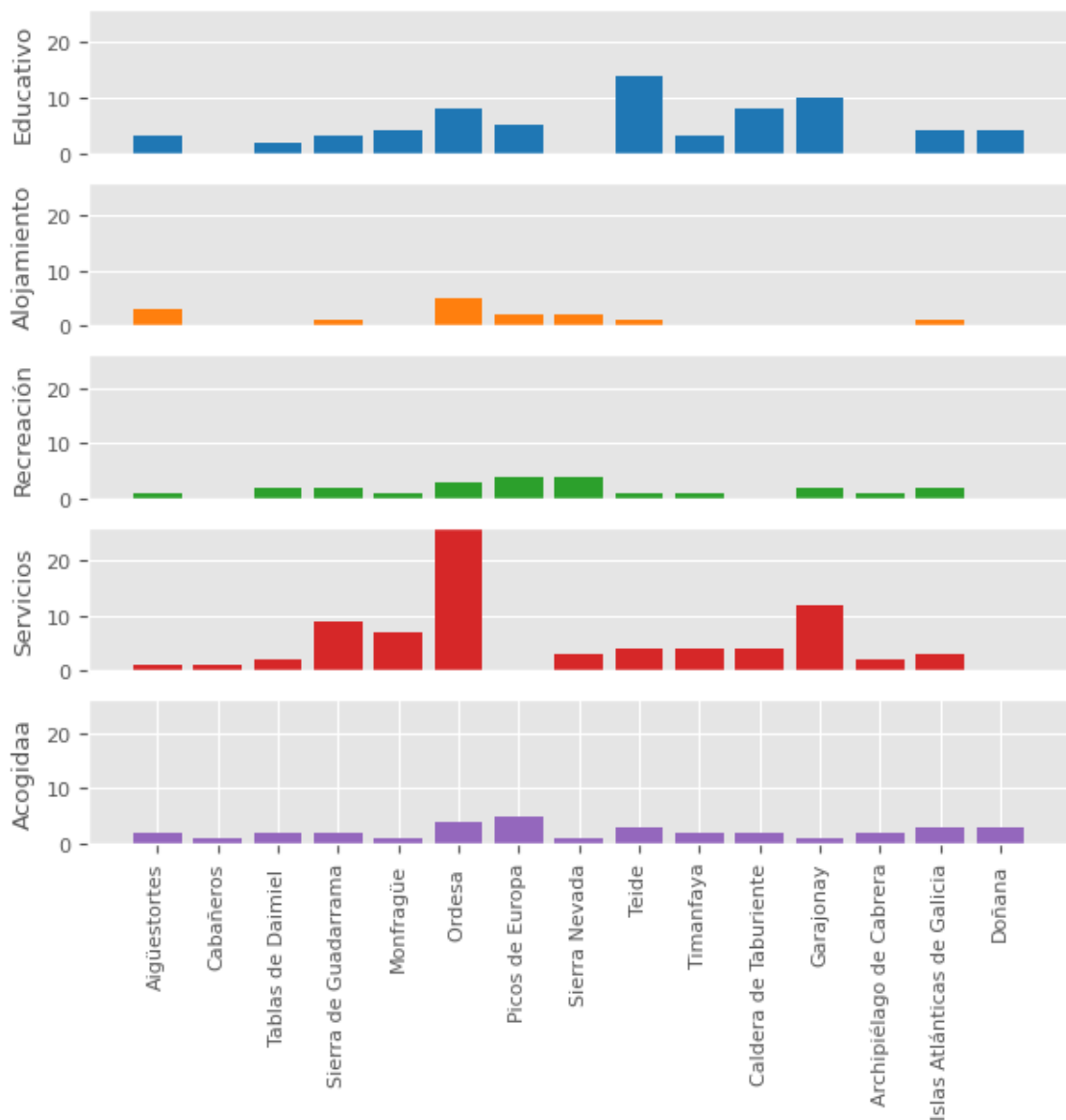
En la Tabla 30 se presenta el resultado de la clasificación, en la se puede resaltar que la mayoría de los equipamientos pertenecen a las categorías de servicios y educación evidenciando su vocación de posibilitar su uso y disfrute.

Tabla 30. Número de equipamientos y puntos de interés por tipo

Clasificación	# de equipamientos	Descripción
Equipamiento de acogida	121	Centros de visitantes, puntos de información, oficina de información, caseta de información, servicio de interpretación ambiental.
Equipamiento de alojamiento	70	Camping, refugio.
Equipamiento de recreación	154	Punto de interés, área de fondeo, apeadero de camellos, área recreativa.
Equipamiento de servicios	257	Aparcamiento, Puesto de socorro, puente, pasarela, aseos, muelle, parada de autobús, barrera, parador, bar.
Equipamiento educativo	242	Mirador, observatorio, faro, aula de educación.

El resultado resumido del análisis de superposición se presenta en la **Figura 74**. Para una comprensión más detallada de cada parque, se proporciona en la sección de Anexos, apartado 1, donde se detalla la identificación de los equipamientos asociados a los hexágonos HH y HL.

Figura 74. Distribución de los clústeres HH por tipo de equipamiento



De manera similar al análisis de asociación de los clústeres HH, también se llevó a cabo un análisis de superposición entre los equipamientos de uso público y los hexágonos HL. Los resultados de este análisis se presentan en la **Tabla 31**.

Tabla 31. Hexágonos HL por tipo de equipamiento.

Parque	Tipo de equipamiento				
	Educativo	Alojamiento	Recreación	Servicios	Acogida
Cabañeros	0	0	1	4	1
Tablas de Daimiel	1	0	0	0	0
Ordesa	1	0	0	0	0
Sierra Nevada	1	0	0	0	0
Garajonay	1	0	0	1	0

Para concluir, en la **Tabla 32** se presentan los clústeres HH que han registrado la mayor afluencia de usuarios en cada parque, junto con su correspondiente punto de interés asociado. Estos clústeres representan los puntos más atractivos en cada parque y, en muchos casos, se convierten en la esencia misma de la experiencia del parque, como sucede en Aigüestortes, Ordesa, Picos de Europa y Tablas de Daimiel donde el clúster HH está asociado con el atractivo que representa la identidad del parque, el Estany San Maurici en Aigüestortes, la zona de Covadonga en Picos de Europa, etc.

Además, la cantidad de hexágonos que componen cada clúster HH proporciona una indicación de la intensidad de la demanda en estos puntos específicos. Un ejemplo de esto es el parque Picos de Europa, donde en tan solo 8 hexágonos se registraron 1496 usuarios, en marcado contraste con el parque menos visitado, Cabañeros, donde el clúster HH solo reunió a 17 usuarios.

Tabla 32. Clústeres HH que atrajeron el mayor número de usuarios en los parques nacionales.

Parque	POI asociado	Hexágonos	Usuarios	Media (u/h)
Aigüestortes	Estany Sant Maurici	14	476	34
Cabañeros	Casa Palillos	4	17	4
Tablas de Daimiel	Isla de Pan	22	311	13
Sierra de Guadarrama	Laguna Grande Peñalara	28	893	32
Monfragüe	Ermita de Monfragüe	14	602	43
Ordesa	Valle de Ordesa	69	1780	26
Picos de Europa	Covadonga	8	1496	187
Sierra Nevada	Estación de esquí	32	597	19
Teide	Roques de García	18	1805	100
Timanfaya	Montañas del fuego	10	408	40
Taburiente	El Roque de los Muchachos	9	318	35
Garajonay	Roque de Agando	9	141	16
Archipiélago de Cabrera	Castillo de Cabrera	11	107	10
Islas Atlánticas de Galicia	Lago dos nenos - Isla Cíes	50	1380	28
Doñana	El Rocío - Charco de la boca	13	915	70

En resumen, los resultados obtenidos del análisis de autocorrelación espacial, tanto a nivel global como local, han desempeñado un papel fundamental en la identificación de los patrones espaciales de los visitantes en los parques nacionales de España. Esto proporciona una comprensión más profunda de cómo los visitantes interactúan con el entorno dentro de estos espacios protegidos. En particular, el análisis LISA ha demostrado ser una herramienta precisa para localizar las áreas de concentración de visitantes y su asociación con puntos de interés turístico y los equipamientos de uso público.

8.5. Explorando la localización óptima de instalaciones de uso público en los parques nacionales – una aplicación del modelo de ubicación-asignación con datos geolocalizados de Flickr.

Los modelos de localización son un conjunto de algoritmos utilizados principalmente en un sistema de información geográfica para determinar una ubicación óptima para una o más instalaciones que atenderán la demanda de un conjunto determinado de puntos. Los algoritmos pueden asignar esos puntos de demanda a una o más instalaciones, teniendo en cuenta factores como la cantidad de instalaciones disponibles, su costo y la impedancia máxima de las instalaciones a los puntos de demanda o viceversa. Por lo tanto, el objetivo de un modelo de ubicación y asignación es determinar la ubicación óptima para cada instalación y asignar los puntos de demanda para cada instalación, según los insumos de cada modelo (Comber et al., 2015). En esta sección se utilizaron los modelos de ubicación y asignación para determinar ubicaciones potenciales de centros de información adicionales en el parque nacional Teide, utilizando como proxy de la demanda las localizaciones de los usuarios de Flickr.

El modelo que se creó toma en cuenta las instalaciones de información existentes (centros de visitantes) y evalúa, en base a su localización, la localización de centros de información nuevos, de manera que los mismos puedan atender a la mayor demanda posible.

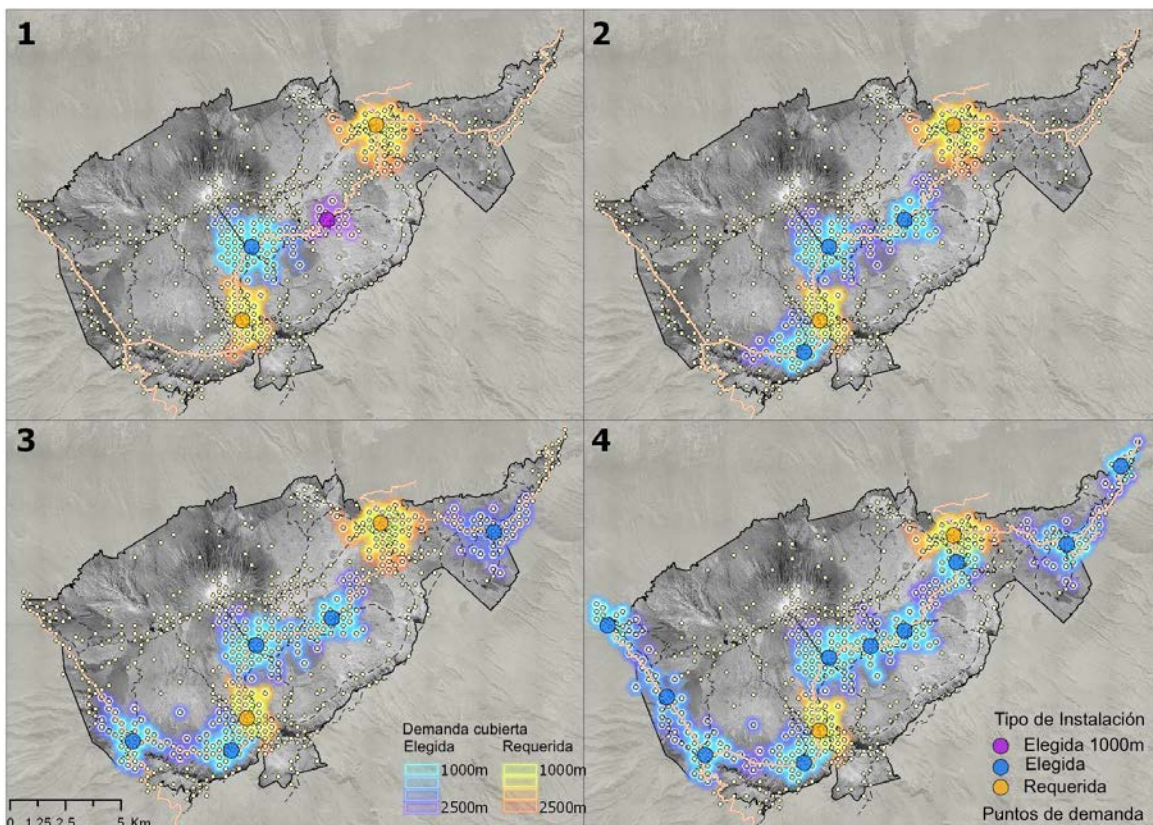
La aplicación propuesta consiste en calibrar y seleccionar un modelo de localización-asignación para calcular la ubicación óptima de nuevos puestos de información para los visitantes del parque nacional del Teide. El parque nacional del Teide ha registrado un número récord de visitantes durante los últimos años, siendo 4.443.628 las personas que acudieron en 2019, justo antes de la pandemia. Debido al incremento constante de la demanda, determinadas zonas del parque como los Roques de García han sobrepasado su capacidad de carga turística, en especial en la franja del mediodía (10:30 – 13:00), lo que conlleva a un deterioro de la experiencia del visitante y a la afectación de los recursos naturales de dichas zonas. Por ello se requiere de medidas que puedan liberar la presión sobre las zonas más concurridas y a su vez promover la visita de otros lugares menos frecuentados.

Aplicación del modelo de localización asignación - análisis de los resultados.

La **Figura 75** muestra las ubicaciones requeridas y elegidas para los cuatro escenarios propuestos. La demanda asignada a cada instalación está representada por el número de hexágonos cubiertos

en cada corte de distancia. En todos los escenarios, las ubicaciones de los centros de información están bien distribuidas en el terreno del parque y coinciden con las ubicaciones de los miradores existentes en el parque. La solución de maximizar la asistencia prioriza las ubicaciones donde los puntos de demanda con mayor peso se encuentran más cerca. En el primer escenario, cuando el modelo tiene que agregar solo una instalación adicional, se eligen dos ubicaciones óptimas diferentes en función de la distancia límite. Para un corte de distancia de 1000 m, la ubicación óptima para una nueva instalación se ubica en el mirador Minas San José, mientras que, para los tres cortes de distancia restantes, la ubicación elegida se ubica en la estación base del teleférico. En el resto de los escenarios, el modelo seleccionó las mismas ubicaciones para los distintos umbrales de distancia establecidos.

Figura 75. Resultados de los modelos de ubicación-asignación: (1) 1 establecimiento adicional, (2) 3 establecimientos adicionales, (3) 5 establecimientos adicionales, (4) 10 establecimientos adicionales.



La ubicación elegida que coincide con la estación base del teleférico parece reunir la mayor demanda potencial para todos los escenarios. Las instalaciones existentes (centros de visitantes El Portillo y Cañada Blanca) pueden cubrir entre el 23% (1000 m de distancia) y el 34% (2500 m) de la demanda potencial (Tabla 33).

Tabla 33. Demanda cubierta de 2 instalaciones existentes

Distancia de corte	Distancia media (m)	Demanda (# de usuarios)	Suma de demanda ponderada	% demanda cubierta por las instalaciones existentes
1000 m	616,63	372	314,83	23%
1500 m	937,37	454	405,17	29%
2000 m	1146,70	494	491,52	31%
2500 m	1306,54	538	559,06	34%

A medida que se asignan instalaciones adicionales, hay un incremento significativo en la demanda cubierta para todas las distancias de corte (Tabla 34). Así, por ejemplo, con la asignación de una sola instalación, la demanda cubierta aumenta del 23% al 36% (1000 m de distancia de corte). Solo cuando se suman 10 instalaciones y se establece la distancia de corte en 2500 m, la demanda estaría totalmente cubierta.

Tabla 34. Demanda cubierta de las instalaciones seleccionadas en todos los escenarios.

Escenarios	Distancia de corte	Distancia media (m)	Demanda (# de usuarios)	Suma de demanda ponderada	% demanda cubierta*	% demanda total cubierta **
1 instalación	1000	560,89	200	192,44	13%	36%
	1500	943,55	337	262,31	22%	51%
	2000	1152,86	404	334,64	26%	57%
	2500	1377,41	444	391,52	28%	62%
3 instalaciones	1000	580,20	493	507,11	31%	54%
	1500	866,87	580	636,16	37%	66%
	2000	1070,64	639	749,63	41%	72%
	2500	1280,18	699	854,41	45%	79%
5 instalaciones	1000	575,94	630	731,19	40%	63%
	1500	841,86	707	902,7	45%	74%
	2000	1080,68	799	1049,41	51%	82%
	2500	1307,86	878	1210,26	56%	90%
10 instalaciones	1000	561,27	840	1161,05	54%	77%
	1500	820,26	912	1425,04	58%	87%
	2000	1006,03	953	1645,14	61%	92%
	2500	1193,04	1038	1844,61	66%	100%

*Demanda cubierta por instalaciones nuevas

**Demanda total cubierta por instalaciones existentes e instalaciones nuevas.

Los resultados obtenidos de los modelos de ubicación – asignación mostraron que los datos de usuarios de Flickr podrían usarse como indicador de la demanda potencial de visitantes para ubicar nuevas instalaciones dentro del parque. Estos resultados pueden ser valiosos para la toma de decisiones y la optimización de recursos en la gestión del parque, ya que la infraestructura de visitantes es un componente clave para atraer visitantes.

Además, los escenarios desarrollados muestran una aplicación nueva para los datos geolocalizados de redes sociales que requiere de un mayor desarrollo y ajuste de los parámetros que influyen en la localización de infraestructura de uso público. Evidentemente, existen otras variables que se pueden utilizar para determinar la localización óptima de nuevas infraestructuras de uso público. Por ejemplo, se podría utilizar el componente temporal de los datos geolocalizados para identificar la demanda potencial en distintos momentos del día, filtrando periodos en los cuáles la afluencia es mayor para evaluar si las localizaciones seleccionadas por los modelos son las adecuadas para todos los momentos del día.

9. Los itinerarios en la red de parques nacionales españoles.

9. Los itinerarios en la red de parques nacionales españoles.

Los senderos son un recurso importante para la recreación y el turismo en los espacios protegidos (Mckercher et al., 2012). Permiten el acceso de los visitantes a los principales puntos de interés y ayudan a preservar los ecosistemas protegidos al restringir la dispersión de los visitantes. Sin embargo, el uso de dicha infraestructura por parte de los visitantes tiene como resultado impactos directos e indirectos sobre los recursos naturales. Por ejemplo, la erosión del suelo y la pérdida de vegetación se encuentran entre los impactos típicos del uso intensivo de senderos. El aumento del uso de visitantes también puede afectar las experiencias de los visitantes, ya que los problemas de hacinamiento, la seguridad de los visitantes y la disminución de la calidad de la experiencia se convierten en un problema real.

Desde la perspectiva de la gestión, es importante medir y monitorear el uso y el comportamiento de los visitantes, en las redes de senderos. Esta información es fundamental para la labor de planificación y gestión de los parques nacionales. La forma convencional de medir la intensidad de uso es a través de encuestas y contadores.

Los avances en la tecnología móvil y de posicionamiento global han hecho posible la producción de grandes cantidades de datos del movimiento de personas y objetos provenientes de dispositivos como receptores GPS, teléfonos móviles, vehículos entre otros (Shoval & Ahas, 2016). La difusión de dispositivos GPS abre oportunidades para recopilar datos precisos y completos sobre cómo se mueven los visitantes y sus itinerarios. Los datos derivados de dichos dispositivos representan la trayectoria de un objeto/individuo en movimiento. Esta trayectoria se puede definir como la secuencia ordenada cronológicamente de ubicaciones espaciales que se miden en 2 dimensiones (longitud y latitud). Además de las coordenadas y el tiempo, cada ubicación puede contener información asociada como la altitud, identificador del punto, entre otros.

Una fuente importante de este tipo de datos tanto por su volumen como disponibilidad son las redes sociales y en particular las redes sociales especializadas en recreación y deportes. En este sentido, el objetivo de esta investigación es utilizar datos de rutas GPS, provenientes de Wikiloc para identificar pautas de movilidad de los visitantes de espacios naturales protegidos. Wikiloc es un sitio web que permite almacenar y compartir trayectorias GPS de distintas actividades. Similar a otras redes como Strava, Alltrails, cuenta con una aplicación móvil.

Las rutas GPS tienen ventajas para estudiar el movimiento de visitantes dentro de los senderos debido a que representan el recorrido que realizó cada usuario al momento de su visita al parque nacional. Además, la gran cantidad de datos que cargan los usuarios brinda oportunidades para

analizar el comportamiento de los visitantes en lugares donde los datos son escasos o inexistentes, como es el caso de los parques nacionales, lugares remotos y extensos que no cuentan con monitoreo de todos los caminos e infraestructura.

En esta sección se utilizan las rutas GPS descargadas de la red social Wikiloc de seis parques nacionales; Archipiélago de Cabrera, Aigüestortes, Caldera de Taburiente, Tablas de Daimiel, Monfragüe e Islas Atlánticas de Galicia, para: 1) Cuantificar el uso de la red de senderos distinguiendo entre dos actividades: senderismo y ciclismo. 2). Mapear los patrones espaciales de las actividades de senderismo y ciclismo y 3) Combinar los datos de Flickr y Wikiloc para calcular y mapear los flujos de visitantes desde los puntos de acceso a un parque nacional utilizando el parque nacional Aigüestortes como área de estudio.

9.1. Metodología para el análisis del uso de las redes de senderos y la distribución de los flujos.

Una vez que los datos descargados de Wikiloc, obtenidos mediante descarga manual del sitio web, fueron depurados, se procedió a realizar un análisis con el propósito de evaluar la intensidad de uso de los senderos dentro de los parques nacionales.

Para conseguir este objetivo, se siguieron una serie de procesos que inician con la transformación de los puntos GPS en capas de geometría lineal que dibujan la ruta seguida por los usuarios en el territorio. Al visualizar las rutas obtenidas, fue evidente que las rutas lineales no se ajustan o alinean con las redes de senderos de los parques. El contar con rutas ajustadas es necesario porque permite la aplicación de herramientas de análisis espacial como el análisis de redes de transporte. En este sentido, el siguiente proceso a aplicar fue el realizar el ajuste de las rutas lineales usando un modelo de map matching.

Finalmente, con las rutas ajustadas, se realizó el cálculo de intensidad de uso de los senderos por tramo y la visualización de los resultados en mapas de densidad que resaltan los tramos de senderos más utilizados y aquellos senderos subutilizados.

A continuación, se explican con mayor detalle los procesos aplicados.

Map Matching - Ajuste de rutas a las redes de senderos.

El Map Matching es un procedimiento en el cual se asignan objetos geográficos a ubicaciones en un mapa. Los objetos geográficos más típicos son posiciones puntuales obtenidas de un sistema de

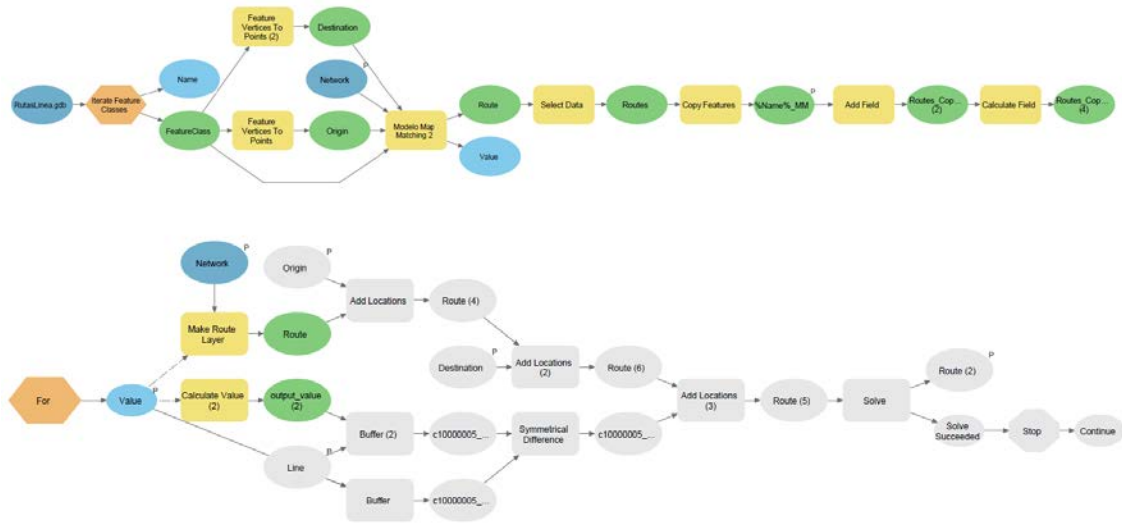
posicionamiento, a menudo un receptor GPS. En usos típicos, las posiciones de GPS se derivan de un receptor ubicado en un vehículo u otro objeto en movimiento que viaja en una red de carreteras, y el mapa digital modela la incrustación en el espacio geográfico de las carreteras por medio de polilíneas que se aproximan a las líneas centrales de las carreteras. Las posiciones GPS generalmente no se cruzan con las polilíneas, debido a imprecisiones. El objetivo del proceso de map matching es colocar las posiciones GPS en sus ubicaciones “correctas” en las polilíneas asociadas a los senderos de cada parque.

Aplicamos una herramienta de map matching basada en la implementación del algoritmo desarrollado por (Dalumpines & Scott, 2011) en el entorno de ArcGIS Pro. El procedimiento crea un búfer alrededor de la ruta GPS que restringe la estimación de la ruta más corta entre el origen y el destino, utilizando el algoritmo de camino más corto de Dijkstra (Cherkassky et al., 1994). En este proceso, la definición de las distancias del búfer determina los resultados: un búfer demasiado pequeño puede impedir la coincidencia de muchas rutas, mientras que un búfer demasiado ancho puede conducir a rutas inexactas o incorrectas (Figura 76). La herramienta basada en GIS utiliza herramientas de análisis de red de la extensión de análisis de red y de geoprocésamiento de ArcGIS para generar la ruta más corta entre el origen y el destino de la ruta GPS a lo largo de la red de senderos creada para cada parque. La ruta resultante es el track GPS ajustado a la red (Romanillos & Gutiérrez, 2020).

Figura 76. Ejemplo ilustrativo del proceso de map matching o ajuste de rutas



Figura 77. Herramienta para implementar algoritmo de map matching utilizando Model Builder de ArcGIS pro

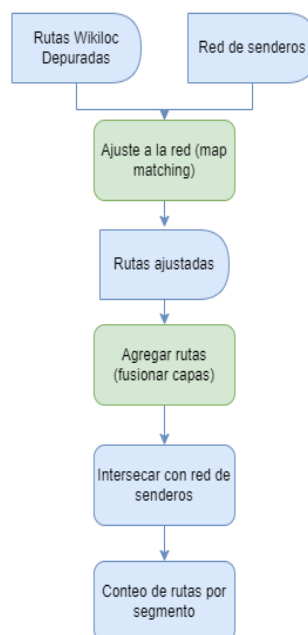


Cálculo de la intensidad de uso de las redes de senderos.

Tras el ajuste de las rutas de Wikiloc conforme a la cartografía específica de la red de senderos en cada parque, se procedió a la consolidación de estas rutas en una única capa geográfica. A continuación, esta capa fue superpuesta con la red de senderos existente, permitiendo así calcular el número de rutas por cada segmento de sendero. Este proceso generó un indicador crucial para evaluar el uso de la red de senderos según las actividades de senderismo y ciclismo.

El siguiente diagrama representa el flujo de trabajo implementado para derivar la intensidad de uso de las redes de senderos.

Figura 78. Diagrama de flujo de trabajo para obtención de densidades de uso de senderos



Cálculo de los flujos de visitantes desde los accesos a los parques nacionales.

Para estimar los flujos de visitantes a partir de los datos geolocalizados de Flickr y Wikiloc se llevó a cabo el siguiente flujo de procesos:

- Preparación de los datos: Se llevó a cabo la limpieza y tratamiento de los datos de ambas fuentes para eliminar registros inconsistentes y prepararlos para el análisis.

En el caso de Flickr, se filtraron los datos para seleccionar usuarios que tomaron al menos tres fotografías en un mismo día y que estas correspondieran a localizaciones diferentes, lo que resultó en 431 fotos de 53 usuarios.

En cuanto a las rutas GPS de Wikiloc, los puntos obtenidos de la descarga manual se transformaron en capas de líneas y se calcularon los parámetros de velocidad, duración y distancia entre cada par de puntos GPS. Estos parámetros son necesarios para el análisis posterior. Además, se eliminaron inconsistencias, como puntos GPS que registraban una velocidad superior a la máxima permitida para senderistas y ciclistas (10 km/h y 20 km/h, respectivamente).

- Construcción de los itinerarios:

Para los itinerarios de Flickr y Wikiloc, se ajustaron las rutas de los usuarios a la red de senderos del parque utilizando la herramienta track analysis de ArcMap 10.6 y aplicando el proceso de map matching, respectivamente.

- Mapas de flujo de accesos:

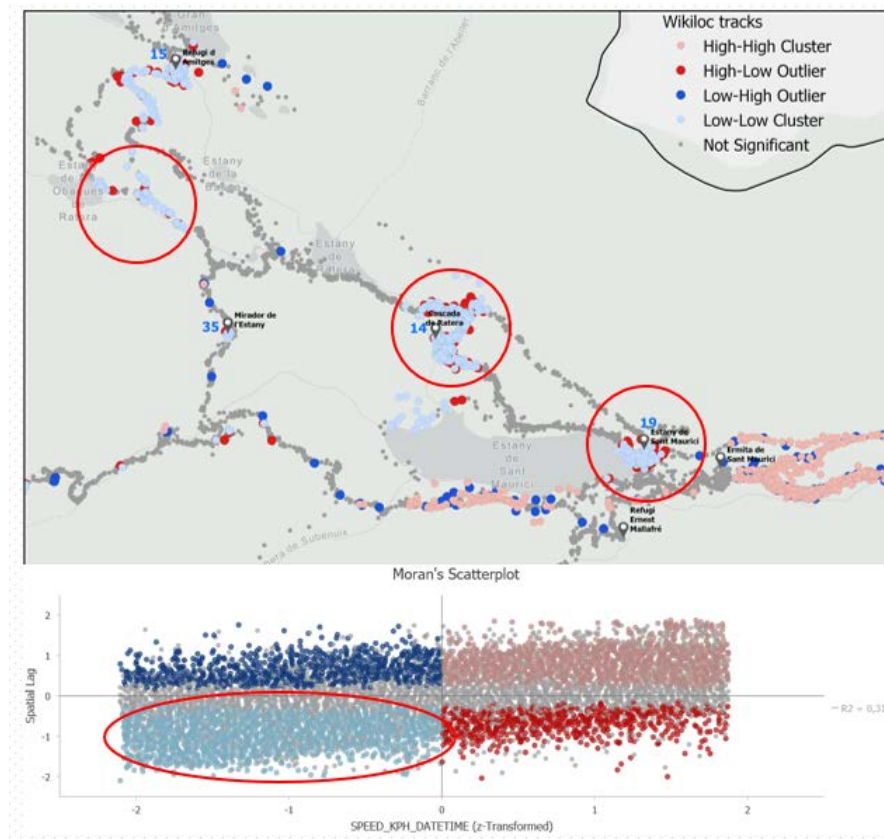
El proceso para mapear el flujo de los visitantes desde los accesos involucró cuatro pasos:

- Se asignó el acceso de cada ruta utilizando la herramienta de ubicación de instalación más cercana, asumiendo que dicho acceso correspondía al que el visitante utilizó para ingresar al parque. Los puntos de inicio de cada ruta fueron tratados como incidentes, mientras que las localizaciones de los accesos se consideraron como instalaciones.

- Se establecieron los puntos de interés donde los visitantes realizaron paradas. Para Flickr, cada fotografía se consideró un punto de parada, mientras que para Wikiloc, se aplicó el índice local de autocorrelación espacial de Moran usando como variable la velocidad de cada punto para identificar puntos de interés donde la mayoría de los visitantes se detuvieron (clústeres LL), es decir los clústeres donde las velocidades eran bajas y se concentraban los puntos de las rutas se asumieron como puntos de parada para los datos de Wikiloc.

- Se reconstruyeron los itinerarios específicos para cada actividad a partir de los puntos de acceso y paradas detectados en los pasos anteriores y usando la herramienta de análisis de rutas.
- Se contó el número de segmentos de sendero utilizados por cada visitante desde el acceso hasta los puntos de interés. Con base en estos resultados, se construyó un mapa de flujo que representaba el análisis del primer punto donde los visitantes comenzaron su recorrido y las paradas que realizaron durante sus recorridos.

Figura 79. Ejemplo de detección de paradas en las rutas GPS a partir del índice local de Moran.



9.2. ¿Existen diferencias importantes entre el movimiento de visitantes que practican senderismo y ciclismo?

La duración promedio de los recorridos de senderismo se encuentra entre los 128 y 330 minutos (Tabla 35), siendo el Archipiélago de Cabrera el parque con las rutas de senderismo que registraron el menor tiempo de duración. En el extremo opuesto, se encuentra el parque nacional Aigüestortes, que registró las rutas con mayor tiempo de duración.

Al igual que en senderismo, Aigüestortes registró la mayor duración promedio de las rutas de ciclismo con 275 minutos, en tanto que en las Islas Atlánticas de Galicia, la isla de Ons, tiene la menor duración media de las rutas de ciclismo Tabla 36.

La velocidad media de los recorridos de senderismo estuvo entre 3,67 y 5,09 km/h. Con respecto a las rutas de ciclismo, la duración de los recorridos presentó valores similares a las rutas de senderismo, registrando una duración entre 123 y 275 minutos.

Como se esperaba, la velocidad promedio de los recorridos de ciclismo, si presentó una diferencia notable y los valores de velocidad fueron superiores encontrándose entre los 7 y 18 km/h. En este sentido, los parques con un relieve plano como Monfragüe y Tablas de Daimiel registraron las mayores velocidades promedio, en tanto que el parque de montaña Aigüestortes registró la velocidad promedio más baja. La velocidad promedio de los recorridos ciclistas en los parques estudiados es de 11,1 km/h, destacando el parque Monfragüe como el espacio donde se registran las rutas con mayor velocidad.

En los parques de Aigüestortes, Caldera de Taburiente y Monfragüe, los recorridos de senderismo presentaron la mayor distancia media (>13 km), pudiendo indicar que los visitantes de estos espacios tienden a realizar itinerarios más largos. La distancia promedio de los recorridos en bicicleta en los 5 parques nacionales fluctúa entre 9 km en las Islas Atlánticas de Galicia y 45.5 km en Monfragüe.

Tabla 35. Estadísticas rutas senderismo.

Parque	# Rutas	Distancia total (Km)	Distancia promedio (Km)	Duración media (minutos)	Velocidad media (Km/h)	# Puntos	
Archipiélago de Cabrera	137	833,5	6,08	128,6	4,92	23837	
Aigüestortes	433	5462,7	13,02	329,9	3,67	88007	
Caldera de Taburiente	216	3063,1	14,24	324,7	4,24	56525	
Tablas de Daimiel	164	1087,4	6,63	130,8	4,55	27443	
Monfragüe	152	1631,4	10,73	241,5	3,65	23009	
Islas Atlánticas de Galicia	Cíes	608	4957,3	8,15	185,3	4,40	104659
	Ons	315	2598,8	8,25	159,5	4,28	48894
	Sálvora	21	113,9	5,43	135,7	5,09	3847
	Cortegada	15	50,5	3,36	108,8	3,79	4701

Tabla 36. Estadísticas rutas ciclismo

Parque	# Rutas	Distancia total (Km)	Distancia promedio (Km)	Duración media (minutos)	Velocidad media (Km/h)	# Puntos	
Aigüestortes	238	3782,7	15,89	275,9	9,08	107145	
Caldera de Taburiente	31	469,2	15,14	235,6	8,22	7396	
Tablas de Daimiel	29	922,3	31,80	145,5	14,70	6315	
Monfragüe	130	5916,8	45,51	243,8	18,47	41170	
Islas	Cíes	20	181,4	9,07	183,3	6,84	2789

Parque		# Rutas	Distancia total (Km)	Distancia promedio (Km)	Duración media (minutos)	Velocidad media (Km/h)	# Puntos
Atlánticas de Galicia	Ons	37	360,4	9,74	123,7	9,27	5993

Uso de las redes de senderos de los parques nacionales.

En esta sección se exponen los resultados derivados del cálculo de la intensidad de uso de las redes de senderos, basados en los datos de Wikiloc. El análisis ha sido segmentado de acuerdo con las actividades documentadas en las rutas: senderismo y ciclismo. Para brindar un entendimiento más detallado de cada parque, se encuentra disponible en la sección de anexos (apartado 2) la descripción de las redes de senderos e itinerarios propios de cada parque. Estas capas de información se utilizaron como insumos para la evaluación de la intensidad de uso, tal como se describe en el siguiente análisis.

Senderismo.

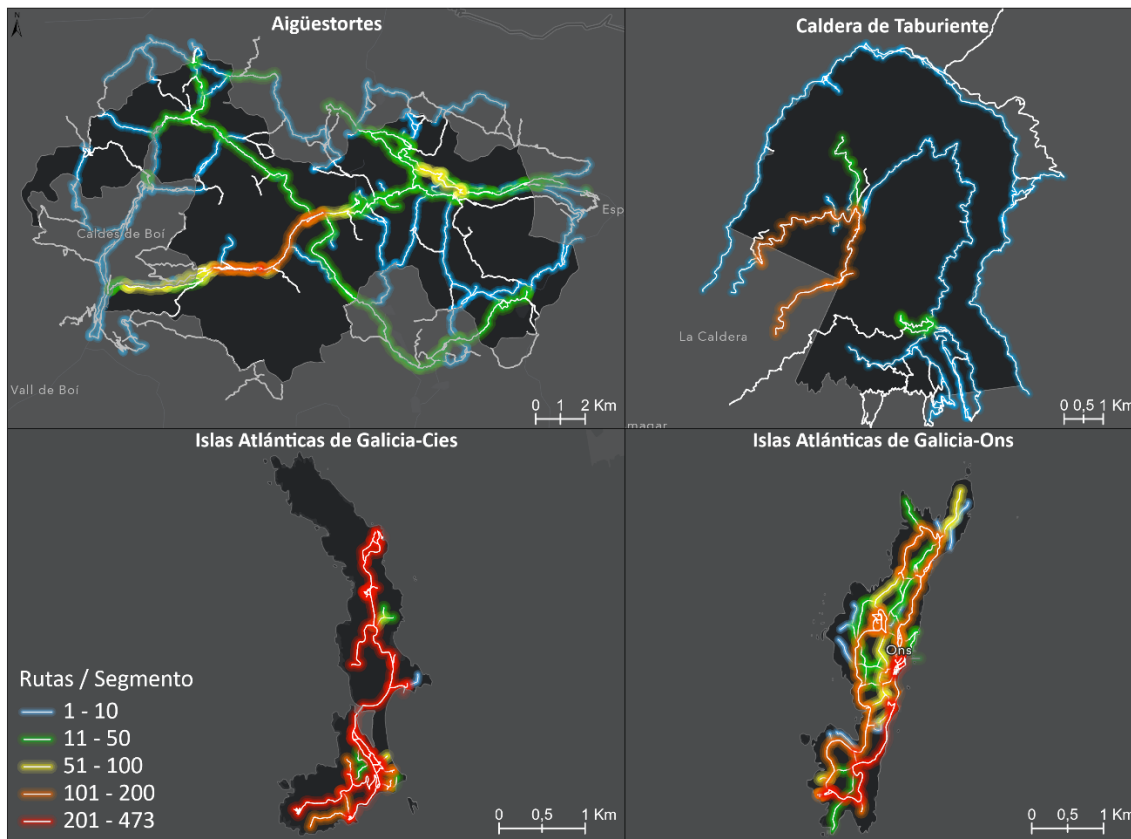
En el contexto del senderismo en los parques nacionales examinados, se destacan variaciones en la utilización de las redes de senderos y que se evidencian en los mapas de densidad de rutas por segmento de sendero (**Figuras 79 y 80**). Se identifican dos grupos: uno con densidades más altas (hasta 473 rutas por segmento) y otro con densidades moderadas (hasta 150 rutas por segmento).

El primer grupo abarca a los parques de Aigüestortes, Caldera de Taburiente y las Islas Atlánticas de Galicia. En todos estos parques, se observa un uso extendido de los senderos, sobresaliendo en este aspecto Islas Atlánticas de Galicia, donde todos los senderos tienen un uso significativo.

En el parque Aigüestortes, el senderismo se concentra en rutas centrales que conducen a puntos clave como los lagos Estany de Sant Maurici, Estany Llong y Estany de Llebre. Por otro lado, en Caldera de Taburiente, la actividad es mayor en el centro y destaca en itinerarios como Brecitos y Barranco de las Angustias, que llevan a zonas de acampada.

Por último, en Islas Atlánticas de Galicia se aprecian dos patrones de actividad en senderismo. En la isla Cíes, los senderos son recorridos en su totalidad, con énfasis en itinerarios del OAPN como Monte Faro y Faro da Porta, con más de 400 rutas por sendero. En la isla Ons, el itinerario Ruta Sur es el más frecuentado.

Figura 80. Mapas de densidad de uso de la actividad de senderismo (Grupo 1).



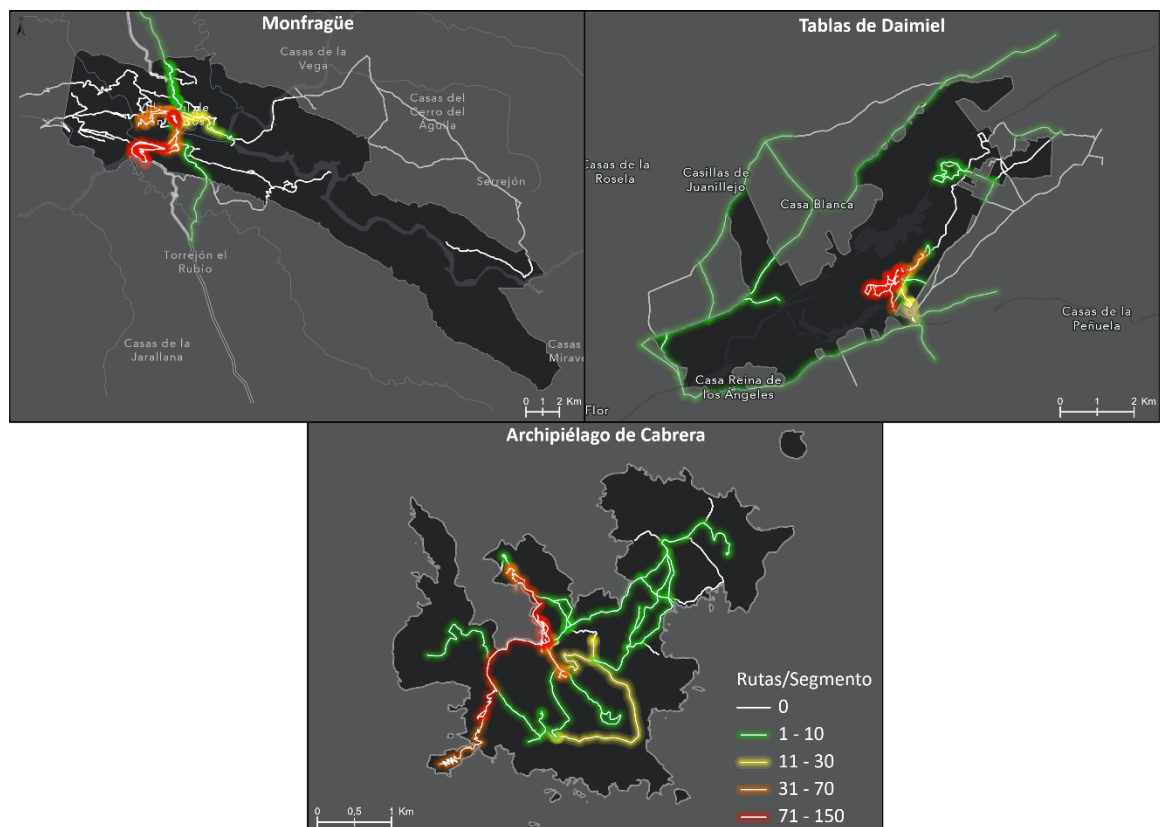
En el segundo grupo, que incluye a los parques de Monfragüe, Tablas de Daimiel y Archipiélago de Cabrera, se destaca una densidad de uso moderada en la actividad de senderismo. Sin embargo, es evidente que no todos los senderos han experimentado la misma intensidad de uso, siendo este fenómeno más marcado en el caso del parque de Monfragüe, donde se observa que varios senderos no han sido utilizados.

En Monfragüe, los senderos que atraviesan el centro del parque han registrado una densidad significativa, superando las 70 rutas por segmento. Estos senderos conducen a los principales atractivos, como el emblemático Castillo de Monfragüe. En contraste, los senderos menos transitados corresponden a los itinerarios de Puerto de la Serrana y Torrejón El Rubio, situados en la parte septentrional del parque.

En el caso del Parque Tablas de Daimiel, los senderos más concurridos se encuentran en el itinerario de La Isla del Pan, donde se localizan instalaciones de uso público como el observatorio de la Isla de Pan, generando una densidad entre 71 y 105 rutas por segmento. En contraste, los senderos menos frecuentados abarcan las áreas periféricas del parque, siendo un ejemplo el sendero correspondiente al itinerario de la Torre de Prado Ancho.

Finalmente, en el Parque Archipiélago de Cabrera, se observa un patrón de uso desigual de la red de senderos. Los visitantes tienen una preferencia marcada por los senderos adyacentes al puerto y la playa, lo que resulta en una concentración de la actividad de senderismo en el centro de la isla de Cabrera, donde se ubican la mayoría de las instalaciones de uso público. Se destacan los senderos de los itinerarios que conducen al Faro de L'Encionla. En contraste, el sector noroeste de la isla presenta una menor afluencia, siendo el itinerario de Set Siges, que abarca el norte de la isla, el que registra el menor número de rutas por segmento.

Figura 81. Mapas de densidad de uso de la actividad de senderismo (Grupo 2)



Ciclismo.

El primer grupo incluye los parques de Aigüestortes, Monfragüe y Tablas de Daimiel. En esta categoría es notable la mayor extensión de uso de la red de senderos para la actividad de ciclismo en comparación con el senderismo en los parques Monfragüe y Tablas de Daimiel.

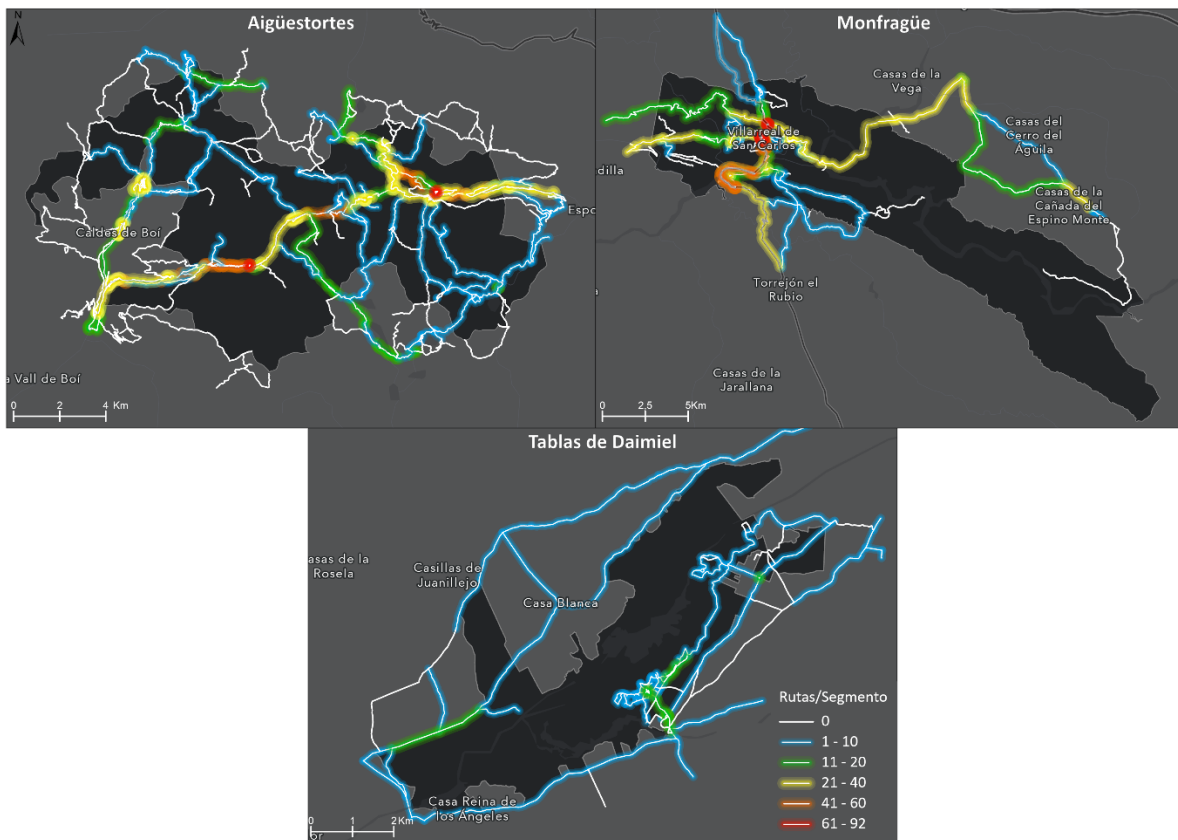
En Aigüestortes, la actividad ciclística se distingue por presentar un volumen menor de rutas en comparación con el senderismo, lo cual se refleja también en el uso de menos ejes dentro de la red de senderos. En esta dinámica, las rutas de ciclismo priorizaron los senderos centrales del parque, siendo especialmente sobresaliente el sendero que conduce al Estany de Cavallers. Cabe resaltar que estos senderos exhibieron densidades superiores en comparación con el senderismo,

indicando una concentración más enfocada de la actividad de ciclismo. El itinerario más concurrido para el ciclismo es el de las Agulles Damitges (entre 665 y 99 rutas), ubicado en el suroeste del parque.

En Monfragüe, los ciclistas aprovecharon la mayoría de los senderos, representando el 85% de la red total. Los senderos más solicitados en esta actividad se distribuyen en todo el territorio del parque, destacando en el corazón del mismo los senderos del puente del Tajo y el camino al aula de naturaleza de los Chozos. La actividad ciclística abarcó todos los itinerarios propuestos por el OAPN, con especial énfasis en los recorridos al Castillo de Monfragüe, Serradilla por la pista de Barbaon y Arroyo Malvecino.

Finalmente, en Tablas de Daimiel, se observó una mayor amplitud de senderos utilizados para la actividad de ciclismo. No obstante, en términos de cantidad de usuarios, se notó una menor intensidad en la utilización de la red para el ciclismo. En este contexto, los senderos en los límites del parque tuvieron mayor afluencia para esta actividad. Los itinerarios de la Laguna Permanente y La torre de Prado Ancho destacaron como los más utilizados en ciclismo (entre 11 y 30 rutas).

Figura 82. Mapas de densidad de uso de la actividad de ciclismo (Grupo 1).

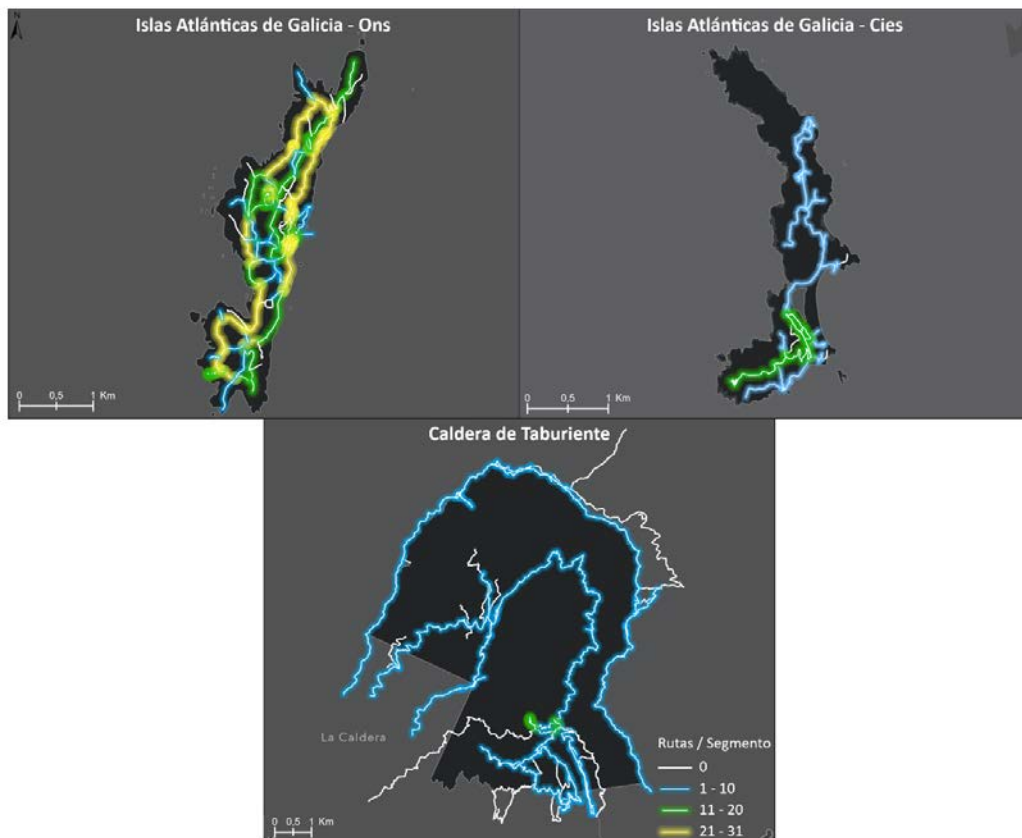


Dentro del segundo grupo de parques, encontramos a Islas Atlánticas de Galicia y Caldera de Taburiente, que muestran una baja densidad de uso en la actividad de ciclismo, con menos de 30 rutas por segmento.

En Islas Atlánticas de Galicia, la práctica del ciclismo se manifiesta con menor intensidad en comparación con el senderismo. Sin embargo, es notable que los visitantes aprovecharon la mayoría de los senderos de la red. En este parque, se observan dos patrones distintos: En la isla Cíes, la actividad de ciclismo presentó una intensidad de uso más reducida (entre 11 y 20 rutas por segmento). Los senderos más transitados pertenecen al itinerario Ruta Monte Faro, mientras que los menos concurridos corresponden al itinerario Ruta de Monteagudo, situados al norte de la isla. Por otro lado, en la isla de Ons, la actividad de ciclismo mostró una intensidad de uso más alta en toda la red de senderos, siendo los senderos más populares los ubicados en el centro de la isla, cerca del centro de visitantes.

En Caldera de Taburiente, la actividad de ciclismo también se desarrolla con menor intensidad en comparación con el senderismo. Esto es evidenciado por la densidad de uso en la mayoría de los senderos, que fue inferior a 10 senderos por segmento. Los senderos preferidos por los ciclistas incluyen los itinerarios de Brecitos y partes del itinerario del Lomo de las Chozas.

Figura 83. Mapas de densidad de uso de la actividad de ciclismo (Grupo 2).



9.3. Flujos de visitantes y movimiento – una aplicación del análisis de redes con rutas GPS de Wikiloc y Flickr.

La movilidad es un concepto inseparable de la actividad turística. Para el turismo de naturaleza, comprender la movilidad de los visitantes y los flujos que generan dentro del espacio protegido es fundamental para la gestión de los impactos derivados de la actividad turística y la planificación de actividades e itinerarios que permitan al visitante tener una experiencia plena. Sin embargo, para poder analizar la movilidad turística es necesario contar con información de alta granularidad espacial y temporal. En este sentido los datos geolocalizados de redes sociales, que son generados por los mismos visitantes presentan oportunidades para analizar la movilidad de los visitantes dentro de los parques nacionales.

En este apartado se utilizan los geolocalizados de Flickr y Wikiloc del parque nacional de Aigüestortes i Estany de Sant Maurici para identificar la afluencia de visitantes desde los puntos de acceso del parque y representar los flujos internos de visitantes.

La elección de Flickr y Wikiloc como fuentes de datos se basó en que representan a diferentes tipos de turistas. Por ejemplo, los usuarios de Flickr, que publican fotos de sus lugares y experiencias más memorables, reflejan el atractivo turístico de la zona. Por otro lado, los usuarios de Wikiloc, que comparten sus rutas GPS junto con la actividad realizada, representan a los turistas de aventura o interesados en actividades deportivas. Esta diferenciación nos proporcionó indicadores valiosos sobre la demanda y el comportamiento turístico en espacios naturales. Los resultados obtenidos reflejan el comportamiento diferenciado de dos tipos de turistas que permiten contar con indicadores potentes de la demanda y comportamiento turísticos en espacios naturales.

Características de los itinerarios.

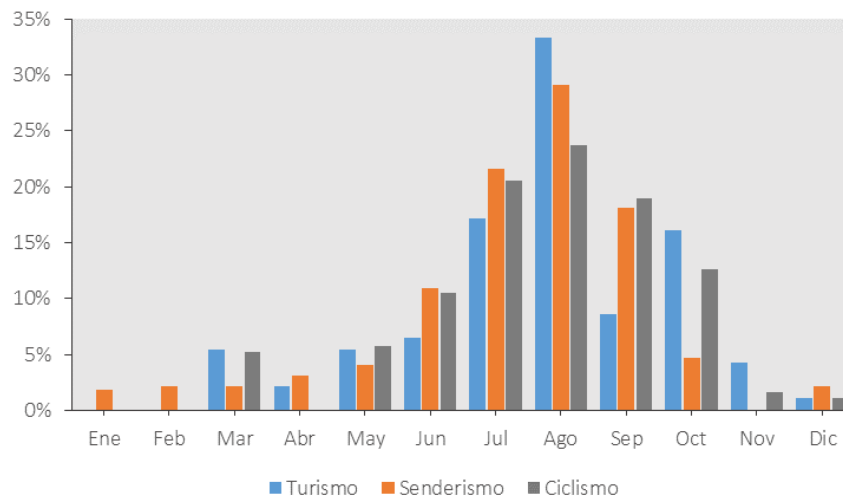
Un total de 93 usuarios de Flickr cumplieron con los criterios establecidos, que requerían un mínimo de tres fotografías capturadas en un día y desde distintas ubicaciones. Utilizando estos datos, se generaron 93 itinerarios que representan la actividad turística. En el caso de Wikiloc, se obtuvieron 320 rutas para el análisis de la actividad de senderismo y 190 rutas para el análisis de ciclismo (ver **Tabla 37**). Al comparar las características de ambas fuentes de datos, se destacan algunas diferencias interesantes. Por ejemplo, la velocidad promedio de los visitantes registrados en Flickr es menor que la de los usuarios de Wikiloc. Asimismo, los visitantes de Wikiloc recorrieron distancias más extensas en comparación con los usuarios de Flickr.

Tabla 37. Características del movimiento de los itinerarios de Flickr y Wikiloc.

Fuente de datos	Número de rutas	Distancia (Km)			Duración (min)			Velocidad (Km/h)
		Media	Min	Max	Media	Min	Max	Media
Flickr Turismo	93	3,6	0,5	14	215	8	643	1,69
Wikiloc Senderismo	320	10	1,1	24	337	22	718	3,33
Wikiloc Ciclismo	190	12	0,8	51	296	15	960	16,71

El análisis de las rutas por mes revela un patrón estacional de visitas que se alinea consistentemente con las cifras oficiales de visitantes. Es notorio que la temporada de verano se destaca como el período de mayor afluencia en todas las actividades que hemos observado. Este incremento es coherente con unas temperaturas más altas y condiciones más propicias para el turismo.

Figura 84. Distribución mensual de los itinerarios según actividad.



Mapas de flujos desde accesos

Se empleó el índice de autocorrelación espacial de Moran, en los datos de Wikiloc, para identificar las ubicaciones donde los visitantes de senderismo y ciclismo realizaron paradas a lo largo de sus rutas. Los resultados de este índice señalaron que el 5,4% de los puntos GPS correspondían a puntos donde se registró una interrupción en el movimiento, es decir, clústeres de valores bajos (low-low values). A partir de estos clústeres, se lograron identificar 46 agrupaciones que pueden relacionarse con puntos de interés o equipamientos dentro del parque. En este contexto, se encontró que los lugares más visitados tanto por senderistas como por ciclistas correspondían a atractivos naturales emblemáticos del parque, como el lago Sant Maurici, la Cascada de Ratera y el mirador de Sant Esperit (Tabla 38 y 39).

Tabla 38. Frecuencia de paradas o puntos de interés en los Itinerarios de senderismo (Wikiloc).

Nombre	Descripción	Frecuencia
Sant Esperit	Mirador	141
Sant Maurici	Lago	121
Llong	Lago	116
Sant Esperit	Cascada	79
Llebrea	Lago	77
Ratera	Cascada	77
Portarró	Lago	61
Espol	Punto de información	57
d'Amitges	Refugio	54
Boi	Punto de información	49
Portarró	Punto de información	46
Ventosa i Calvell	Lago	38
Port de Ratera	Cima	29
Dellui	Lago	28
Tort de Pequera	Lago	28
Vall Fosca	Teleférico	26
Llastra	Lago	22
Dellui	Lago	21
Colomers	Refugio	21
Tart	Lago	20
Crestada	POI	19
Saburó	Lago	14
Cap de Port	Lago	13
Portarró	Mirador	12
Port de Colomers	POI	10
Monestero	POI	10
Gran de Pequera	Lago	9
Portarró	Cima	8

Tabla 39. Frecuencia de paradas o puntos de interés en los Itinerarios de ciclismo (Wikiloc)

Nombre	Descripción	Frecuencia
Sant Maurici	Lago	57
Cascada de Ratera	Cascada	38
Refugi Estany Llong	Refugio	38
Cabana del Pont de Paller	Cabaña	36
Espol	Punto de información	34
Sant Esperit	Mirador	29
Refugi d'Amitges	Refugio	29
Font del Planell Gran	Fuente	23
Portarró	Punto de información	20
Boi	Punto de información	19
Estanyol del Portarró	Lago	15
Cabana del Magre	Lago	14
Cascada de Sant Esperit	Cascada	14
Estany de Dellui	Lago	13
Estany Negre de Pequera	Lago	11
Estany Tart	Lago	10
Refugi Colomina	Refugio	10
Refugi Ventosa i Calvell	Refugio	10
Dellui	Lago	9
Saburó	Lago	8
Estany de la Llastra	Lago	8
Port de Ratera de Colomers	POI	8

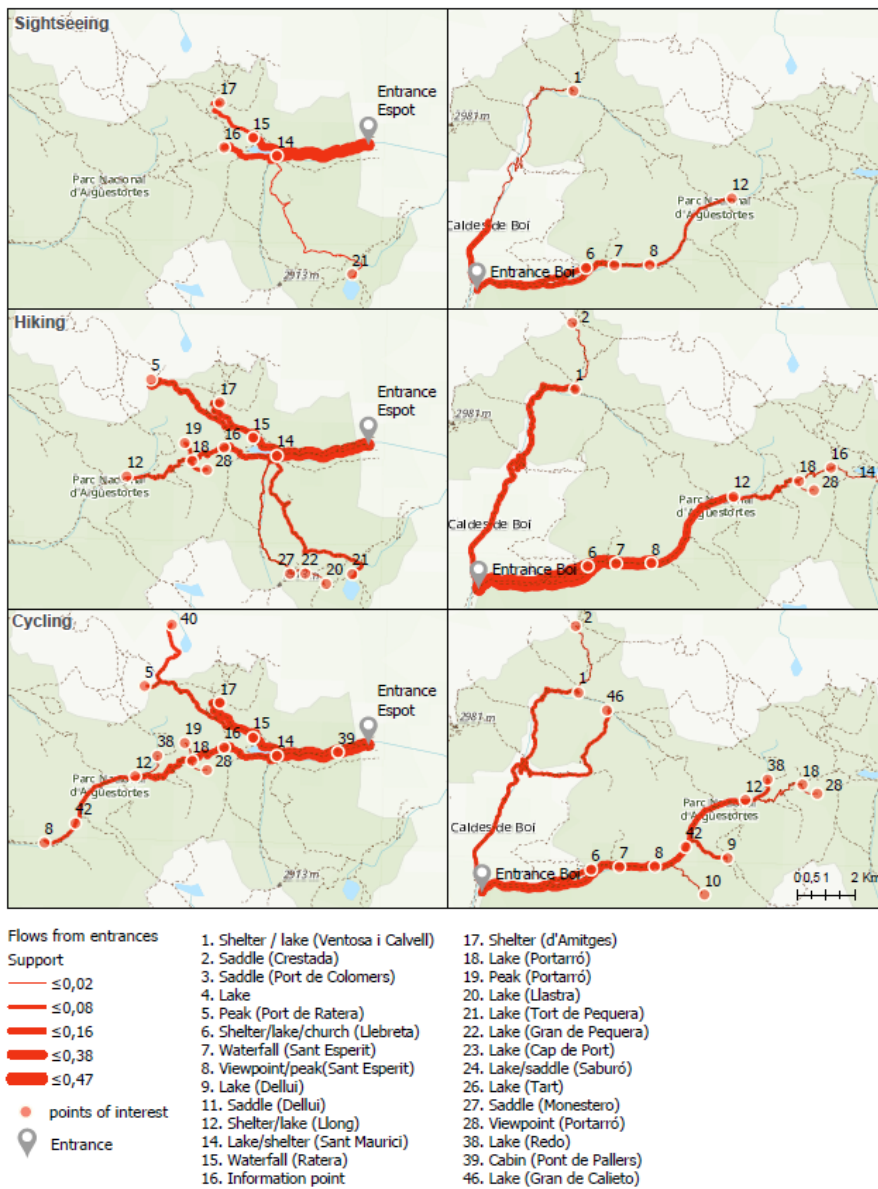
Con respecto a los visitantes captados con la red social Flickr, las paradas realizadas durante los itinerarios corresponden con las localizaciones de las fotografías. Al igual que con los visitantes de senderismo y ciclismo, los lugares más visitados fueron el lago Sant Maurici y la cascada Ratera (Tabla 40). Entre los equipamientos más frecuentados en todas las actividades se destacaron los refugios y puntos de información. Específicamente, los refugios del Estany Llong y D' Atmigés, junto con los puntos de información de Espot y Boi, captaron un notable interés por parte de los visitantes.

Tabla 40. Frecuencia de paradas o puntos de interés en los Itinerarios de turistas (Flickr).

Nombre	Descripción	Frecuencia
Sant Maurici	Lago	26
Ratera	Cascada	26
Portarró	Lago	17
D'Atmigés	Refugio	14
Espot	Punto de información	12
Boi	Punto de información	12
Llebreta	Refugio	11
Sant Esperit	Cascada	11
Sant Esperit	Mirador	10
Llong	Lago	9
Ventosa i Calvell	Lago	8
Tort de Pequera	Lago	8

El análisis de los flujos de visitantes desde los accesos a los atractivos naturales y equipamientos reveló la importancia de la entrada desde la localidad de Espot para las actividades de turismo y ciclismo. En contraste, en el caso de la actividad de senderismo, la entrada desde Boi fue la más utilizada para iniciar la exploración del parque (ver Figura 86). Sin embargo, se pudo observar que, una vez dentro del parque, los visitantes de todas las actividades tuvieron la oportunidad de recorrer y acceder a múltiples atracciones. En cuanto al volumen de los flujos desde los puntos de acceso, se notó que los valores más altos correspondían a las áreas cercanas a ambos puntos de entrada. Esto sugiere que la mayoría de los usuarios optaron por realizar recorridos cortos y escogieron visitar las atracciones más próximas a su punto de partida. Por ejemplo, desde el acceso en Espot, los lugares más visitados en todas las actividades fueron el Lago Sant Murici y la cascada Ratera, que se hallan cercanos a dicho acceso. Del mismo modo, desde el acceso de Boi, los lugares de mayor afluencia fueron aquellos que se encontraban en las proximidades, como el refugio Llebre y la cascada de Sant Esperit.

Figura 85. Mapa de flujos de visitantes desde puntos de acceso para las actividades de turismo (Flickr), Senderismo (Wikiloc) y Ciclismo (Wikiloc).



En última instancia, el análisis de los flujos de visitantes desde los accesos a las atracciones y equipamientos permitió explorar la relevancia de los puntos de acceso Espot y Boi para tres actividades. Estos resultados proporcionan un conocimiento valioso para la gestión ya que se podría utilizar para optimizar la gestión del acceso a los parques nacionales.

10. Conclusiones y futuras líneas de investigación.

10. Conclusiones y futuras líneas de investigación.

En esta tesis se plantearon siete preguntas que guían la investigación y que se explicaron en el capítulo introductorio. Además, se definieron objetivos relacionados a responder estas preguntas. En esta sección, se presentan las conclusiones derivadas de la investigación, estructuradas en torno a las preguntas de investigación.

10.1. Conclusiones.

PI.1: ¿Por qué es relevante estudiar el uso turístico de los espacios protegidos?

Investigar el uso turístico de los espacios naturales protegidos es esencial para entender los impactos del turismo en los recursos naturales y culturales de dichos espacios, al igual que para aportar información útil para mejorar su gestión. En el desarrollo de esta investigación, al revisar la literatura sobre el turismo en espacios naturales protegidos se confirmó la importancia de entender y caracterizar el uso turístico de estos espacios y que se refleja en el variado cuerpo de literatura existente. La protección del medio natural lleva ya más de 150 años desde la primera declaración de un parque nacional y es necesario ahora el paso hacia una gestión eficiente a largo plazo. La investigación es necesaria tanto para la conservación de los valores naturales de estos espacios como para el bienestar y disfrute futuro de los recursos que ofrecen. Surgen dos aspectos clave que esta investigación ha abordado.

En primer lugar, monitorear el uso público requiere de datos fiables y continuos sobre el comportamiento de los visitantes para medir los impactos de esta actividad. La investigación enfocada en identificar fuentes de datos alternativas y desarrollar métodos de análisis puede contribuir a la gestión de los parques con información relevante. En la fase inicial de esta tesis, se constató la falta de datos completos y homogéneos en la red de parques nacionales, debido principalmente a la variedad de metodologías utilizadas para el seguimiento de uso público en cada parque. En este sentido la exploración inicial de los datos descargados de Flickr y Wikiloc, nos permitió confirmar que estos datos podían aportar información complementaria del uso turístico de la red de parques nacionales. Por ejemplo, el número de usuarios únicos por día de Flickr refleja los patrones de afluencia a los parques. La distribución espacial de ambas fuentes de datos permite visualizar la presencia de visitantes en la infraestructura y equipamientos de los parques.

En segundo lugar, investigar los aspectos que modelan el uso turístico de los espacios protegidos implica el cálculo de indicadores sobre aspectos importantes del uso de los dichos espacios, incluidos los volúmenes de uso, los tipos y la ubicación de esos usos y varias medidas de impacto,

como la satisfacción del visitante y los impactos ambientales. A lo largo de esta investigación se desarrollaron métodos para la obtención de indicadores que describen y miden el uso turístico. Por ejemplo, a partir de los datos geolocalizados de Flickr, fue posible identificar los lugares más visitados en los parques nacionales españoles y obtener las distribuciones espaciales de sus usuarios en diferentes escalas de tiempo. Así mismo, a partir del análisis de las rutas de Wikiloc, fue posible localizar y cuantificar el uso de las redes de senderos en los parques nacionales.

P.1.2: ¿Qué interés tiene el uso de fuentes de datos alternativas para el monitoreo del uso público de espacios protegidos como los parques nacionales?

A partir de la revisión de la literatura, se confirmó la hipótesis de esta investigación de que estos datos son una fuente de gran interés para el estudio del uso turístico de los espacios protegidos, debido a su heterogeneidad. Esta característica hace posible extraer características de los visitantes como la duración de la visita y el país o región de origen. Además, es posible la obtención de indicadores de uso turístico como la densidad de los visitantes, los lugares o puntos calientes de actividad y la intensidad de uso de los senderos.

Es importante destacar que el uso de datos geolocalizados en este contexto se encuentra en una etapa inicial. La mayoría de los estudios disponibles se enfocaron en evaluar la representatividad de los datos en comparación con los conteos de visitantes oficiales. En este sentido los resultados expuestos por los estudios (Martinez-Harms et al., 2018; Sessions et al., 2016; Sinclair et al., 2020; Wood et al., 2013) y los obtenidos de esta investigación han demostrado una relación positiva y estadísticamente significativa entre los indicadores de afluencia oficiales y los obtenidos a partir de fotografías geolocalizadas.

La componente temporal y espacial de las publicaciones en las redes sociales ofrece una alta resolución espacio-temporal para explorar y cuantificar los movimientos de los visitantes dentro de los espacios naturales protegidos (Mancini et al., 2018). Además, enfoques científicos menos explorados, como la geografía del tiempo, ahora se pueden redescubrir y aplicar para arrojar luz sobre cómo los visitantes se mueven y se comportan.

Otro aspecto explorado fue la identificación del origen de los visitantes. La literatura revisada mostró coeficientes consistentes y positivos al comparar datos geolocalizados con encuestas y estadísticas de parques nacionales y zonas de conservación (Sinclair et al., 2020; Sonter et al., 2016). Al respecto, en esta investigación fue posible extraer el país de origen de los visitantes de la red de parques nacionales españoles. Sin embargo, las estimaciones del país de origen revelaron

una pequeña sobreestimación de los visitantes internacionales que puede verse influenciada por el valor percibido que los visitantes dan a los destinos más alejados de su hogar, por lo que eligen publicar en redes sociales.

Un aspecto ampliamente subrayado en la literatura es la necesidad de anonimizar los datos de los usuarios que pueden comprometer el derecho a la privacidad. La obtención y análisis de datos geolocalizados de usuarios sin su consentimiento pueden plantear cuestiones éticas, por lo que la utilización de estos datos debe cumplir con las regulaciones de privacidad. En este sentido, para la realización de esta investigación se descargaron los datos de usuarios que escogieron compartir de forma pública sus publicaciones dentro de las redes sociales. Además, los datos se han agregado a diferentes escales espaciales y temporales.

Los estudios revisados reflejan el interés y utilidad de usar datos geolocalizados de redes sociales para el análisis del turismo de naturaleza, en particular el uso de fotografías geolocalizadas. A diferencia de fuentes tradicionales, las nuevas fuentes de datos proporcionan un registro continuo del comportamiento de los turistas a un costo muy bajo y de acceso fácil. Por lo tanto, estos datos pueden ser una alternativa para monitorear la actividad turística y contar con indicadores para la planificación y gestión de destinos con pocos recursos y de difícil acceso.

Pl. 3: ¿Qué fuentes de datos de redes sociales son más apropiadas para el análisis geográfico del turismo en espacios naturales protegidos?

La selección de la fuente de datos más apropiada para estudiar el uso público de los espacios naturales protegidos depende de varios factores, entre los cuales destacan la disponibilidad de los datos, la popularidad de la red social y su cobertura geográfica. En esta investigación se preseleccionaron tres redes sociales: Flickr, Twitter y Wikiloc, debido a su accesibilidad y a la diversidad de datos que ofrecen. Por ejemplo, en Flickr, los usuarios comparten fotos que revelan actividades y experiencias turísticas, en cambio los usuarios de Twitter suelen compartir actualizaciones en tiempo real y organizar sus tweets en torno a temas específicos que puede ser útil para identificar preferencias de los visitantes. Finalmente, las rutas de Wikiloc proporcionan datos geolocalizados precisos sobre las actividades al aire libre lo que puede ser muy útil para estudiar el movimiento de visitantes.

La descarga y exploración inicial de los datos de las tres redes sociales permitieron dar respuesta a esta pregunta. Con respecto a la disponibilidad de los datos, tanto Flickr como Twitter cuentan con

APIs que permiten el acceso y descarga automatizada de sus datos y tienen una amplia cobertura y popularidad en España. En tanto que Wikiloc presenta un limitante relacionado con la disponibilidad de datos ya que no dispone de una API para acceder a sus datos.

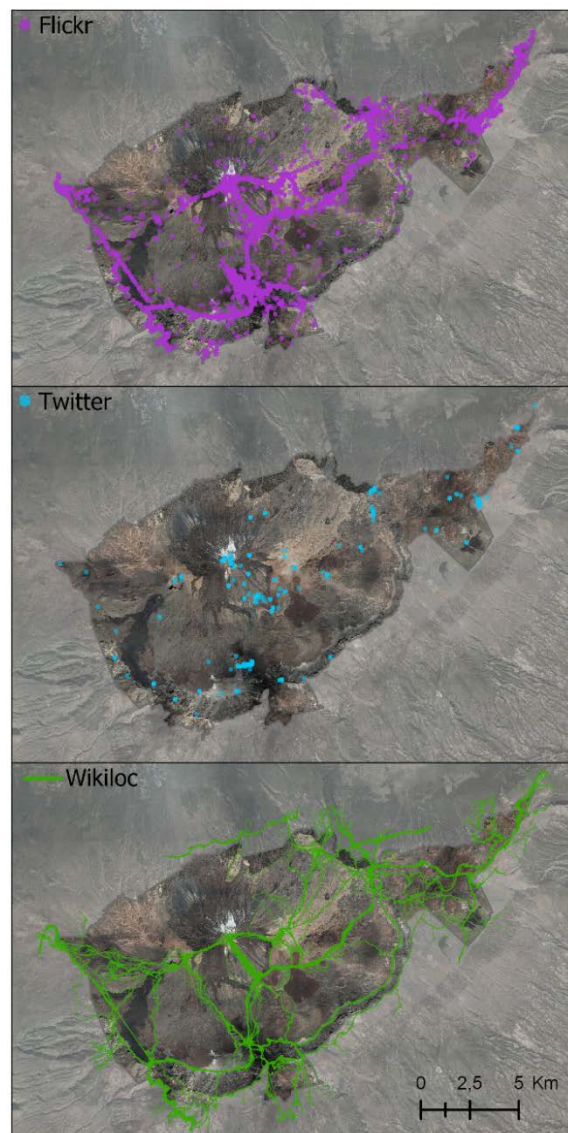
En relación con el volumen de datos descargados, las tres plataformas contenían información de la red de parques nacionales españoles, entre las cuáles Twitter destacó como la fuente de datos con un mayor número de registros. Sin embargo, al examinar los datos de Twitter en un SIG se observó que la mayoría de los tweets estaban concentrados en una misma localización (coordenadas) y no mostraban una distribución adecuada para el análisis espacial (**Figura 87**). Por lo tanto, se descartó su uso como fuente de datos, a pesar de su comprobada utilidad para la investigación del turismo urbano.

Los datos geolocalizados de Flickr se presentaron como la fuente con mayor potencial para extraer información relevante sobre el uso público de los espacios protegidos. Además de aprovechar la componente espacial y temporal para el análisis geográfico de los visitantes de parques nacionales, los datos de perfil de los usuarios demostraron ser útiles para la extracción del país de origen o residencia de los visitantes.

La fuente de datos con el menor volumen fue Wikiloc. Sin embargo, presenta la mayor granularidad espacial y temporal, además de que muestra un uso específico de los parques en actividades al aire libre. Los datos de Wikiloc proveen información detallada sobre el comportamiento de los turistas, incluidos los patrones de movimiento, la ubicación y la duración de las actividades.

En términos de integridad, los datos geolocalizados tanto de Flickr como de Wikiloc fueron más abundantes para los parques nacionales que tenían mayores tasas de visitas, accesibilidad y cercanía a ciudades grandes. Por lo tanto, la utilidad de los datos geolocalizados para estimar los indicadores de monitoreo de visitantes parece ser mayor y más confiable en áreas protegidas populares y conocidas, como los parques del Teide, Picos de Europa y Timanfaya. En los sitios menos populares (Cabañeros y Archipiélago de Cabrera), había un menor volumen de datos geolocalizados. Los resultados del análisis espacial y temporal de tales áreas deben tomarse con cautela.

Figura 86. Despliegue de datos geolocalizados de las tres redes sociales examinadas en el parque nacional Teide.



PI.4: ¿Permiten los datos de Flickr medir e identificar el ritmo temporal del turismo en la red de parques nacionales?

Esta investigación encontró correlaciones sólidas y significativas entre los datos geolocalizados de Flickr y los datos oficiales de visitantes a nivel mensual, validando así el uso de datos de redes sociales para inferir patrones temporales y como un indicador indirecto de las tasas de visitas. Los análisis de las distribuciones semanales y diarias proporcionaron información que podría ayudar a los responsables de la gestión de parques a identificar con mayor precisión los períodos de mayor saturación. Los resultados obtenidos del análisis están alineados con estudios previos (Sonter et al., 2016) que utilizaron enfoques similares.

Sin embargo, se observó que la correlación con los datos oficiales de los parques nacionales fue menor en algunos parques y que puede atribuirse a: (1) La falta de datos de visitantes, como en el caso de los parques de Cabañeros y Sierra Nevada, que resultaron en una correlación moderada, pero estadísticamente no significativa, (2) La aplicación de diferentes métodos para el conteo de visitantes, como el Parque Nacional de Guadarrama, que ha cambiado su método de conteo dos veces desde su creación en 2013 (PNSG, 2016); (3) Los visitantes recurrentes toman menos fotografías.

Muchos autores han postulado que las condiciones naturales son determinantes para los patrones de visita en las áreas protegidas (Butler, 2001; Hadwen et al., 2011). En la red de parques nacionales españoles, la ubicación geográfica parece ser un factor importante para la estacionalidad. Los parques de montaña, Aigüestortes y Picos de Europa, reportaron índices altos de estacionalidad. Esto también ocurrió con los parques insulares (Archipiélago de Cabrera e Islas Atlánticas de Galicia), que son atractivos para el turismo náutico y de playa. En parques con valores de estacionalidad medios, el entorno natural específico de cada parque podría ser el factor de picos estacionales fuera de la temporada de verano. Por ejemplo, en el parque nacional de Doñana, el mayor índice de visitas se da en mayo. Esto coincide con la temporada de reproducción de muchas especies de aves, lo que significa que los visitantes tienen más oportunidades de observar aves.

Por el contrario, la baja estacionalidad en algunos parques podría deberse a factores externos. La proximidad a grandes núcleos de población podría aumentar el número de visitantes en cada mes, como es el caso, de los parques de Sierra Nevada y Sierra de Guadarrama, muy próximos a las ciudades de Granada y Madrid, respectivamente. Además, los parques ubicados en destinos turísticos con alta demanda podrían verse influenciados por los flujos turísticos regulares donde los turistas optan por visitar los parques nacionales como una actividad adicional durante sus vacaciones, como es el caso de los parques ubicados en las Islas Canarias.

La alta concentración diaria de visitantes en algunos parques podría indicar problemas de aglomeración. En tan solo tres horas, los parques de Timanfaya, Garajonay, Aigüestortes, Ordesa y la Sierra de Guadarrama concentran hasta el 50% de la actividad observada en Flickr. Si bien esta distribución puede estar condicionada por el “efecto retorno”, todo parece indicar una mayor presencia de visitantes durante la mañana.

Las distribuciones temporales extraídas de los datos geolocalizados de Flickr pueden ser útiles para los administradores de parques. Por ejemplo, la distribución de visitantes a diferentes escalas se puede utilizar para tareas de gestión como la programación de equipos, la contratación de personal adicional, el seguimiento de puntos de interés o la planificación de horarios de visitas. En esta investigación, ninguno de los parques tenía una distribución semanal o diaria de los visitantes, por lo que los datos geolocalizados de los usuarios de Flickr ayudaron a ilustrar el comportamiento temporal de los visitantes durante el día y su afluencia durante la semana.

PI.5: ¿Puede el componente espacial de los datos de Flickr usarse como indicador de la distribución espacial del turismo en la red de parques nacionales?

Los datos geolocalizados de redes sociales revelaron patrones espaciotemporales relevantes, como la intensidad de uso y la localización de puntos calientes que no pueden ser capturados por encuestas predefinidas. Los resultados de dicho análisis fueron consistentes con el comportamiento esperado e informado de los visitantes en parques nacionales.

La geolocalización de los datos de Flickr se utilizó en esta investigación para determinar la distribución espacial de los visitantes en los 15 parques nacionales españoles. En este sentido, Flickr brinda información adicional, no disponible en fuentes oficiales. La actividad social de compartir imágenes deja indicadores digitales de preferencias espaciales, con personas que comparten fotos específicas considerando que el lugar representado no solo "vale la pena visitarlo" sino también "vale la pena compartirlo visualmente" (Gliozzo et al., 2016).

La densidad Kernel, se utilizó para estimar en términos relativos la densidad de visitantes en los parques nacionales. Los mapas de densidad mostraron una aproximación inicial por la cual se pueden identificar áreas con alta concentración de visitantes. En esta investigación se identificaron en todos los parques zonas de densidad media y alta, destacando los parques de las islas Canarias y el parque Islas Atlánticas de Galicia como los espacios que presentaron el mayor número de zonas con densidad alta. Además, al utilizar el componente temporal de los datos es posible analizar la densidad de visitantes en diferentes momentos y comprobar si las zonas de densidad varían en cuanto a número y magnitud dentro de los parques. Estos resultados pueden ser utilizados para la toma de decisiones informadas sobre la asignación de recursos.

El índice global de Moran se utilizó para verificar la existencia de patrones de concentración espacial en los datos geolocalizados de Flickr. Los resultados del índice de autocorrelación global de Moran fueron positivos y estadísticamente significativos para todos los parques, indicando que existen patrones de agrupación de los valores altos y bajos de las distribuciones espaciales de

usuarios en los parques nacionales. Los parques de Islas Atlánticas de Galicia, Tablas de Daimiel y Doñana presentaron los valores más altos de autocorrelación lo que sugiere que en estos parques los patrones de agrupamiento son más fuertes.

Al confirmar la existencia de patrones de agrupamiento en los datos, se aplicó el índice local de autocorrelación LISA para localizar los clústeres de usuarios dentro de los parques. La mayoría de los clústeres en los parques coinciden con los principales puntos de interés y los equipamientos de los parques nacionales. Los datos sobre la concentración espacial y temporal de los visitantes pueden alertar a los administradores del parque sobre los problemas de aglomeración de visitantes y brindar información útil para la toma de decisiones para enfrentarlos.

PI.6: ¿Pueden los datos de Wikiloc ser indicadores de la movilidad turística dentro de los parques nacionales?

Las rutas GPS obtenidas de Wikiloc han proporcionado datos altamente precisos acerca de los movimientos de los visitantes dentro de la red de parques nacionales en España. Con relación a estas rutas descargadas, se utilizaron para estudiar la intensidad de uso de las redes de senderos en seis parques nacionales, enfocándose en dos actividades: senderismo y ciclismo.

El análisis de intensidad de uso ha revelado patrones claros en todos los parques, identificando senderos ampliamente transitados en contraste con aquellos que presentan menor interés para los visitantes. Adicionalmente, al contrastar estos resultados con los datos de itinerarios oficiales, se ha logrado descubrir nuevos recorridos de interés para los visitantes. La incorporación de estos itinerarios potenciales a la oferta de los parques podría enriquecer significativamente la experiencia ofrecida a los visitantes.

Las rutas de GPS provenientes de redes sociales pueden proporcionar datos precisos sobre el movimiento de los visitantes dentro de los parques nacionales, incluidos los senderos recorridos, las distancias recorridas y el tiempo que pasan en varios lugares. Esta información puede ayudar a los gestores de parques a comprender el comportamiento y las preferencias de los visitantes para tomar decisiones informadas sobre el mantenimiento de senderos y el desarrollo de infraestructura.

Además, las rutas GPS se pueden utilizar para monitorear el cumplimiento de los visitantes con las reglas y regulaciones del parque, como permanecer en los senderos designados y minimizar el impacto en los ecosistemas frágiles. Sin embargo, es importante tener en cuenta que el uso de las rutas GPS de redes sociales debe hacerse de manera que se respete la privacidad del visitante y se

mantenga la confidencialidad de la información personal, lo que se puede lograr agregando los datos a aplicando técnicas de anonimización.

PI. 7: ¿Qué tipo de aplicaciones prácticas para el manejo y planificación de espacios pueden ofrecerse a partir de la huella digital?

En esta investigación se realizaron dos aplicaciones con los datos geolocalizados de Flickr y Wikiloc.

En la primera aplicación se utilizó la geolocalización de los usuarios de Flickr para modelar la demanda potencial de visitantes en el parque nacional del Teide. Con los datos de la demanda como parámetro se probaron modelos de ubicación y asignación para identificar las localizaciones óptimas de nuevos puestos de información dentro del parque en distintos escenarios. Los resultados obtenidos muestran que los datos de Flickr pueden ser valiosos para la toma de decisiones y la optimización de recursos en la gestión del parque, ya que la infraestructura de visitantes es un componente clave para atraer visitantes y contribuir a su educación ambiental. Los resultados obtenidos de los modelos de ubicación-asignación, en el caso del Teide mostraron que las instalaciones existentes cubren menos del 50% de la demanda potencial. Por lo tanto, la adición de una o más instalaciones podría mejorar la experiencia de los visitantes con respecto al suministro de información sobre los atributos y actividades del parque.

En la segunda aplicación se utilizaron los datos de Flickr y Wikiloc para analizar el movimiento de los visitantes dentro del parque Aigüestortes. En base a los itinerarios obtenidos de los usuarios de ambas aplicaciones fue posible extraer el punto de inicio de cada itinerario y utilizarlo para inferir el acceso del visitante y cuantificar de forma relativa el flujo de visitantes desde los accesos al parque. Para analizar los patrones de movimiento en las rutas se utilizó el índice local de Moran. A partir de sus resultados se puede identificar las paradas que realizan los usuarios de Wikiloc como las localizaciones de clústeres de valores bajos de velocidad. Los resultados obtenidos permitieron descubrir los lugares más atractivos para senderistas y ciclistas y turistas y comprender los patrones comunes en el orden de visita de esos lugares.

En este sentido, Flickr y Wikiloc proporcionan información complementaria sobre el uso de la red de caminos y senderos: mientras que Flickr detecta principalmente los flujos de visitantes cuando hacen turismo, Wikiloc identifica los senderos utilizados para practicar actividades como senderismo o ciclismo.

10.2. Futuras líneas de investigación:

El uso de nuevas fuentes de datos para la investigación es un área de estudio de gran interés. En esta sección se describen algunas de las líneas de investigación que están surgiendo y han quedado pendientes de estudio.

Es necesario realizar nuevos estudios que utilicen y comparen datos de diferentes plataformas sociales con el objetivo de identificar las ventajas y limitantes de cada plataforma y establecer cuándo y cómo utilizar cada una de las nuevas fuentes de datos. De igual manera, estudios que combinen los datos de diferentes plataformas sociales pueden ser útiles para mejorar la representatividad y la longevidad de los datos (Kim et al., 2019; Norman & Pickering, 2017; Tenkanen et al., 2017). Cada red social está dirigida a un sector definido de la población y a su vez su popularidad varía entre países, por lo que combinar varias fuentes de datos puede agregar representatividad a los datos analizados. La combinación de múltiples fuentes de datos también puede contribuir a reducir el sesgo en los análisis ya que cada plataforma puede estar más orientada hacia un segmento específico de la población. Además, el uso de datos de diferentes plataformas sociales aporta información complementaria para clasificar a los visitantes y las actividades realizadas en las áreas protegidas. En los estudios revisados, solo 8 estudios utilizaron más de una red social como fuente de datos, por lo que es oportuno seguir investigando en el uso combinado de las nuevas fuentes de datos.

Sin embargo, el acceso a las fuentes puede suponer un limitante a la hora de seleccionar la plataforma social. En este aspecto, la revisión mostró que la mayoría de los estudios obtuvieron los datos de redes sociales que cuentan con APIs para la descarga de datos. Los cambios en el acceso a la API, la creación de nuevas restricciones al uso de datos y las limitaciones de datos pueden restringir el uso de datos de redes sociales para la investigación. En general, Flickr sigue siendo la fuente de datos más utilizada debido a su API de fácil acceso. Se están empezando a aplicar nuevos métodos de recuperación de datos, como el web scraping, que pueden aumentar el uso de más sitios de redes sociales como fuente de datos.

Otra línea de investigación es la aplicación de técnicas de minería de datos para examinar y explicar el comportamiento de los visitantes de espacios naturales protegidos. Se podrían utilizar algoritmos de minería de patrones secuenciales, por ejemplo, el algoritmo a priori, para identificar los patrones y orden de visita de los diferentes puntos de interés dentro de un espacio protegido. En este sentido, los datos de sitios de intercambio de rutas GPS como Wikiloc, Strava, GPSies y Map my fitness parecen ser los más adecuados para realizar estos análisis.

La aplicación de técnicas de procesamiento de lenguaje natural (NLP) para analizar el sentimiento presente en los textos relacionados con las publicaciones de datos geolocalizados ofrece una valiosa fuente de información acerca de las percepciones y experiencias de los visitantes durante su estadía en los espacios protegidos.

Además, la investigación en el uso de modelos de series temporales, una técnica de minería de datos aún poco explorada con datos geolocalizados puede resultar útil para realizar predicciones. Estos modelos pueden prever las tendencias futuras en la afluencia de visitantes en momentos específicos. Dada la granularidad de los datos geolocalizados provenientes de las redes sociales, estas predicciones tienen el potencial de brindar conocimientos significativos para la gestión eficiente de los parques nacionales y demás espacios protegidos.

La exploración y aplicación de técnicas de análisis de Big Data, como el aprendizaje automático, también abre una interesante línea de investigación futura. La investigación ha comenzado a aplicarlos, pero es necesario producir más investigaciones sobre el uso de técnicas de detección de imágenes y etiquetas para validar los datos de las redes sociales como indicadores de las experiencias y satisfacción de los visitantes. Las fotografías geolocalizadas pueden proporcionar información sobre los tipos de actividades que realizan los visitantes dentro de los parques nacionales. Por ejemplo, las fotos con etiquetas específicas como "senderismo" o "vida salvaje" pueden ser útiles para identificar las actividades más populares entre los visitantes. Esta información puede ser valiosa para fines de gestión y marketing turístico.

11. Bibliografía

11. Bibliografía

- Adrienko, N., & Andrienko, G. (2011). Spatial generalization and aggregation of massive movement data. *IEEE Transactions on Visualization and Computer Graphics*, *17*(2), 205–219. <https://doi.org/10.1109/TVCG.2010.44>
- AEITG. (2019). *Encuesta de Turismo en destino de Galicia*. https://aei.turismo.gal/osdam/filestore/5/1/1/8/0_ea2f0f1b3b4afa0/51180_15a9489233c8fd3.pdf
- Amo, L., López, P., & Martín, J. (2006). Nature-based tourism as a form of predation risk affects body condition and health state of *Podarcis muralis* lizards. *Biological Conservation*, *131*(3), 402–409. <https://doi.org/10.1016/j.biocon.2006.02.015>
- Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Anselin, L. (1999). Interactive techniques and exploratory spatial data analysis. *Geographic Information Systems: Principles, Techniques, Management and Applications*, 253–266.
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology: Theory and Practice*, *8*(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Ballantyne, R., Packer, J., & Falk, J. (2011). Visitors' learning for environmental sustainability: Testing short- and long-term impacts of wildlife tourism experiences using structural equation modelling. *Tourism Management*, *32*(6), 1243–1252. <https://doi.org/10.1016/j.tourman.2010.11.003>
- Balmford, A., Beresford, J., Green, J., Naidoo, R., Walpole, M., & Manica, A. (2009). A global perspective on trends in nature-based tourism. *PLoS Biology*, *7*(6), 1–6. <https://doi.org/10.1371/journal.pbio.1000144>
- Balmford, A., Green, J., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M., & Manica, A. (2015). Walk on the Wild Side: Estimating the Global Magnitude of Visits to Protected Areas. *PLoS Biology*, *13*(2), e1002074. <https://doi.org/10.1371/journal.pbio.1002074>
- Barros, C., Moya-Gómez, B., & García-Palomares, J. C. (2019). Identifying Temporal Patterns of Visitors to National Parks through Geotagged Photographs. *Sustainability*, *11*(24), 1–16. <https://doi.org/10.3390/su11246983>
- Barros, C., Moya-Gómez, B., & Gutiérrez, J. (2020). Using geotagged photographs and GPS tracks from social networks to analyse visitor behaviour in national parks. *Current Issues in Tourism*, *23*(10), 1291–1310. <https://doi.org/10.1080/13683500.2019.1619674>
- Bassolas, A., Lenormand, M., Tugores, A., Gonçalves, B., & Ramasco, J. J. (2016). Touristic site attractiveness seen through Twitter. *EPJ Data Science*, *5*(1). <https://doi.org/10.1140/epjds/s13688-016-0073-5>
- Batrinca, B., & Treleaven, P. C. (2014). Social media analytics: a survey of techniques, tools and platforms. *AI and Society*, *30*(1), 89–116. <https://doi.org/10.1007/s00146-014-0549-4>
- Baum, T., & Lundtorp, S. (2001). Seasonality in tourism. In T. Baum & S. Lundtorp (Eds.), *Seasonality in tourism* (pp. 23–50). Elsevier.
- Bojic, I., Massaro, E., Belyi, A., Sobolevsky, S., & Ratti, C. (2015). *Choosing the Right Home Location*

- Definition Method for the Given Dataset BT - Social Informatics* (T.-Y. Liu, C. N. Scollon, & W. Zhu (eds.); pp. 194–208). Springer International Publishing.
- Brandt, T., Bendler, J., & Neumann, D. (2016). Social media analytics and value creation in urban smart tourism ecosystems. *Information and Management*.
<https://doi.org/10.1016/j.im.2017.01.004>
- Buckley, R. (2008). *Ecotourism: principles and practices*. CABI.
- Butler, R. W. (2001). Seasonality in tourism: Issues and implications. In T. Baum & S. Lundtorp (Eds.), *Seasonality in tourism* (pp. 5–22). Pergamon Oxford.
- Buultjens, J., Ratnayake, I., Gnanapala, A., & Aslam, M. (2005). Tourism and its implications for management in Ruhuna National Park (Yala), Sri Lanka. *Tourism Management*, 26(5), 733–742. <https://doi.org/10.1016/j.tourman.2004.03.014>
- Callau, A. À., Pérez Albert, M. Y., Rota, J. J., & Giné, D. S. (2019). Landscape characterization using photographs from crowdsourced platforms: Content analysis of social media photographs. *Open Geosciences*, 11(1), 558–571. <https://doi.org/10.1515/geo-2019-0046>
- Campelo, M. B., & Nogueira Mendes, R. M. (2016). Comparing webshare services to assess mountain bike use in protected areas. *Journal of Outdoor Recreation and Tourism*, 15, 82–88. <https://doi.org/10.1016/j.jort.2016.08.001>
- Cessford, G., & Muhar, A. (2003). Monitoring options for visitor numbers in national parks and natural areas. *Journal of Nature Conservation*, 11(4), 240–250. <https://doi.org/10.1078/1617-1381-00055>
- Chareyron, G. G., Da-Rugna, J. J., & Raimbault, T. (2014). Big data: A new challenge for tourism. *Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014*, 5–7. <https://doi.org/10.1109/BigData.2014.7004475>
- Cherkassky, B. V., Goldberg, A. V., & Radzik, T. (1994). Shortest paths algorithms: Theory and experimental evaluation. *Proceedings of the Annual ACM SIAM Symposium on Discrete Algorithms*, 516–525.
- Chua, A., Servillo, L., Marcheggiani, E., & Moere, A. Vande. (2016). Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. *Tourism Management*, 57, 295–310. <https://doi.org/10.1016/j.tourman.2016.06.013>
- Chun, J., Kim, C. K., Kim, G. S., Jeong, J. Y., & Lee, W. K. (2020). Social big data informs spatially explicit management options for national parks with high tourism pressures. *Tourism Management*, 81, 104136. <https://doi.org/10.1016/j.tourman.2020.104136>
- Cifuentes, M. (1992). Determinacion De Capacidad De Carga Turistica En Areas Protegidas. In *Turialba (Costa Rica): CATIE* (p. 23).
- Clemente, P., Calvache, M., Antunes, P., Santos, R., Cerdeira, J. O., & Martins, M. J. (2019). Combining social media photographs and species distribution models to map cultural ecosystem services: The case of a Natural Park in Portugal. *ECOLOGICAL INDICATORS*, 96(1), 59–68. <https://doi.org/10.1016/j.ecolind.2018.08.043>
- Comber, A., Dickie, J., Jarvis, C., Phillips, M., & Tansey, K. (2015). Locating bioenergy facilities using a modified GIS-based location–allocation–algorithm: Considering the spatial distribution of resource supply. *Applied Energy*, 154, 309–316. <https://doi.org/10.1016/J.APENERGY.2015.04.128>
- Da Rugna, J., Chareyron, G., & Branchet, B. (2012). Tourist behavior analysis through geotagged

photographies: A method to identify the country of origin. *CINTI 2012 - 13th IEEE International Symposium on Computational Intelligence and Informatics, Proceedings*, 347–351. <https://doi.org/10.1109/CINTI.2012.6496788>

- Dalumpines, R., & Scott, D. M. (2011). *GIS-based Map-matching: Development and Demonstration of a Postprocessing Map-matching Algorithm for Transportation Research*. 101–120. https://doi.org/10.1007/978-3-642-19789-5_27
- De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. *AIP Conference Proceedings*, 1644, 97–104. <https://doi.org/10.1063/1.4907823>
- Demchenko, Y., & Membrey, P. (2013). Addressing Big Data Issues in Scientific Data Infrastructure. *International Conference on Collaboration Technologies and Systems (CTS)*, 48–55. <https://doi.org/10.1201/b21988>
- Di Minin, E., Tenkanen, H., & Toivonen, T. (2015). Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science*, 3(September), 63. <https://doi.org/10.3389/fenvs.2015.00063>
- Do, Y., & Kim, J. Y. (2020). An assessment of the aesthetic value of protected wetlands based on a photo content and its metadata. *Ecological Engineering*, 150(March), 105816. <https://doi.org/10.1016/j.ecoleng.2020.105816>
- Dudley, N. (2008). *Guidelines for applying protected area management categories* (N. Dudley (ed.); Issue 21). IUCN. <https://doi.org/10.2305/IUCN.CH.2008.PAPS.2.en>
- Eagles, P. F. J. (2014). Research priorities in park tourism. *Journal of Sustainable Tourism*, 22(4), 528–549. <https://doi.org/10.1080/09669582.2013.785554>
- Eagles, P. F. J., & McCool, S. F. (2002). *Tourism in national parks and protected areas: planning and management* (CABI (ed.)). [https://doi.org/10.1016/S0261-5177\(03\)00091-8](https://doi.org/10.1016/S0261-5177(03)00091-8)
- Eagles, P. F. J., McCool, S. F., & Haynes, C. D. (2002). Sustainable Tourism in Protected Areas. In A. Phillips (Ed.), *World* (Issue 8). IUCN, Gland, Switzerland, and Cambridge, UK, the United Nations Environment Programme and the World Tourism Organization. <https://doi.org/10.2305/IUCN.CH.2002.PAG.8.en>
- ESRI. (2018). *Location-allocation analysis*. <http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/location-allocation.htm>
- ESRI. (2020). *Track intervals to feature*. <https://desktop.arcgis.com/en/arcmap/latest/tools/tracking-analyst-toolbox/track-intervals-to-feature.htm>
- Farrell, T. A., & Marion, J. L. (2001). Identifying and assessing ecotourism visitor impacts at eight protected areas in Costa Rica and Belize. *Environmental Conservation*, 28(03), 215–225. <https://doi.org/10.1017/S0376892901000224>
- Ferreira, D. (2019). Big data and innovation in tourism, travel, and hospitality: Managerial approaches, techniques, and applications. In *Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques, and Applications* (pp. 1–223). Springer Singapore. <https://doi.org/10.1007/978-981-13-6339-9>
- Flickr. (2017a). *Flickr History*. <https://www.flickr.com/photos/flickr/12433867035>
- Flickr. (2017b). *The Flickr Developer Guide: API*. <https://www.flickr.com/services/developer/api/>

- Flickr. (2018). *Why we're changing Flickr free accounts*.
<https://blog.flickr.net/en/2018/11/01/changing-flickr-free-accounts-1000-photos/>
- Frost, W., & Hall, C. M. (2009). *Tourism and national parks: International perspectives on development, histories, and change* (Vol. 14). Routledge.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144.
<https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- García-Palomares, J. C., Gutiérrez-Puebla, J., & Mínguez, C. (2015). Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS. *Applied Geography*, 63, 408–417.
<https://doi.org/10.1016/j.apgeog.2015.08.002>
- Garrod, B. (2008). Exploring place perception a photo-based analysis. *Annals of Tourism Research*, 35(2), 381–401. <https://doi.org/10.1016/j.annals.2007.09.004>
- Gartner. (2017). *Gartner IT Glossary*. <https://www.gartner.com/en/information-technology/glossary/big-data>
- Getis, A. (2008). A history of the concept of spatial autocorrelation: A geographer's perspective. *Geographical Analysis*, 40(3), 297–309. <https://doi.org/10.1111/j.1538-4632.2008.00727.x>
- Ghermandi, A., & Sinclair, M. (2019). Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change*, 55(February), 36–47.
<https://doi.org/10.1016/j.gloenvcha.2019.02.003>
- Giglio, S., Bertacchini, F., Bilotta, E., & Pantano, P. (2019). Using social media to identify tourism attractiveness in six Italian cities. *Tourism Management*, 72(January 2018), 306–312.
<https://doi.org/10.1016/j.tourman.2018.12.007>
- Girardin, F., Blat, J., Calabrese, F., Dal Fiore, F., & Ratti, C. (2008). Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7(4), 36–44.
<https://doi.org/10.1109/MPRV.2008.71>
- Gliozzo, G., Pettorelli, N., & Haklay, M. (Muki). (2016). Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK. *Ecology and Society*, 21(3), art6.
<https://doi.org/10.5751/ES-08436-210306>
- Global Data. (2017). *Twitter: Monetizable Daily Active Users*. <https://www.globaldata.com/data-insights/technology--media-and-telecom/twitter-monetizable-daily-active-users/>
- González, A., Fosse, J., & Costa, C. (2021). *Evitar la masificación del turismo en los parques nacionales de España : diagnóstico y directrices*. <https://doi.org/http://doi.org/10.5281/zenodo.4748809>
- Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211–221. <https://doi.org/10.1007/s10708-007-9111-y>
- Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1, 110–120. <https://doi.org/10.1016/j.spasta.2012.03.002>
- Gosal, A. S., Geijzendorffer, I. R., Václavík, T., Poulin, B., Ziv, G., Vaclavik, T., Poulin, B., & Ziv, G. (2019). Using social media, machine learning and natural language processing to map multiple recreational beneficiaries. *Ecosystem Services*, 38(June 2018), 100958.
<https://doi.org/10.1016/j.ecoser.2019.100958>

- Grekousis, G. (2020). *Spatial analysis methods and practice : describe, explore, explain through GIS*.
- Grombridge, B. (1992). Global biodiversity status of the earth's living resources. In *Chapman and Hall* (Vol. 1). <https://doi.org/10.1088/1751-8113/44/8/085201>
- Gutiérrez-Puebla, J. (2018). Big Data y nuevas geografías: la huella digital de las actividades humanas. *Documents d'Anàlisi Geogràfica*, 64(2), 195. <https://doi.org/10.5565/rev/dag.526>
- Gutiérrez-Puebla, J., & García-Palomares, J. C. (2016). Redes sociales y estadística espacial en el estudio de la población extranjera. *XV Congreso de La Población Española, June*, 651–666.
- Hadwen, W. L., Arthington, A. H., Boon, P. I., Taylor, B., & Fellows, C. S. (2011). Do climatic or institutional factors drive seasonal patterns of tourism visitation to protected areas across diverse climate zones in eastern Australia? *Tourism Geographies*, 13(2), 187–208. <https://doi.org/10.1080/14616688.2011.569568>
- Hadwen, W. L., Hill, W., & Pickering, C. M. (2007). Icons under threat: Why monitoring visitors and their ecological impacts in protected areas matters. *Ecological Management and Restoration*, 8(3), 177–181. <https://doi.org/10.1111/j.1442-8903.2007.00364.x>
- Hamstead, Z. A., Fisher, D., Ilieva, R. T., Wood, S. A., McPhearson, T., & Kremer, P. (2018). Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers, Environment and Urban Systems*, 72(February), 38–50. <https://doi.org/10.1016/j.compenvurbsys.2018.01.007>
- Hausmann, A., Toivonen, T., Fink, C., Heikinheimo, V., Tenkanen, H., Butchart, S. H. M., Brooks, T. M., & Di Minin, E. (2019). Assessing global popularity and threats to Important Bird and Biodiversity Areas using social media data. *Science of the Total Environment*, 683, 617–623. <https://doi.org/10.1016/j.scitotenv.2019.05.268>
- Hausmann, A., Toivonen, T., Heikinheimo, V., Tenkanen, H., Slotow, R., Di Minin, E., & Minin, E. Di. (2017). Social media reveal that charismatic species are not the main attractor of ecotourists to sub-Saharan protected areas. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-017-00858-6>
- Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., & Di Minin, E. (2018). Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas. *Conservation Letters*, 11(1), 1–10. <https://doi.org/10.1111/conl.12343>
- Hecht, B., Hong, L., Suh, B., & Chi, E. H. (2011). Tweets from Justin Bieber's heart. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 237–246. <https://doi.org/10.1145/1978942.1978976>
- Heikinheimo, V., Minin, E. Di, Tenkanen, H., Hausmann, A., Erkkonen, J., Toivonen, T., Di Minin, E., Tenkanen, H., Hausmann, A., Erkkonen, J., Toivonen, T., Minin, E. Di, Tenkanen, H., Hausmann, A., Erkkonen, J., & Toivonen, T. (2017). User-Generated Geographic Information for Visitor Monitoring in a National Park: A Comparison of Social Media Data and Visitor Survey. *ISPRS International Journal of Geo-Information*, 6(3), 85. <https://doi.org/10.3390/ijgi6030085>
- Hernández Álvarez, J. C. (2017). *Parque Nacional Teide: Visita Actual y retos del futuro*.
- Hernández de la Obra, J., & Gómez Limón García, J. (2005). *Manual sobre conceptos de uso público en los espacios naturales protegidos* (Issue 11).
- Hochmair, H., & Cvetojevic, S. (2014). Assessing the Usability of Georeferenced Tweets for the

- Extraction of Travel Patterns: A Case Study for Austria and Florida. *GI_Forum 2014*, 30–39. <https://doi.org/10.1553/giscience2014s30.The>
- Holdgate, M. (1999). The green web: a union for world conservation. In Earthscan (Ed.), *Journal of Environmental Policy & Planning* (Vol. 2, Issue 3). <https://doi.org/10.1080/714038559>
- Holmes, J., Williams, F. B., & Brown, L. A. (1972). Facility Location under a Maximum Travel Restriction: An Example Using Day Care Facilities. *Geographical Analysis*, 4(3), 258–266. <https://doi.org/10.1111/j.1538-4632.1972.tb00474.x>
- INE. (2022). *Número de turistas según país de residencia*. Número de turistas según país de residencia
- INE. (2023). *Encuesta de ocupación en alojamientos de turismo rural*. <https://www.ine.es/jaxi/Tabla.htm?tpx=53558&L=0>
- Ishwarappa, & Anuradha, J. (2015). A brief introduction on big data 5Vs characteristics and hadoop technology. *Procedia Computer Science*, 48(C), 319–324. <https://doi.org/10.1016/j.procs.2015.04.188>
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Jenkins, Z. (2017). *Happy 13th Birthday, Flickr!* <https://blog.flickr.net/en/2017/02/10/happy-13th-birthday-flickr/>
- Jiménez, A. M. (1991). mediante estimadores Kernel '. *Treballs de La Societat Catalana de Geografia*, 155(445), 445–460.
- Jones, T., & Ohsawa, T. (2016). Monitoring nature-based tourism trends in Japan's national parks: Mixed messages from domestic and inbound visitors. *Parks*, 22(1), 25–36. <https://doi.org/10.2305/IUCN.CH.2016.PARKS-22-1TJ.en>
- Jurado Rota, J., Perez Albert, M. Y., Serrano Giné, D., Pérez Albert, M. Y., Serrano Giné, D., Perez Albert, M. Y., & Serrano Giné, D. (2019). Visitor monitoring in protected areas: an approach to Natura 2000 sites using Volunteered Geographic Information (VGI). *Geografisk Tidsskrift - Danish Journal of Geography*, 119(1), 69–83. <https://doi.org/10.1080/00167223.2019.1573409>
- Kádár, B., & Gede, M. (2013). Where Do Tourists Go? Visualizing and Analyzing the Spatial Distribution of Geotagged Photography. *Cartographica*, 48(2), 78–88. <https://doi.org/10.3138/carto.48.2.1839>
- Kastner, M., Tricco, A. C., Soobiah, C., Lillie, E., Perrier, L., Horsley, T., Welch, V., Cogo, E., Antony, J., & Straus, S. E. (2012). What is the most appropriate knowledge synthesis method to conduct a review? Protocol for a scoping review. *BMC Medical Research Methodology*, 12, 1–10. <https://doi.org/10.1186/1471-2288-12-114>
- Kim, Y., Kim, C. ki, Lee, D. K., Lee, H. woo, & Andrada, R. I. T. I. T. (2019). Quantifying nature-based tourism in protected areas in developing countries by using social big data. *TOURISM MANAGEMENT*, 72(December 2018), 249–256. <https://doi.org/10.1016/j.tourman.2018.12.005>
- Lee, Y., Kwon, P., Yu, K., & Park, W. (2016). Method for Determining Appropriate Clustering Criteria of Location-Sensing Data. *ISPRS International Journal of Geo-Information*, 5(9), 151. <https://doi.org/10.3390/ijgi5090151>
- Letouzé, E. (2012). *Big Data for Development: Challenges & Opportunities* (Issue May).

- Levin, N., Kark, S., & Crandall, D. (2015). Where have all the people gone? Enhancing global conservation using night lights and social media. *Ecological Applications : A Publication of the Ecological Society of America*, 25(8), 2153–2167. <https://doi.org/10.1890/15-0113.1.sm>
- Levin, N., Lechner, A. M., & Brown, G. (2017). An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied Geography*, 79(INE. (n., 115–126. <https://doi.org/10.1016/j.apgeog.2016.12.009>
- Lewis, L. (2021). *What happens in an Internet Minute*. <https://www.allaccess.com/merge/archive/32972/infographic-what-happens-in-an-internet-minute>
- Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of twitter and flickr. *Cartography and Geographic Information Science*, 40(2), 61–77. <https://doi.org/10.1080/15230406.2013.777139>
- Lin, T. P. (2010). Carbon dioxide emissions from transport in Taiwan's national parks. *Tourism Management*, 31(2), 285–290. <https://doi.org/10.1016/j.tourman.2009.03.009>
- Livina, A., & Reddy, M. (2017). Nature park as a resource for nature based tourism. *Vide. Tehnologija. Resursi - Environment, Technology, Resources*, 1, 179–183. <https://doi.org/10.17770/etr2017vol1.2590>
- López, I., & Pardo, M. (2018). Tourism versus nature conservation: reconciliation of common interests and objectives — an analysis through Picos de Europa National Park. *Journal of Mountain Science*, 15(11), 2505–2516. <https://doi.org/10.1007/s11629-018-4943-0>
- Lothian, W. F. (1987). *A Brief history of Canada's national parks*.
- Lundtorp, S. (2001). Measuring tourism seasonality. In S. Lundtorp & T. Baum (Eds.), *Seasonality in tourism* (First edit, pp. 23–50). Pergamon Oxford.
- Lyon, K., Cottrell, S. P., Siikamäki, P., & Marwijk van, R. (2011). Biodiversity Hotspots and Visitor Flows in Oulanka National Park, Finland. *Scandinavian Journal of Hospitality and Tourism*, 11(SUPPL. 1), 100–111. <https://doi.org/10.1080/15022250.2011.629909>
- Mancini, F., Coghill, G. M., & Lusseau, D. (2018). Using social media to quantify spatial and temporal dynamics of nature-based recreational activities. *PLOS ONE*, 13(7), e0200565. <https://doi.org/10.1371/journal.pone.0200565>
- Manning, R. E. (2002). How much is too much? Carrying capacity of national parks and protected areas. ... *of Visitor Flows in Recreational and Protected Areas. ...*, 306–313.
- MAPAMA, & OAPN. (2017). *Boletín de la red de parques nacionales*. <http://www.mapama.gob.es/es/red-parques-nacionales/boletin/visitantes-teide.aspx>
- Marques, C., Reis, E., & Menezes, J. (2010). Profiling the segments of visitors to Portuguese protected areas. *Journal of Sustainable Tourism*, 18(8), 971–996. <https://doi.org/10.1080/09669582.2010.497222>
- Márquez-Pérez, J., Vallejo-Villalta, I., & Álvarez-Francoso, J. I. (2017). Estimated travel time for walking trails in natural areas. *Geografisk Tidsskrift - Danish Journal of Geography*, 117(1), 53–62. <https://doi.org/10.1080/00167223.2017.1316212>
- Martinez-Harms, M. J., Bryan, B. A., Wood, S. A., Fisher, D. M., Law, E., Rhodes, J. R., Dobbs, C., Biggs, D., & Wilson, K. A. (2018). Inequality in access to cultural ecosystem services from protected areas in the Chilean biodiversity hotspot. *Science of the Total Environment*, 636, 1128–1138. <https://doi.org/10.1016/j.scitotenv.2018.04.353>

- Mc Cool, S. F., & Lime, D. W. (2001). Tourism carrying capacity: Tempting fantasy or useful reality? *Journal of Sustainable Tourism*, 9(5), 372–388. <https://doi.org/10.1080/09669580108667409>
- Mckercher, B., Shoval, N., Ng, E., Birenboim, A., Mckercher, B. O. B., Shoval, N., & Ng, E. (2012). First and Repeat Visitor Behaviour : GPS Tracking and GIS Analysis in Hong Kong First and Repeat Visitor Behaviour : GPS Tracking and GIS Analysis in Hong Kong. *Tourism Geographies*, 14(1), 147–161. <https://doi.org/10.1080/14616688.2011.598542>
- Mínguez García, C., Calle Vaquero, M. de la, & García Hernández, M. (2016). Nuevas lecturas territoriales de las actividades turístico-deportivas: huella digital y SIG. *Leituras Geográficas: Ensaios Teóricos Sobre Temas Da Contemporaneidade*, 53(9), 1689–1699. <https://doi.org/10.1017/CBO9781107415324.004>
- Mitchell, A. (2012). *The ESRI Guide to GIS Analysis: Modelling Suitability, Movement, and Interaction. Vol. 3* (Vol. 3). ESRI press.
- Moore, S., & Weiler, B. (2009). Editorial: Tourism-protected area partnerships: Stoking the fires of innovation. *Journal of Sustainable Tourism*, 17(2), 129–132. <https://doi.org/10.1080/09669580802582506>
- Moran, P. A. P. (1948). The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society*, 10(2), 243–251.
- Newsome, D., Moore, S., & Dowling, R. (2012). *Natural Area Tourism*.
- Norman, P., & Pickering, C. M. (2017). Using volunteered geographic information to assess park visitation: Comparing three on-line platforms. *Applied Geography*, 89(October), 163–172. <https://doi.org/10.1016/j.apgeog.2017.11.001>
- Norman, P., Pickering, C. M., & Castley, G. (2019). What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landscape and Urban Planning*, 185(February), 180–190. <https://doi.org/10.1016/j.landurbplan.2019.02.015>
- OAPN. (2011). *La Red de Parques Nacionales en la Sociedad*. https://www.miteco.gob.es/es/parques-nacionales-oapn/publicaciones/parques-nacionales-sociedad_tcm30-69562.pdf
- Plan Director de la Red de Parques Nacionales, 690 (2011). <http://www.boe.es>
- OAPN. (2015). *Memoria de la red de parques nacionales. 2015*. <https://www.miteco.gob.es/es/red-parques-nacionales/divulgacion/divulgacion-memorias.aspx>
- OAPN. (2018). *Parques Nacionales: límites y zonas periféricas de protección*. <https://www.miteco.gob.es/es/red-parques-nacionales/sig/parques-nacionales.aspx>
- Oku, K., Hattori, F., & Kawagoe, K. (2015). Tweet-mapping method for tourist spots based on now-tweets and spot-photos. *Procedia Computer Science*, 60(1), 1318–1327. <https://doi.org/10.1016/j.procs.2015.08.202>
- OpenStreetMap, W. (2021). *Map Features*. https://wiki.openstreetmap.org/wiki/Map_features
- Orams, M. B. (1996). Using Interpretation to Manage Nature-based Tourism. *Journal of Sustainable Tourism*, 4(2), 81–94. <https://doi.org/10.1080/09669589608667260>
- Orellana, D., Bregt, A. K., Ligtenberg, A., & Wachowicz, M. (2012). Exploring visitor movement patterns in natural recreational areas. *Tourism Management*, 33(3), 672–682. <https://doi.org/10.1016/j.tourman.2011.07.010>

- Orsi, F., & Geneletti, D. (2013). Using geotagged photographs and GIS analysis to estimate visitor flows in natural areas. *JOURNAL FOR NATURE CONSERVATION*, 21(5), 359–368. <https://doi.org/10.1016/j.jnc.2013.03.001>
- Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., Plieninger, T., Martín-López, B., Fagerholm, N., Bieling, C., & Plieninger, T. (2017). Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators*, 94(2), 74–86. <https://doi.org/10.1016/j.ecolind.2017.02.009>
- Paracchini, M. L., Zulian, G., Kopperoinen, L., Maes, J., Schägner, J. P., Termansen, M., Zandersen, M., Perez-Soba, M., Scholefield, P. A., & Bidoglio, G. (2014). Mapping cultural ecosystem services: A framework to assess the potential for outdoor recreation across the EU. *Ecological Indicators*, 45, 371–385. <https://doi.org/10.1016/j.ecolind.2014.04.018>
- Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of K in K-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 219(1), 103–119. <https://doi.org/10.1243/095440605X8298>
- Pickering, C., Walden-Schreiner, C., Barros, A., & Rossi, S. D. (2020). Using social media images and text to examine how tourists view and value the highest mountain in Australia. *Journal of Outdoor Recreation and Tourism*, 29(September 2019), 100252. <https://doi.org/10.1016/j.jort.2019.100252>
- PNSG, G. coordinadora del. (2016). *Memoria de Actividades 2016*.
- Rahm, E., & Do, H. H. (2000). Data Cleaning: Problems and Current Approaches. *Data Engineering*, 23(4), 3–13.
- Richards, D. R., & Friess, D. A. (2015). A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. *Ecological Indicators*, 53, 187–195. <https://doi.org/10.1016/j.ecolind.2015.01.034>
- Rocha, C. H. B., Fontoura, L. M., Vale, W. B. do, Castro, L. F. de P., da Silva, A. L. F., Prado, T. de O., & da Silveira, F. J. (2021). Carrying capacity and impact indicators: analysis and suggestions for sustainable tourism in protected areas—Brazil. *World Leisure Journal*, 63(1), 73–97. <https://doi.org/10.1080/16078055.2021.1888000>
- Romanillos, G., & Gutiérrez, J. (2020). Cyclists do better. Analyzing urban cycling operating speeds and accessibility. *International Journal of Sustainable Transportation*, 14(6), 448–464. <https://doi.org/10.1080/15568318.2019.1575493>
- Salas-Olmedo, M. H., Moya-Gómez, B., García-Palomares, J. C., & Gutiérrez, J. (2018). Tourists' digital footprint in cities: Comparing Big Data sources. *Tourism Management*, 66, 13–25. <https://doi.org/10.1016/j.tourman.2017.11.001>
- Sayan, M. S., & Atik, M. (2011). Koruma alanlarında rekreasyonel taşıma kapasitesi tahminleri: Termessos milli parkı. *Ekoloji*, 78, 66–74. <https://doi.org/10.5053/ekoloji.2011.7811>
- Schägner, J. P., Brander, L., Maes, J., Paracchini, M. L., & Hartje, V. (2016). Mapping recreational visits and values of European National Parks by combining statistical modelling and unit value transfer. *Journal for Nature Conservation*, 31, 71–84. <https://doi.org/10.1016/j.jnc.2016.03.001>
- Scheyvens, R. (1999). Ecotourism and the empowerment of local communities. *Tourism Management*, 20(2), 245–249. [https://doi.org/10.1016/S0261-5177\(98\)00069-7](https://doi.org/10.1016/S0261-5177(98)00069-7)
- Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). Analytics: The real-

- world use of big data: How innovative enterprises extract value from uncertain data. In *IBM Institute for Business Value*.
https://www.ibm.com/smarterplanet/global/files/se__sv_se__intelligence__Analytics_-_The_real-world_use_of_big_data.pdf
- Sessions, C., Wood, S. A., Rabotyagov, S., & Fisher, D. M. (2016). Measuring recreational visitation at U.S. National Parks with crowd-sourced photographs. *Journal of Environmental Management*, 183(3), 703–711. <https://doi.org/10.1016/j.jenvman.2016.09.018>
- Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: the first decade. *Tourism Geographies*, 18(5), 587–606. <https://doi.org/10.1080/14616688.2016.1214977>
- Shoval, N., & Isaacson, M. (2009). Tourist Mobility and Advanced Tracking Technologies. In *Routledge advances in tourism* (1st editio, Vol. 19). Routledge.
- Silverman, B. (1986). Density estimation for statistics and data analysis. *Chapman and Hall*, 37(1), 1–22. <https://doi.org/10.2307/2347507>
- Sinclair, M., Ghermandi, A., Moses, S. A., & Joseph, S. (2019). Recreation and environmental quality of tropical wetlands: A social media based spatial analysis. *Tourism Management*, 71(2751), 179–186. <https://doi.org/10.1016/j.tourman.2018.10.018>
- Sinclair, M., Ghermandi, A., & Sheela, A. M. (2018). A crowdsourced valuation of recreational ecosystem services using social media data: An application to a tropical wetland in India. *Science of the Total Environment*, 642, 356–365.
<https://doi.org/10.1016/j.scitotenv.2018.06.056>
- Sinclair, M., Mayer, M., Woltering, M., & Ghermandi, A. (2020). Using social media to estimate visitor provenance and patterns of recreation in Germany's national parks. *JOURNAL OF ENVIRONMENTAL MANAGEMENT*, 263. <https://doi.org/10.1016/j.jenvman.2020.110418>
- Somoza Medina, J. (2009). The national park concept in Spain: patriotism, education, romanticism and tourism. In W. Frost & C. M. Hall (Eds.), *Tourism and national parks: International perspectives on development, histories, and change* (Vol. 14, pp. 143–154). Routledge.
- Sonter, L. J., Watson, K. B., Wood, S. A., & Ricketts, T. H. (2016). Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLoS ONE*, 11(9).
<https://doi.org/10.1371/journal.pone.0162372>
- Statista. (2021). *Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025*. <https://www.statista.com/statistics/871513/worldwide-data-created/#:~:text=The total amount of data,replicated reached a new high.>
- Stienmetz, J. L., & Fesenmaier, D. (2016). Validating Volunteered Geographic Information: Can We Reliably Trace Visitors' Digital Footprints? *Tourism Travel and Research Association: Advancing Tourism Research Globally*. <http://scholarworks.umass.edu/ttra>
- Sumner, E. L. (1936). *Special Report on a Wildlife Study of the High Sierra in Sequoia and Yosemite National Parks and Adjacent Territory*.
- Syn, S. Y., & Oh, S. (2015). Why do social network site users share information on Facebook and Twitter? *Journal of Information Science*, 41(5), 553–569.
<https://doi.org/10.1177/0165551515585717>
- Teles da Mota, V., & Pickering, C. (2020). Using social media to assess nature-based tourism: Current research and future trends. *Journal of Outdoor Recreation and Tourism*, 30(May), 100295. <https://doi.org/10.1016/j.jort.2020.100295>

- Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., & Toivonen, T. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, 7(1), 1–11. <https://doi.org/10.1038/s41598-017-18007-4>
- The Internet Archive. (2016). *Yahoo timeline*. <https://web.archive.org/web/20080713214826/http://yhoo.client.shareholder.com/press/timeline.cfm>
- Torres-Delgado, A., & Saarinen, J. (2014). Using indicators to assess sustainable tourism development: a review. *Tourism Geographies*, 16(1), 31–47. <https://doi.org/10.1080/14616688.2013.867530>
- Tran, T. N., Wehrens, R., & Buydens, L. M. C. (2006). KNN-kernel density-based clustering for high-dimensional multivariate data. *Computational Statistics and Data Analysis*, 51(2), 513–525. <https://doi.org/10.1016/j.csda.2005.10.001>
- Troitiño, M. Á. (1995). Espacios naturales protegidos y desarrollo rural: una relación territorial conflictiva. In *Boletín de la A.G.E.N* (Vol. 20, pp. 23–37). <file:///C:/Users/39954177-A/Downloads/Dialnet-EspaciosNaturalesProtegidosYDesarrolloRural-1318470.pdf>
- UNWTO. (2021). UNWTO Tourism Definitions. In *UNWTO Tourism Definitions*. <https://www.e-unwto.org/doi/pdf/10.18111/9789284420858>
- Valentine, P. (1992). Review: Nature-based tourism. *Special Interest Tourism*, 105–127. http://link.springer.com/chapter/10.1007/978-1-4020-6799-0_8%5Cnhttp://researchonline.jcu.edu.au/1632/
- Vranešević, T., Perić, N., & Marušić, T. (2019). Perception of Social Media as a Source of Relevant Information. *Zagreb International Review of Economics and Business*, 22(1), 133–144. <https://doi.org/10.2478/zireb-2019-0016>
- Vu, H. Q., Li, G., Law, R., & Ye, B. H. (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *Tourism Management*, 46, 222–232. <https://doi.org/10.1016/j.tourman.2014.07.003>
- Wagar, J. A. (1964). The carrying capacity of wild lands for recreation. *Forest Science*, 10(7), 1–24.
- Walden-Schreiner, C., Leung, Y.-F., & Tateosian, L. (2018). Digital footprints: Incorporating crowdsourced geographic information for protected area management. *Applied Geography*, 90, 44–54. <https://doi.org/10.1016/j.apgeog.2017.11.004>
- Walden-Schreiner, C., Rossi, S. D., Barros, A., Pickering, C., Leung, Y.-F. F., Dario Rossi, S., Barros, A., Pickering, C., & Leung, Y.-F. F. (2018). Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas. *Ambio*, 47(7), 1–13. <https://doi.org/10.1007/s13280-018-1020-4>
- Watson, G. L., & Kopachevsky, J. P. (1996). *Tourist Carrying Capacity* : 2(June 1995), 169–179.
- West Sellars, R. (2009). Preserving Nature in the National Parks: A History. In *Yale University Press*. <https://doi.org/10.1086/384142>
- Wikiloc. (2017). *Wikiloc's history*. <https://es.wikiloc.com/wikiloc/about-us.do#history>
- Willemen, L., Cottam, A. J., Drakou, E. G., & Burgess, N. D. (2015). Using social media to measure the contribution of red list species to the nature-based tourism potential of african protected areas. *PLOS ONE*, 10(6), 1–14. <https://doi.org/10.1371/journal.pone.0129785>

- Wood, S. A., Guerry, A. D., Silver, J. M., & Lacayo, M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific Reports*, 3(1), 2976. <https://doi.org/10.1038/srep02976>
- Zambrano, A. M. A., Broadbent, E. N., & Durham, W. H. (2010). Social and environmental effects of ecotourism in the Osa Peninsula of Costa Rica: the Lapa Rios case. *Journal of Ecotourism*, 9(1), 62–83. <https://doi.org/10.1080/14724040902953076>
- Zeng, B., & Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism Management Perspectives*, 10, 27–36. <https://doi.org/10.1016/j.tmp.2014.01.001>
- Zeng, Z., Zhang, R., Liu, X., Guo, X., & Sun, H. (2012). Generating Tourism Path from Trajectories and Geo-Photos. In X. S. Wang, I. Cruz, A. Delis, & G. Huang (Eds.), *Web Information Systems Engineering - WISE 2012: 13th International Conference, Paphos, Cyprus, November 28-30, 2012. Proceedings* (pp. 199–212). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35063-4_15
- Zhao, B. (2020). Encyclopedia of Big Data. *Encyclopedia of Big Data*, May 2017. <https://doi.org/10.1007/978-3-319-32001-4>

12. Anexos

12. Anexos

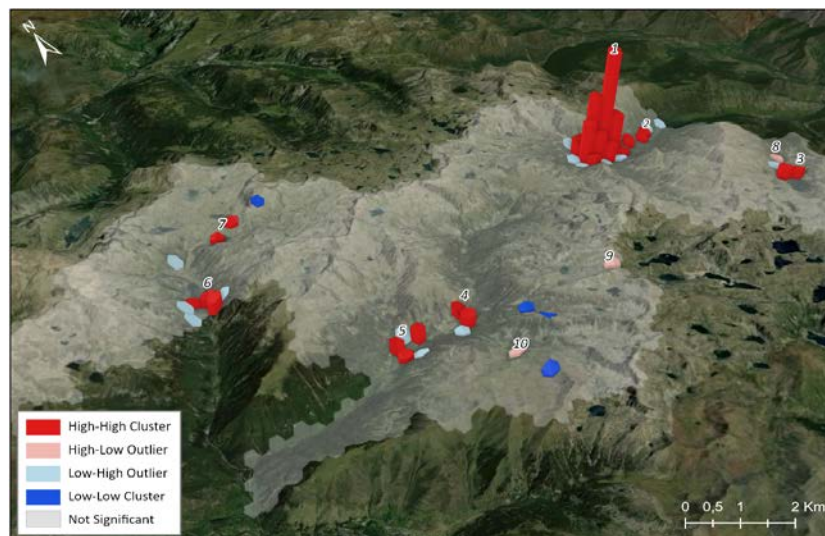
12.1. Detalle de análisis de autocorrelación espacial:

Parque Nacional Aigüestortes.

Tabla 41. Clústeres de valores altos y atípicos altos en el parque nacional Aigüestortes.

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Mirador	14	476
		Refugio		
		Punto de información		
		Estany Sant Maurici		
2	HH	Punto de interés	2	10
		Exposición de Pont Palleres		
3	HH	Refugio Josep María Blanc	2	27
		Estany Negre de Pequera		
4	HH	c	2	19
		Punto de información		
		Río de San Nicolau		
5	HH	Puente	3	22
		Cascada de Sant Esperit		
		Estany de Llebreta		
		Mirador de la cascada de Sant Esperit		
6	HH	Cabaña del Magre	3	30
		Estany de Cavallers		
		Presa de Cavallers		
7	HH	Aparcamiento	2	11
		Estany Negre		
8	HL	Refugio Joan Ventosa i Calvell	1	4
		Estany de Trescuro		
9	HL	Collada de Deliu	1	5
10	HL	S/I	1	4

Figura 87. Mapa de clústeres LISA del parque nacional Aigüestortes



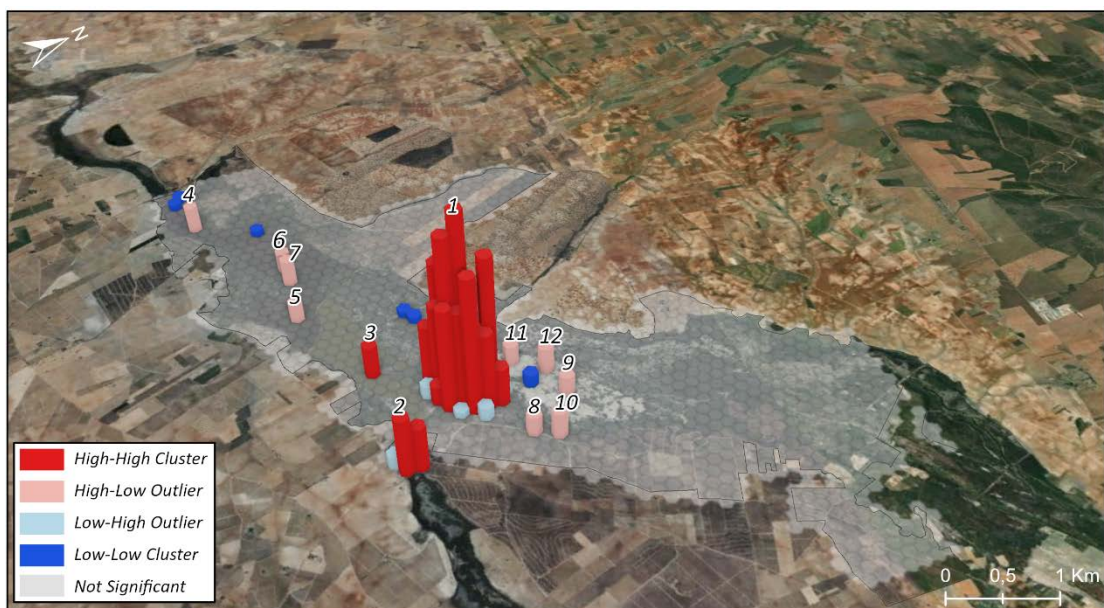
*Parque Nacional Cabañeros.**Tabla 42. Clústeres de valores altos y atípicos altos en el parque nacional Cabañeros.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Centro de visitantes Casa Palillos	4	17
2	HL	N/A	1	2
3	HL	N/A	1	2
4	HL	Centro de visitantes Torre de Abraham	1	6
		Área recreativa		
		Embalse de Torre de Abraham		
5	HL	Camino	1	2
6	HL	El risco Tirapanes	1	3
		Ruta Boquerón		
7	HL	N/A	1	2
8	HL	Pico Rocigalgo	1	4
9	HL	Carretera CM-4157	1	4
10	HL	Ruta de los Navalucillos	1	2
11	HL	Ruta del Macizo del Rocigalgo	1	2

Figura 88. Mapa de clústeres LISA del parque nacional Cabañeros.

*Parque Nacional Tablas de Daimiel.**Tabla 43. Clústeres de valores altos y atípicos altos en el parque nacional Tablas de Daimiel.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Isla de Pan	22	290
		Observatorio 3		
		Casilla Pescadores		
		Centro de visitantes		
		Bosque de Tarayes		
2	HH	Puente	2	16
		Molino de Molemocho		
		Centro de visitantes		
3	HH	Río Guadiana	1	5
4	HL	Río Guadiana	1	4
5	HL	Camino Río Guadiana	1	3
6	HL	Isla de Layos	1	3
7	HL	Isla de Layos	1	4
8	HL	Ruta la torre de prado ancho	1	3
9	HL	Río Gigüela	1	3
10	HL	Observatorio torre de prado ancho	1	4
11	HL	Río Gigüela	1	4
12	HL	Isla de los Asnos	1	3

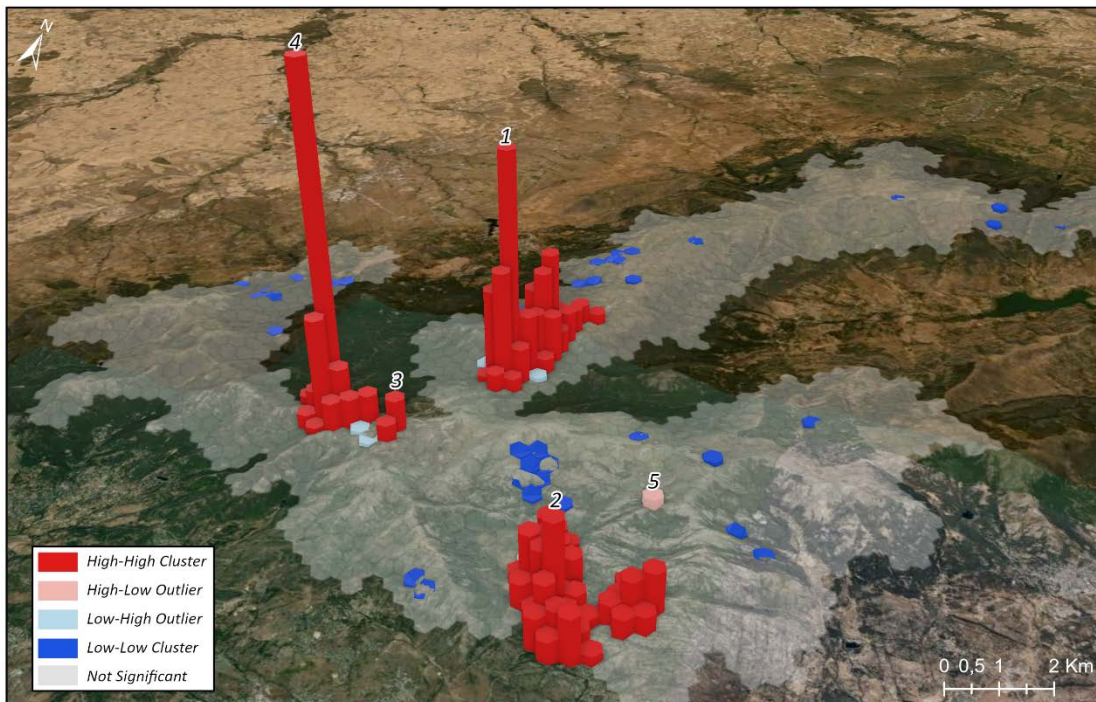
Figura 89. Mapa de clústeres LISA del parque nacional Tablas de Daimiel.

Parque Nacional Sierra de Guadarrama.

Tabla 44. Clústeres de valores altos y atípicos altos en el parque nacional Sierra de Guadarrama.

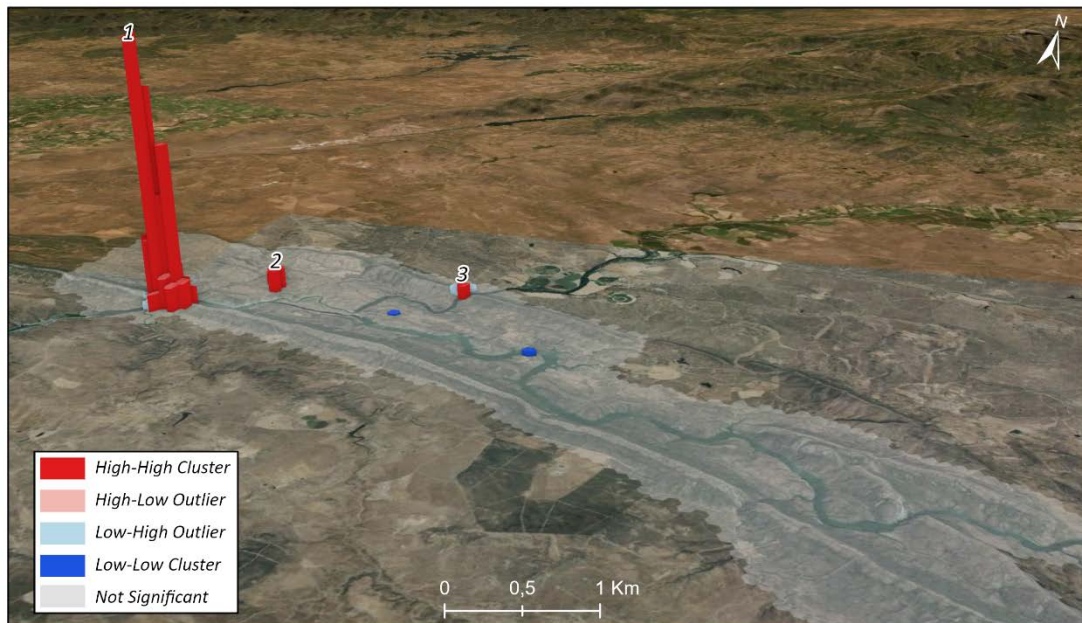
Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Centro de visitantes Peñalara	28	893
		Mirador 2		
		Los Cotos		
		Refugio 2		
		Laguna Grande Peñalara		
2	HH	La Pedriza - Cantocochino	32	731
		Mirador 3		
		Área recreativa		
		Punto de información		
		Refugio		
3	HH	Alto de las Guaramillas	2	35
		Ruta Cuerda alta		
4	HH	Puerto de Navacerrada	11	527
		Mirador		
		Refugio		
		Estación de esquí		
5	HL	Collado del Miradero	1	8

Figura 90. Mapa de clústeres LISA del parque nacional Sierra de Guadarrama.



*Parque Nacional Monfragüe.**Tabla 45. Clústeres de valores altos y atípicos altos en el parque nacional Monfragüe.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Mirador 4	14	602
		Salto del Gitano		
		Castillo		
		Ermita de Monfragüe		
		Pinturas rupestres		
2	HH	Centro de visitantes Villareal	3	32
		Oficina de uso público		
		Aula de educación		
		Área recreativa		
3	HH	Observatorio Portilla del Tietar	1	8
		Mirador		

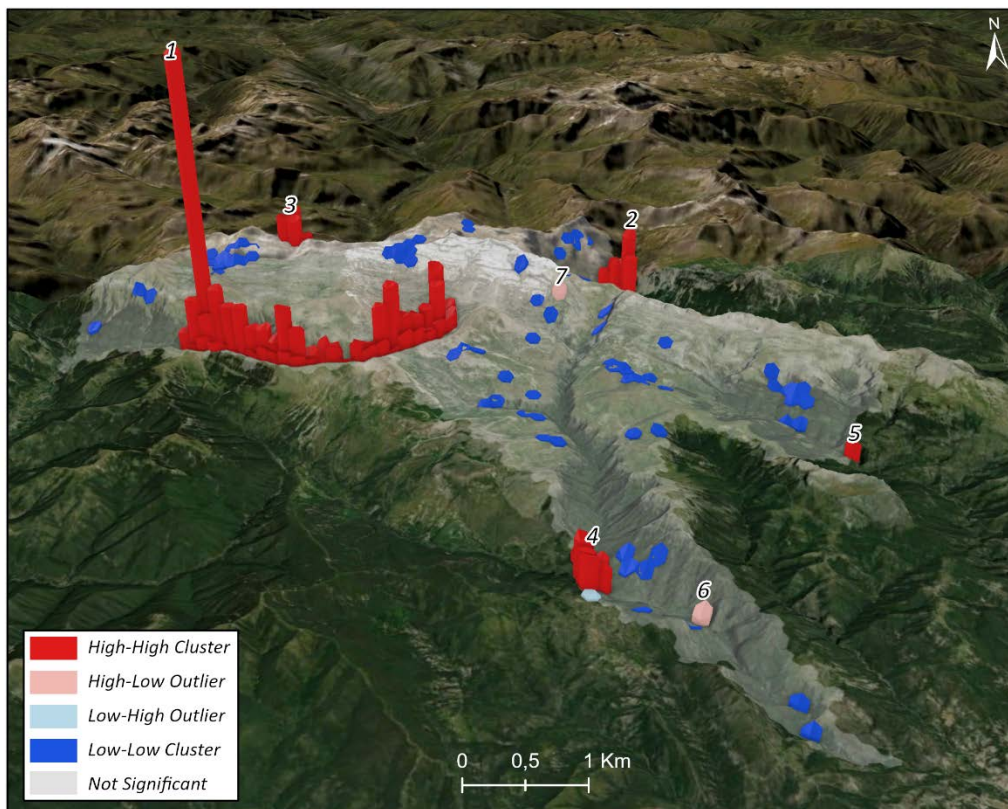
Figura 91. Mapa de clústeres LISA del parque nacional Monfragüe.

Parque Nacional Ordesa y Monte Perdido.

Tabla 46. Clústeres de valores altos y atípicos altos en el parque nacional Ordesa y Monte Perdido.

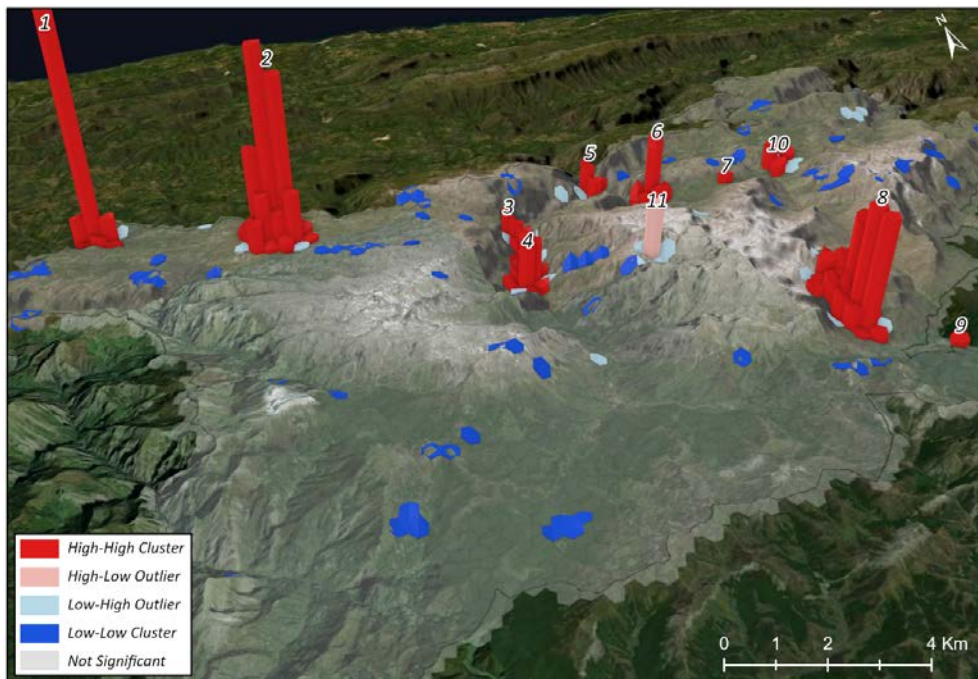
Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Valle de Ordesa	69	1780
		Punto de información		
		Mirador 8		
		Refugio 4		
		Aparcamiento 2		
		Cascada Cola de caballo		
2	HH	Valle de Pineta	8	190
		Punto de información		
		Ermita de la virgen de Pineta		
		Parador		
3	HH	Brecha de Rolando	4	86
		Frontera		
4	HH	Ermita de San Urbez de Añisclo	6	132
		Molino de Aso		
		Aparcamiento 2		
5	HH	Punto de información Escuaín	1	13
		Mirador 3		
6	HL	Mirador	1	6
7	HL	Ruta Faja de las Olas	1	6

Figura 92. Mapa de clústeres LISA del parque nacional Ordesa y Monte Perdido.



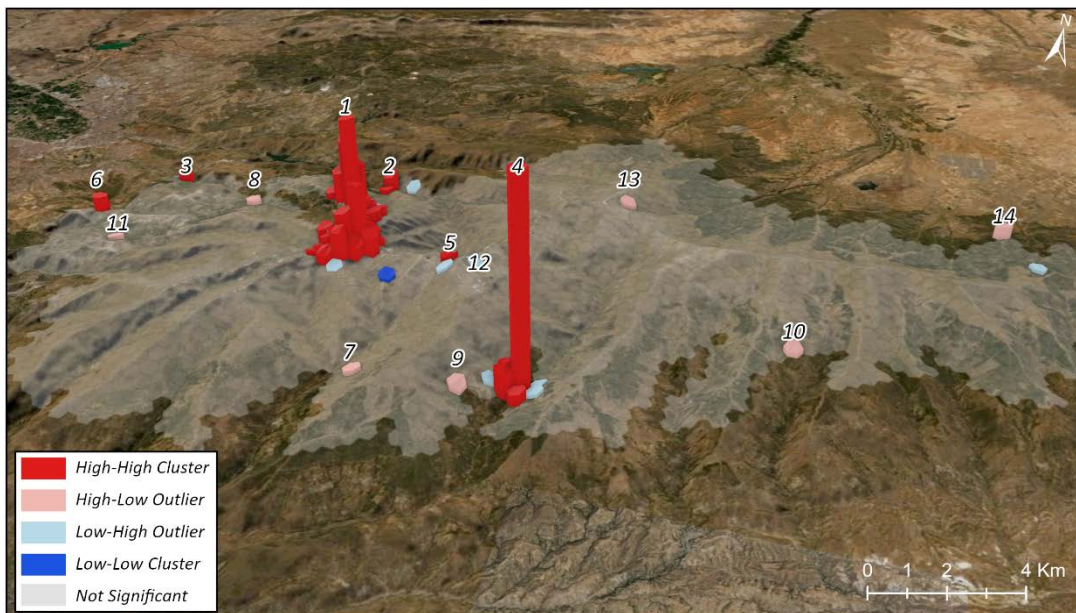
*Parque Nacional Picos de Europa.**Tabla 47. Clústeres de valores altos y atípicos altos en el parque nacional Picos de Europa.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Basílica de Covadonga	8	1496
		Mirador Cueva de Covadonga		
2	HH	Lago Enol	17	1468
		Refugio		
		Centro de visitantes		
		Mirador 2		
		Área recreativa		
		Punto de información		
3	HH	Pasarela de los Martínez	6	169
4	HH	Casa Cuevas	10	351
		Poblado Caín		
5	HH	Camarmeña	3	129
		Mirador		
6	HH	Bulnes de abajo	5	205
		Mirador 2		
7	HH	Refugio	1	20
8	HH	Teleférico Fuente Dé	17	927
		Área recreativa 2		
		Punto de información		
		Mirador		
9	HH	Espinama	1	11
10	HH	Sotres	3	82
		Área recreativa		
11	HL	Valle Hoyo Grande Bajero	1	95

Figura 93. Mapa de clústeres LISA del parque nacional Picos de Europa.

*Parque Nacional Sierra Nevada.**Tabla 48. Clústeres de valores altos y atípicos altos en el parque nacional Sierra Nevada.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Estación de esquí	32	597
		Mirador		
		Punto de información		
		Refugio		
2	HH	Ruta vereda de la Estrella	2	16
		Río Genil		
3	HH	Cueva de la Sabina	1	9
4	HH	Mirador de los Alayos	1	13
5	HH	Vivac del Mulhacén (refugio)	1	6
6	HH	Trevez	4	205
		Iglesia		
		Mirador		
7	HL	Sendero Acequias del Poqueira	1	4
8	HL	Ruta Transnevada	1	4
9	HL	Sendero Trevez - 7 lagunas	1	4
10	HL	Ruta Transnevada -Yegen	1	4
11	HL	Pico de la Virgen	1	4
12	HL	Fuente de la Higuera	1	4
13	HL	Arroyo Alhori	1	4
14	HL	Mirador del Marquesado	1	11

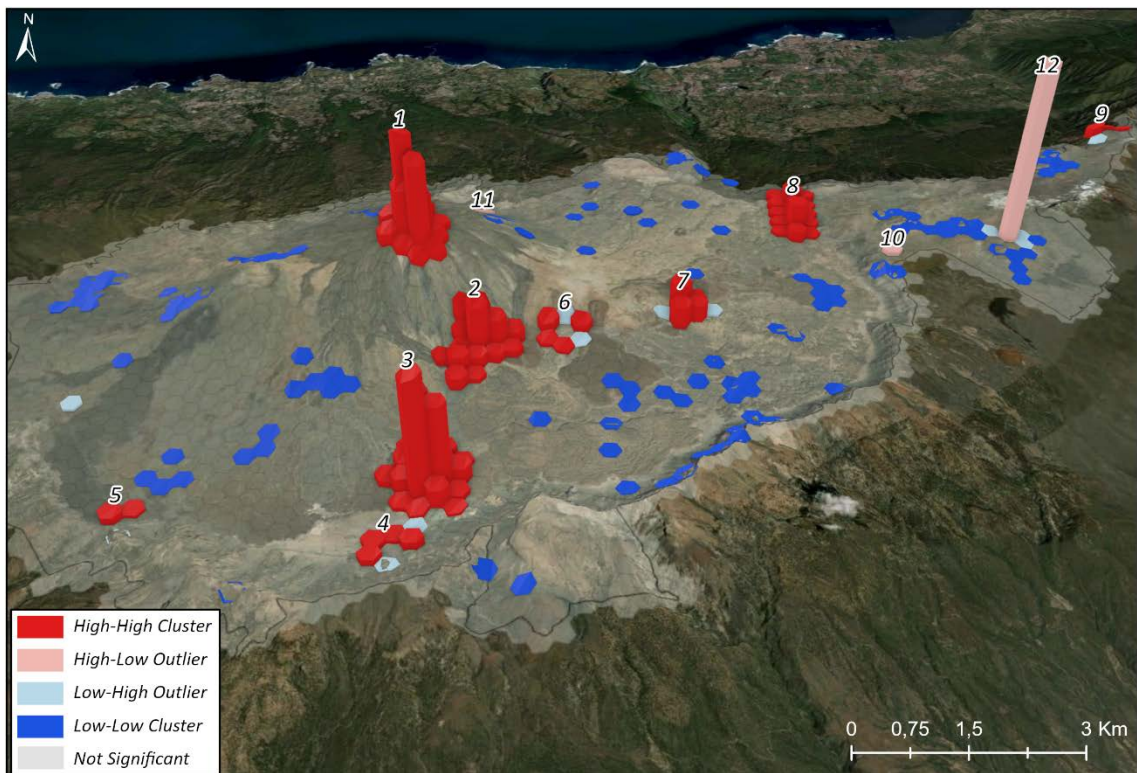
Figura 94. Mapa de clústeres LISA del parque nacional Sierra Nevada.

Parque Nacional Teide.

Tabla 49. Clústeres de valores altos y atípicos altos en el parque nacional Teide.

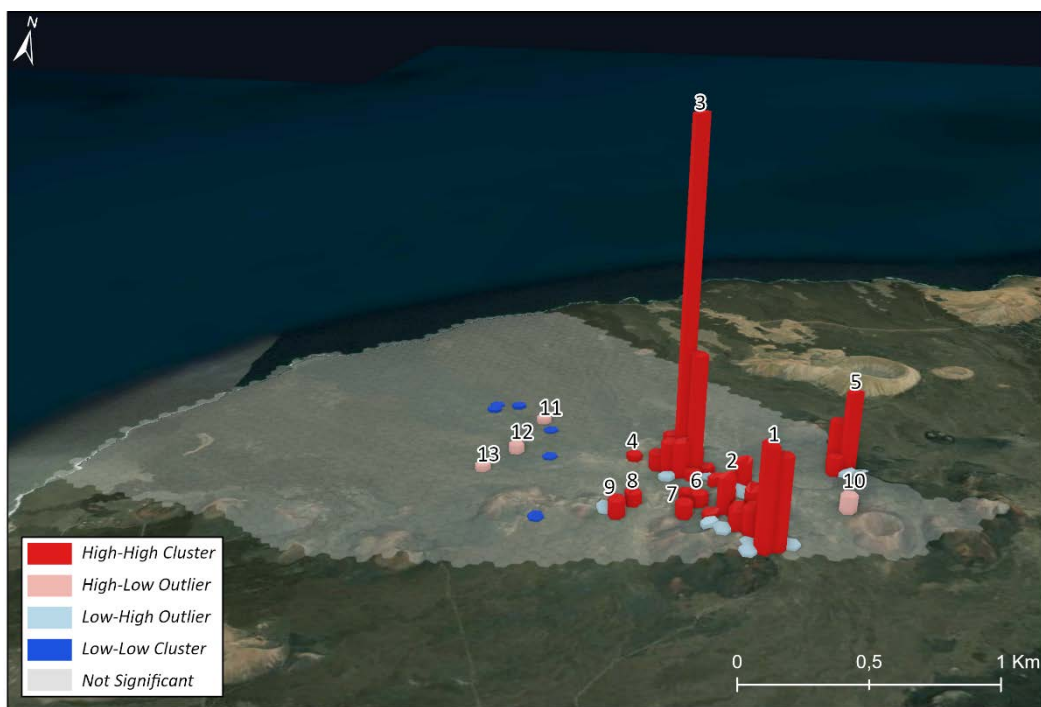
Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Pico del Teide	14	1425
		Mirador 4		
2	HH	Teleférico	16	1013
		Mirador 4		
3	HH	Roques de García	18	1805
		Centro de visitantes Cañada Blanca		
		Mirador 7		
4	HH	Mirador Llano de Ucanca	4	95
5	HH	Mirador Boca Tauce	2	39
		Punto de información		
6	HH	Mirador Tabonal Negro	4	106
7	HH	Mirador Minas San José	4	333
8	HH	Centro de visitantes El Portillo	15	662
		Mirador 2		
		Refugio		
9	HH	Mirador de El Valle	3	73
10	HL	N/A	1	19
11	HL	Montaña Abejera	1	13
12	HL	Sendero Volcán de Fasnía	1	763

Figura 95. Mapa de clústeres LISA del parque nacional Teide.



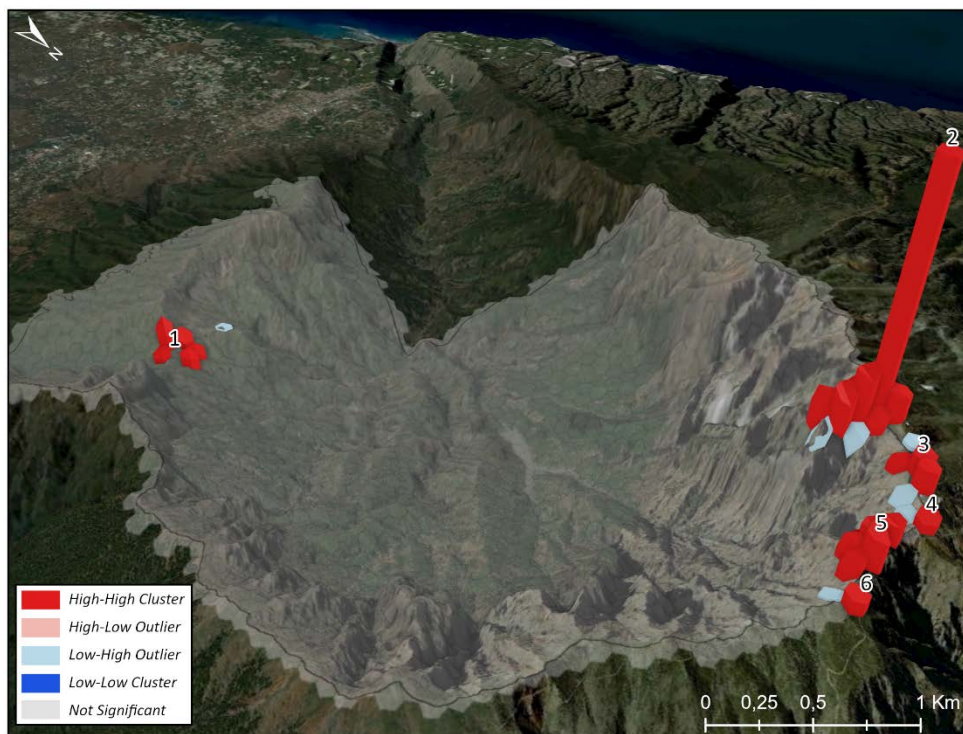
*Parque Nacional Timanfaya.**Tabla 50. Clústeres de valores altos y atípicos altos en el parque nacional Timanfaya.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Apeadero de los camellos	18	362
		Punto de información		
		Mirador		
2	HH	Ruta de los Volcanes - 4 fondos	4	33
		Caldera Boca del infierno		
3	HH	Montañas del fuego	10	408
		Punto de información		
		Mirador		
4	HH	Manto de la Virgen	1	4
5	HH	Barrera de control de acceso	4	102
6	HH	Montaña de los camellos	2	14
7	HH	Valle de la Tranquilidad	1	9
8	HH	Caldera Quemada	1	8
9	HH	Caldera Quemada Bajo	1	11
10	HL	N/A	1	11
11	HL	N/A	1	4
12	HL	N/A	1	6
13	HL	N/A	1	4

Figura 96. Mapa de clústeres LISA del parque nacional Timanfaya.

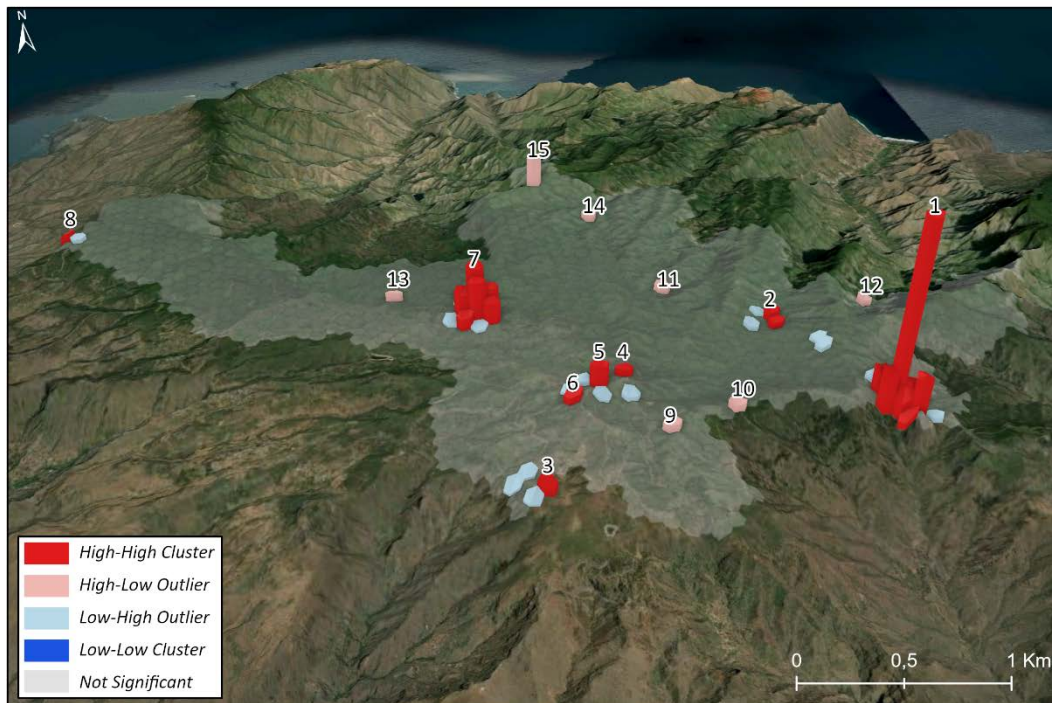
*Parque Nacional Caldera de Taburiente.**Tabla 51. Clústeres de valores altos y atípicos altos en el parque nacional Caldera de Taburiente.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Punto de información La cumbrecita	5	39
		Mirador 2		
2	HH	El Roque de los Muchachos	9	318
		Mirador 5		
		Punto de información		
3	HH	Telescopios (observatorio)	3	37
4	HH	Mirador de Taburiente	1	7
5	HH	Pared de Roberto	6	72
		Mirador		
6	HH	Camino Llano del Tauce	1	10

Figura 97. Mapa de clústeres LISA del parque nacional Caldera de Taburiente.

*Parque Nacional Garajonay.**Tabla 52. Clústeres de valores altos y atípicos altos en el parque nacional Garajonay.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Roque de Agando	9	141
		Mirador 6		
2	HH	Barranco del Cedro	2	8
3	HH	Ermita San Francisco de Asís	1	3
		Área recreativa		
4	HH	Ruta Alto del Garajonay	1	3
5	HH	3 senderos	1	9
6	HH	Mirador Alto del Garajonay	1	3
7	HH	La Laguna Grande	8	82
		Centro de visitantes		
		Área recreativa		
8	HH	Mirador de Alojera	1	4
9	HL	Ruta pajarito lajugal	1	3
10	HL	Ruta Camino Cumbres la Gomera	1	3
11	HL	N/A	1	3
12	HL	Ermita Guadalupe -ruta 15	1	3
13	HL	N/A	1	4
14	HL	Presa de Meriga	1	3
15	HL	Mirador Valle Hermoso	1	14

Figura 98. Mapa de clústeres LISA del parque nacional Garajonay.

*Parque Nacional Archipiélago de Cabrera.**Tabla 53. Clústeres de valores altos y atípicos altos en el parque nacional Archipiélago de Cabrera.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Castillo de Cabrera	11	107
		Punto de información		
		Centro administrativo		
		Cantina		
2	HL	N/A	1	3
3	HL	N/A	1	3
4	HL	Baliza de Na Foradada	1	3

Figura 99. Mapa de clústeres LISA del parque nacional Archipiélago de Cabrera.

Parque Nacional Islas Atlánticas de Galicia.

Tabla 54. Clústeres de valores altos y atípicos altos en el parque nacional Islas Atlánticas de Galicia.

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	Isla Cíes	50	1380
		Lago dos nenos		
		Centro de visitantes		
		Muelle		
		Área de fondeo 2		
		Mirador		
		Faro		
		Punto de información		
2	HH	Isla de Ons	1	10
		Centro de visitantes		
		Puesto de socorro		

Figura 100. Mapa de clústeres LISA del parque nacional Islas Atlánticas de Galicia – Isla Cíes.

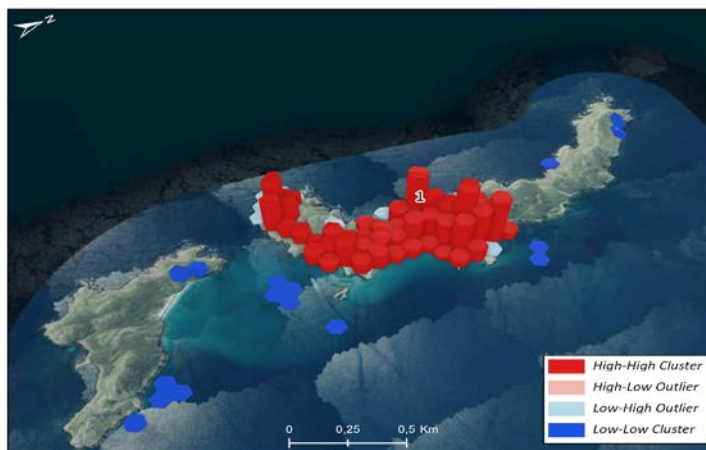


Figura 101. Mapa de clústeres LISA del parque nacional Islas Atlánticas de Galicia – Isla Ons.



*Parque Nacional Doñana.**Tabla 55. Clústeres de valores altos y atípicos altos en el parque nacional Doñana.*

Clúster	Tipo	POI asociado	Hexágonos	Usuarios
1	HH	El Rocío	13	915
		Centro de visitantes		
		Centro ornitológico		
		Charco de la boca		
2	HH	Centro de visitantes J. A. Valverde	2	12
3	HH	Playa del Coto (Matalascañas)	4	73
4	HH	Observatorio Llano de Velázquez	1	5
5	HL	N/A, Marisma del Rocío	1	4
6	HL	N/A	1	4
7	HL	N/A, Marisma del Rocío	1	8
8	HL	N/A, Marisma del Rocío	1	6
9	HL	N/A, Marisma del Rocío	1	4

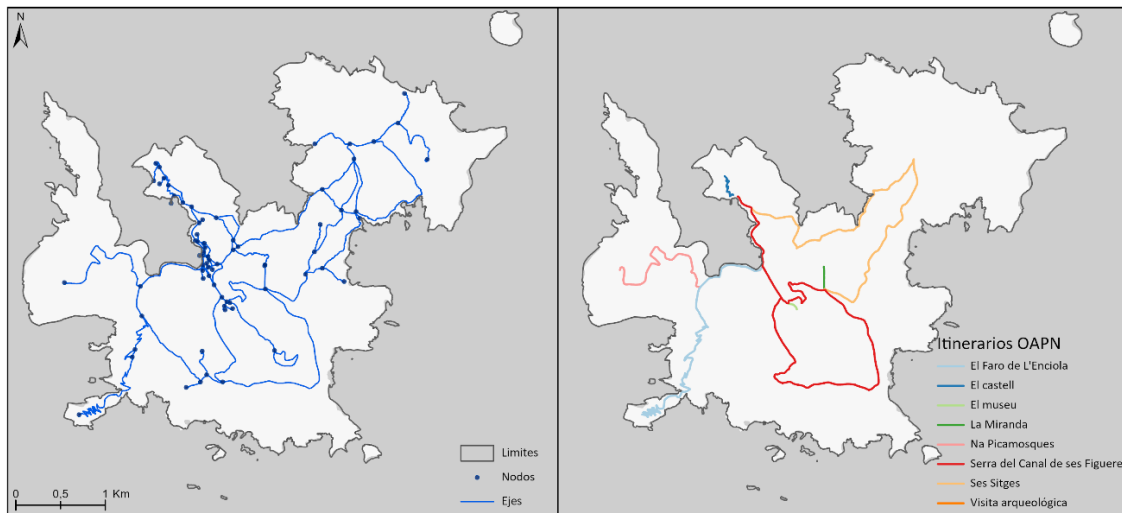
Figura 102. Mapa de clústeres LISA del parque nacional Doñana.

12.2. Descripción de las redes de senderos (grafos) y red de itinerarios de los parques nacionales:

Parque Nacional Archipiélago de Cabrera.

La red de senderos del parque nacional Archipiélago de Cabrera está compuesta por 141 ejes, de los cuales 55 (39%) forman parte de los itinerarios establecidos por el OAPN (Figura 104). La red tiene una longitud de 33,9 kilómetros y conecta todas las infraestructuras y puntos de información de uso público disponibles en la isla. Existen un total de 8 itinerarios diseñados por el OAPN repartidos en toda la superficie del parque.

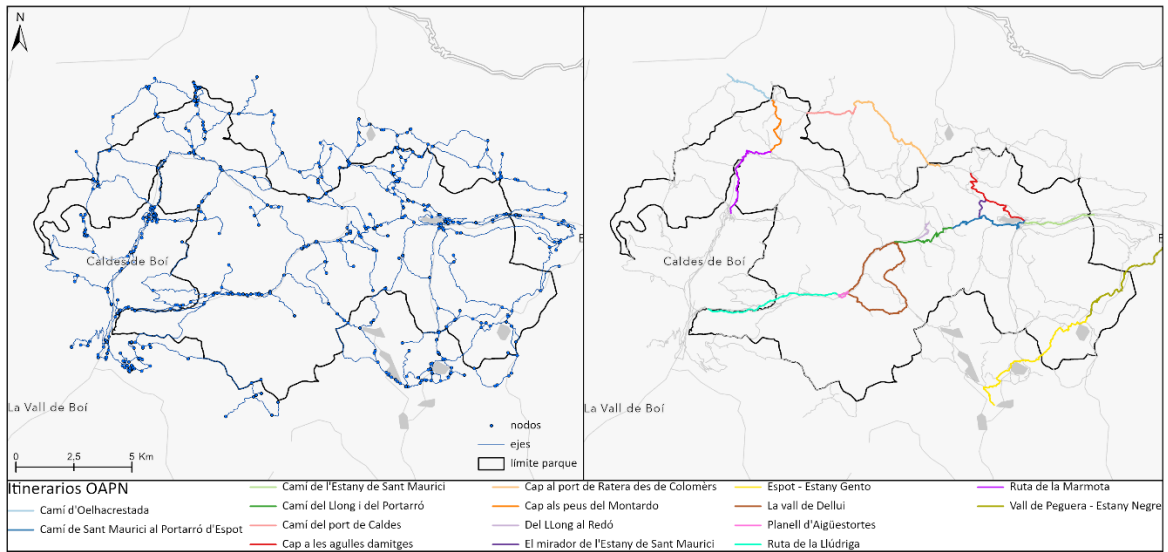
Figura 103. Red de grafos y red de itinerarios del P.N. Archipiélago de Cabrera.



Parque Nacional Aigüestortes.

La red de senderos del parque nacional Aigüestortes está compuesta por 571 nodos y 732 ejes (Figura 105), de los cuales solo el 28% (200 ejes) forman parte de los itinerarios establecidos por el OAPN. La red de grafos creada a partir de datos de Open Street Maps tiene senderos que se encuentran fuera del límite del parque y senderos que no han sido identificados en la red de senderos del parque, por ejemplo, los senderos localizados en el sector oriental del parque. La red tiene una longitud de 380,5 kilómetros y conecta todos los ingresos del parque con las diferentes instalaciones de uso público y puntos de interés disponibles. El OAPN ha establecido un total de 16 itinerarios en el parque con una duración entre 1 y 6 horas.

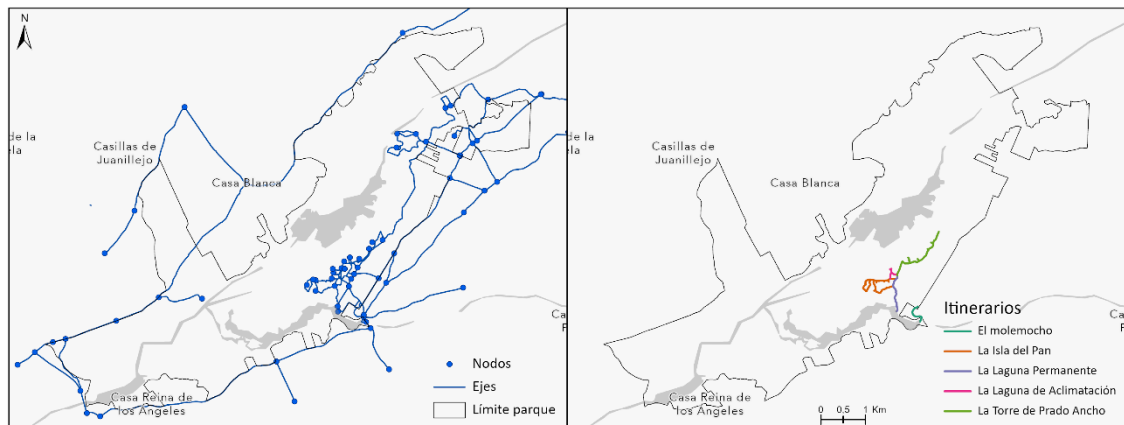
Figura 104. Red de grafos y red de itinerarios del P.N. Aigüestortes.



Parque Nacional Tablas de Daimiel.

La red de senderos del parque nacional Tablas de Daimiel está compuesta por 135 nodos y 167 ejes, de los cuales solo 49 ejes (29%) forman parte de los itinerarios establecidos por el OAPN. La red de grafos creada a partir de datos de Open Street Maps tiene senderos que se encuentran fuera del límite del parque. La red de senderos tiene una longitud de 184,8 kilómetros, que recorren los itinerarios diseñados por el OAPN y rodean los límites del parque. Los senderos que rodean el parque son utilizados para actividades de ciclismo. Existe solo un sendero dentro del parque que no forma parte de los itinerarios y se localiza en el sector norte del parque. El OAPN ha establecido un total de 5 itinerarios en el parque con una duración entre 20 minutos y 1 hora (Figura 106).

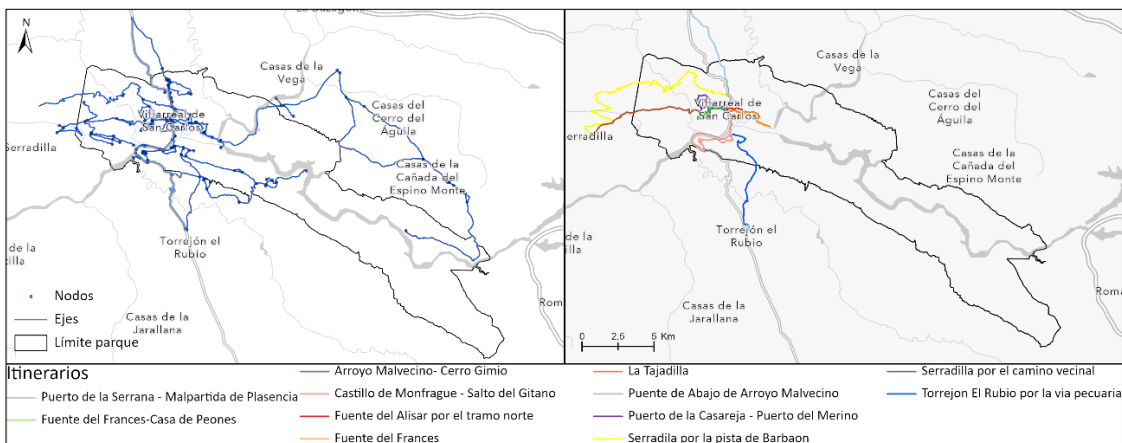
Figura 105. Red de grafos y red de itinerarios del P.N. Tablas de Daimiel.



Parque Nacional Monfragüe.

La red de senderos del parque nacional Monfragüe está compuesta por 211 nodos y 272 ejes, de los cuales 110 ejes (40%) forman parte de los itinerarios establecidos por el OAPN. La red de grafos creada a partir de datos de Open Street Maps tiene una longitud de 204,1 kilómetros y recorre los itinerarios del OPAN y senderos que atraviesan el parque. El OAPN ha establecido un total de 16 itinerarios en el parque con una duración entre 1 y 5 horas (Figura 107).

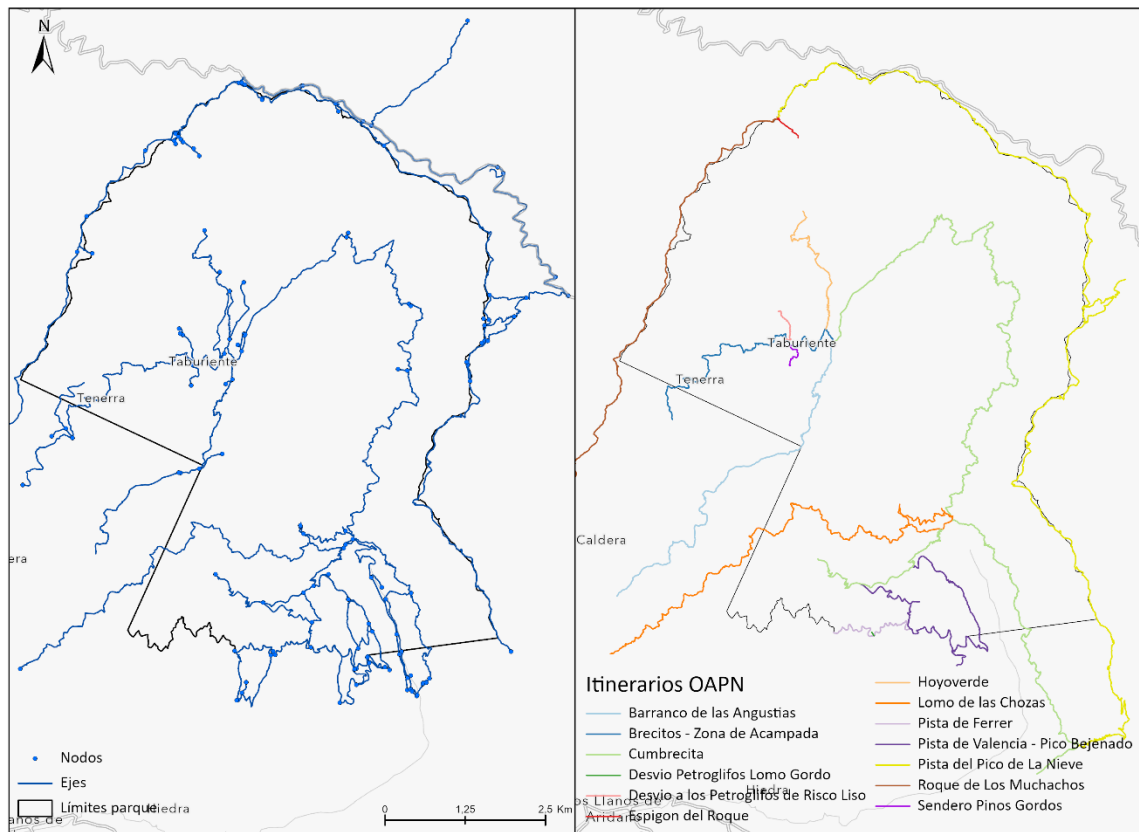
Figura 106. Red de grafos y red de itinerarios del P.N. Monfragüe.



Parque Nacional Caldera de Taburiente.

La red de senderos del parque nacional Caldera de Taburiente está compuesta por 170 nodos y 200 ejes, de los cuales 111 ejes (56%) forman parte de los itinerarios establecidos por el OAPN. La red de senderos creada a partir de los datos de Open Street Maps tiene una longitud de 119,9 kilómetros. La mayoría de los senderos de la red forman parte algún itinerario diseñado por el OAPN. El OAPN ha establecido un total de 13 itinerarios en el parque con una duración entre 1 y 7 horas (Figura 108).

Figura 107. Red de grafos y red de itinerarios del P.N. Caldera de Taburiente.



Parque Nacional Islas Atlánticas de Galicia.

La red de senderos de la Isla Cíes está compuesta por 69 nodos y 81 ejes, de los cuales 34 ejes (60,2%) forman parte de los itinerarios establecidos por el OAPN. La red de senderos creada a partir de los datos de Open Street Maps tiene una longitud de 16 kilómetros. En la Isla Cíes, el OAPN ha establecido un total de 4 itinerarios en el parque con una duración entre 40 minutos y 3 horas (Figura 109).

En la Isla Ons, la red de senderos está compuesta por 101 nodos y 196 ejes, de los cuales 118 ejes (49,3%) forman parte de los itinerarios establecidos por el OAPN. La red de senderos creada a partir de los datos de Open Street Maps tiene una longitud de 7,1 kilómetros. En la Isla Ons, el OAPN ha establecido un total de 4 itinerarios en el parque con una duración entre 45 minutos y 2 horas (Figura 110).

Figura 108. Red de grafos e itinerarios en la isla Cies.

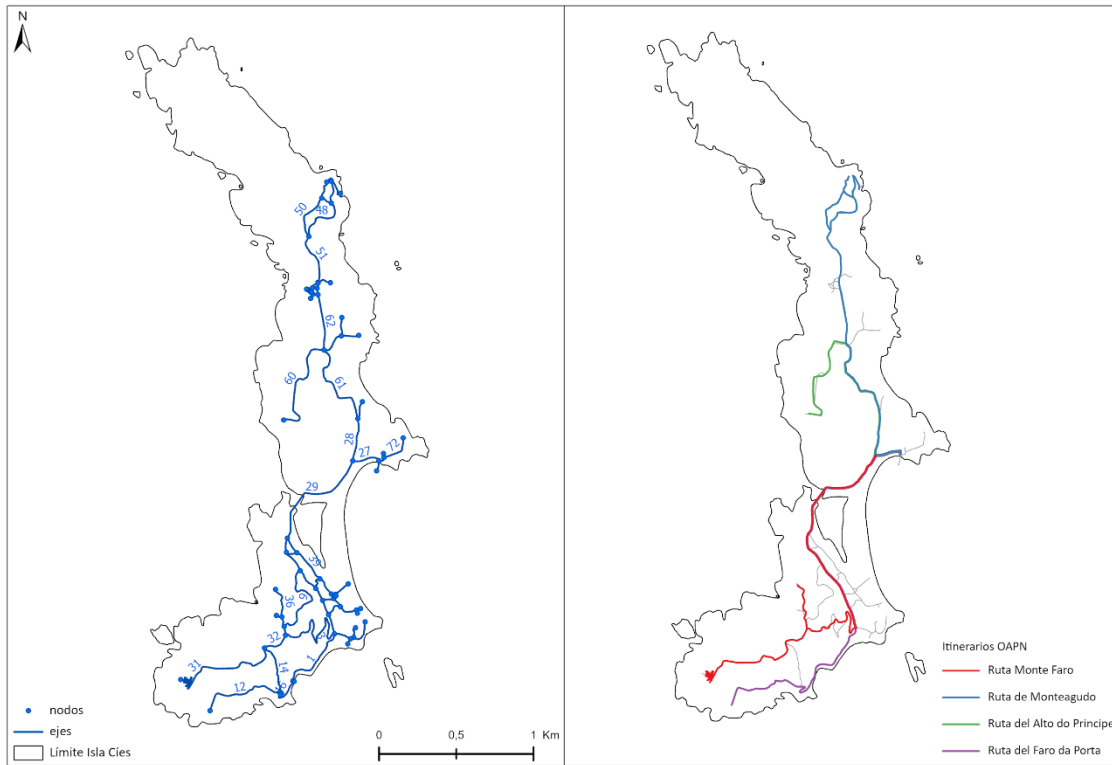


Figura 109. Red de grafos e itinerarios en la isla Ons.

