VNiVERSiDAD
D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL

DEPARTAMENTO DE INGENIERÍA CARTOGRÁFICA
Y DEL TERRENO

Tesis Doctoral

AUTOMATIZACIÓN EN PROCESAMIENTO DE DATOS LIDAR, TELEDETECCIÓN Y PROCESAMIENTO DE IMÁGENES DIGITALES

Programa de Doctorado:
Geotecnologías aplicadas a la Construcción, Energía e Industria.

José Antonio Martín Jiménez

2021

Departamento de Ingeniería Cartográfica y del Terreno

Escuela Politécnica Superior de Ávila

Universidad de Salamanca

AUTOR:

José Antonio Martín Jiménez

DIRECTORES:

Dra. Susana Lagüela López

Dr. Pablo Rodríguez Gonzálvez

2021

Automatización en procesamiento de datos LiDAR, teledetección y procesamiento de imágenes digitales

Tesis doctoral presentada por

José Antonio Martín Jiménez

Informe de los directores de Tesis

La Tesis Doctoral "*Automatización en procesamiento de datos LiDAR, teledetección y procesamiento de imágenes digitales*", realizada y presentada por José Antonio Martín Jiménez, se enmarca en las líneas de investigación de sistemas láser móvil, sistemas de información geográfica, biomasa y energía, y modelado urbano mediante LiDAR.

Los resultados obtenidos con las metodologías presentadas dieron lugar a la publicación de artículos científicos, tres de los cuales se incluyen en el presente documento. Asimismo, las herramientas desarrolladas para la aplicación de cada metodología por parte de los usuarios han sido objeto de registros de propiedad, en todos los casos a nombre de la Universidad de Salamanca. De manera paralela, durante la ejecución de esta tesis se han realizado desarrollos en el marco de proyectos de transferencia, cuya propiedad recaía en empresa, reforzando la utilidad de los avances para el mercado.

Cabe destacar la gran calidad de los artículos incluidos en esta tesis, refrendada por su publicación en revistas internacionales de elevado prestigio, sometidas a un proceso de revisión anónimo por pares. Estos, por orden de calidad de la publicación, se distribuyen del siguiente modo:

- Un artículo en la revista "ISPRS Journal of Photogrammetry and Remote Sensing", indexada en el Journal Citation Report en la 1ª posición dentro de la categoría Geociencia, en la que hay 50 revistas (1º cuartil).
- Un artículo en la revista "Renewable & Sustainable Energy Reviews", indexada en el Journal Citation Report, en la categoría Ciencia y Tecnología Verde y Sostenible, en la que ocupa la 1ª posición de un total de 44 revistas (1º cuartil).
- Un artículo en la revista "Automation in Construction", con indexación en el Journal Citation Report, categoría Tecnología de la Construcción y Edificación, ocupando la posición 3 de 67 revistas (1º cuartil).

Asimismo, el doctorando ha realizado una estancia de investigación, con el objetivo de cumplir los requisitos para la obtención de la Mención Internacional al título de Doctor. Los datos de la estancia son los siguientes:

- Department of Civil Engineering. University of Minho (Portugal), desde Junio 2021 hasta Septiembre 2021 (3 meses). Desarrollo de aplicaciones geoinformáticas y desarrollo de rutinas para la generación de BIM (Building Information Models).

Dadas las condiciones mencionadas, se considera que la presente tesis se ajusta, de modo óptimo, a las condiciones requeridas para la presentación de la misma por la modalidad de "Compendio de publicaciones", conforme los requisitos expuestos en el Reglamento de Doctorado de la Universidad de Salamanca, así como es candidata a la "Mención Internacional". La calidad de las metodologías desarrolladas, así como la validación de su uso mediante la aplicación a casos de estudio reales queda irrefutablemente reconocida después de su aceptación en los ámbitos internacionales de ingeniería civil, tecnologías de la construcción, energías renovables y geoinformática.

Por todos los motivos mencionados, esta tesis es favorable para obtener el grado de Doctor por la Universidad de Salamanca en el programa de doctorado en Geotecnologías Aplicadas a la Construcción, Energía e Industria.

Lo que firman, a todos los efectos oportunos, en Ávila, a 9 de Noviembre de 2021.

LAGÜELA LOPEZ SUSANA - 53186464E

Firmado digitalmente por LAGÜELA LOPEZ SUSANA - 53186464E
Fecha: 2021.11.29 09:55:14 +01'00'

Dra. Susana Lagüela López

RODRIGUEZ GONZALVEZ PABLO - 71769960R

Firmado digitalmente por RODRIGUEZ GONZALVEZ PABLO - 71769960R
Fecha: 2021.11.29 16:37:42 +01'00'

Dr. Pablo Rodríguez Gonzálvez

Listado de artículos publicados

La Tesis Doctoral que se presenta en este documento, está constituida por un compendio de tres artículos científicos. Estos artículos han sido publicados en revistas internacionales de alto impacto. Se enumeran a continuación estas publicaciones.

1. **Road safety evaluation through automatic extraction of road horizontal alignments from Mobile LiDAR System and inductive reasoning based on decision tree.**

   José Antonio Martín-Jiménez[1], Santiago Zazo[2], José Juan Arranz Justel[3], Pablo Rodríguez-Gonzálvez[4], Diego González-Aguilera[1]

   [1]University of Salamanca, Department of Cartographic and Land Engineering, Hornos Caleros, 50, 05003 Ávila, Spain
   [2]University of Salamanca, TIDOP Research Group, Hornos Caleros, 50, 05003 Ávila, Spain
   [3]Technical University of Madrid, Department of Topographic Engineering and Cartography, Camino de la Arboleda, s/n Campus Sur, Autovía de Valencia Km 7, 28031 Madrid, Spain
   [4]Universidad de León, Department of Mining Technology, Topography and Structures, Astorga, s/n, 24401, Ponferrada, León, Spain

   ISPRS Journal of Photogrammetry and Remote Sensing, Diciembre 2018.

   DOI: 10.1016/j.isprsjprs.2018.10.004

2. **Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity**

   J. Martín-Jiménez, S. Del Pozo, M. Sánchez-Aparicio, S. Lagüela
   Department of Cartographic and Land Engineering, University of Salamanca, Hornos Caleros, 50, Ávila 05003, Spain

   Automation in Construction, Enero 2020.

   DOI: 10.1016/j.autcon.2019.102965

3. A geospatial web-based platform for the photovoltaic potential computation: Ener3DMap-SolarWeb Roofs

M. Sánchez-Aparicio, J. Martín-Jiménez, S. Del Pozo, E. González-González, S. Lagüela
Department of Cartographic and Land Engineering, University of Salamanca, Hornos Caleros, 50, Ávila 05003, Spain

Renewable and Sustainable Energy Reviews, Enero 2021.

DOI: 10.1016/j.rser.2020.110203

“La juventud de un ser

no se mide por los años que tiene,

sino por la curiosidad que almacena.”

(Salvador Paniker)

Resumen

Las herramientas de captura de información han ido evolucionando para permitir cada vez mayor precisión en la digitalización de la realidad. Los equipos de medición LiDAR (Light Detection And Ranging) proporcionan nubes de puntos 3D masivas con información geométrica de escenarios complejos. Los vuelos fotogramétricos permiten obtener ortoimágenes aéreas digitales, georreferenciadas y cada vez de mayor resolución. Se puede hablar de Big Data geoespacial para referirnos a esta ingente cantidad de información que se necesita procesar para dotarla de propiedades semánticas según los puntos o los pixeles capturados se relacionan entre sí. En esta Tesis Doctoral, se desarrollan metodologías y algoritmos que automatizan el análisis y procesamiento de estos volúmenes extremadamente grandes de datos, para dotarlos de propiedades semánticas que facilitan su manejo. La investigación realizada aporta soluciones al procesamiento de la información obtenida en la digitalización 3D de infraestructuras viarias y edificios, centrándose en avanzar hacia los ODS (Objetivos de Desarrollo Sostenible) de la Agenda 2030 especialmente en el uso de energías renovables (ODS 7 – energía asequible y no contaminante; ODS 11 – ciudades y comunidades sostenibles) y en el plan de Visión 0 (0 muertos, 0 heridos, 0 atascos y 0 emisiones) para fomentar la seguridad vial y el desarrollo sostenible en movilidad y transporte. La investigación científica se enmarca de manera directa dentro de los proyectos *InRoad: Sistema integral para la prevención y la asistencia al rescate en accidentes de tráfico*, financiado por el Ministerio de Energía, Turismo y Agenda Digital y *Ener3DMap: Gestión de Clientes y recursos energéticos distribuidos: mapeado energético*, dentro de la Cátedra Iberdrola-USAL VIII Centenario, aunque el doctorando ha realizado desarrollos en otros numerosos proyectos de investigación durante la evolución de la tesis.

El primer reto de esta investigación se centra en automatizar el procesamiento de nubes de puntos de una infraestructura viaria capturados con un sistema de cartografiado móvil MLS (Mobile LiDAR System). Se busca conseguir la extracción del eje de la carretera para obtener la alineación en planta, con sus tramos curvos, rectos y curvas de transición. A partir de esta información se pretende calcular tres parámetros de estabilidad y se asigna un índice de consistencia geométrica y un nivel de seguridad para cada tramo. Con ello, se puede disponer de una herramienta preventiva para extremar las precauciones en tramos de mayor peligrosidad. Este procesamiento automático se desarrolla en la aplicación informática *inRoad in Alert* (apéndice B).

El siguiente reto que se quiere afrontar, es combinar el procesado de nubes de puntos de LiDAR aerotransportado y ortoimágenes digitales, procedentes de vuelos fotogramétricos realizados por el Instituto Geográfico Nacional, para caracterizar geométricamente las aguas de los tejados de edificios, calculando sus inclinaciones, orientaciones y superficies. De igual forma, se procede a calcular su potencial solar fotovoltaico. Combinando estos dos procesados se fortalece la debilidad que presentan los datos LiDAR disponibles en cuanto al cálculo de la superficie de tejados, con la información que se puede obtener de las ortoimágenes. Este procesado se desarrolla en el software *Ener3dmap-Solar Roofs* (apéndice B).

La información de salida de este procesado permite generar una capa con los datos geométricos de los tejados, para la aplicación de mapas web desarrollada *Ener3DMap-SolarWeb Roofs* (apéndice B). Esta herramienta utiliza la librería para mapas Leaflet, integra esta capa con mapas base, datos catastrales y el modelo de radiación solar validado PVGIS (Photovoltaic Geographical Information System). Este desarrollo permite calcular la producción solar de una instalación de paneles fotovoltaicos de una forma ágil, proporcionando datos de producción anual, mensual y diaria, tanto para un tejado de un edificio existente, como en una parcela sin edificación, o un edificio proyectado.

Adicionalmente se programa otra infraestructura de datos espaciales *SolarWeb Cities* (apéndice B), con la misma base de funcionamiento que la aplicación anterior, para realizar una prospectiva de potencial solar fotovoltaico a nivel de barrio o ciudad.

Summary

Data acquisition tools have evolved towards greater precision in the digitization of reality. LiDAR (Light Detection And Ranging) measurement equipment provides massive 3D point clouds with geometric information of complex scenarios. Photogrammetric flights make it possible to obtain digital, geo-referenced and increasingly higher resolution aerial orthoimages. The term "Geospatial Big Data" refers to the huge amount of information that needs to be processed in order to assign the semantic properties that relate the points or pixels acquired. This Doctoral Thesis shows the development of methodologies and algorithms for the automation of the analysis and processing of these extremely large volumes of data, to provide them with semantic properties that facilitate their management. The research performed provides solutions to the processing of the information obtained in the 3D digitalization of road infrastructures and buildings, focusing on advancing towards the SDGs (Sustainable Development Goals) of the 2030 agenda especially in the use of renewable energies (SDG 7 - affordable and clean energy; SDG 11 - sustainable cities and communities) and the Vision 0 Plan (0 deaths, 0 injuries, 0 traffic jams and 0 emissions) to promote road safety and sustainable development in mobility and transport. The scientific research is directly framed within the projects *InRoad: Comprehensive system for prevention and rescue assistance in traffic accidents*, funded by the Ministry of Energy, Tourism and Digital Agenda and *Ener3DMap: Management of customers and distributed energy resources: energy mapping*, within the Iberdrola-USAL VIII Centenary Chair, although the PhD student has made developments in numerous other research projects during the evolution of the thesis.

The first challenge of this research focuses on automating the processing of point clouds of a road infrastructure acquired with a mobile mapping system MLS (Mobile LiDAR System). The aim is to extract the road axis to obtain the alignment in plan, with its curved and straight sections, and transition curves. This information is used to calculate three stability parameters and assign a geometric consistency index and a safety level for each section. With this, a preventive tool is available to take extreme precautions in the most dangerous sections. This automatic processing is carried out in the *inRoad in Alert* software application (Appendix B).

The next challenge to be faced is the combination of the processing of airborne LiDAR point clouds and digital orthoimages, coming from photogrammetric flights performed by the National Geographic Institute, to geometrically characterize the slopes of

building roofs, calculating their inclinations, orientations and surfaces. Similarly, their photovoltaic solar potential is calculated. Combining these two processes strengthens the weakness of the available LiDAR data in terms of calculating the roof surface, with the information that can be obtained from orthoimages. This processing is developed in the *Ener3dmap-Solar Roofs* software (appendix B)

The output information from this processing is used for the generation of a layer with the geometric data of the roofs, for the web mapping application developed *Ener3DMap-SolarWeb Roofs* (appendix B). This tool uses the Leaflet map library, integrates this layer with base maps, cadastral data and the validated solar radiation model PVGIS (Photovoltaic Geographical Information System). This development allows to calculate the solar production of a photovoltaic panel installation in an agile way, providing annual, monthly and daily production data, both for an existing building roof, as well as for a plot without building, or a projected building.

In addition, another spatial data infrastructure *SolarWeb Cities* (appendix B) is programmed, with the same operating basis as the previous application, to perform a prospective of solar photovoltaic potential at neighborhood or city level.

Agradecimientos

Esta Tesis Doctoral, es un reto al que llego después de más de veinte años en el mundo laboral. Volver al mundo de la investigación tras esa etapa requiere un importante cambio de enfoque. El entorno de investigación del grupo Tidop, ha creado un clima propicio, que ha favorecido esa labor, todos los compañeros han hecho que ese cambio haya sido muy fructífero. Gracias a todos los compañeros del grupo por ayudarme a disfrutar de este periodo. Quiero dar las gracias también a los compañeros de la Cátedra Iberdrola-USAL VIII Centenario, Susana Lagüela, Susana del Pozo, María, Enrique y Paula, sois el mejor equipo que alguien pueda desear. Gracias a todos los coautores de los artículos, esta Tesis lleva mucho de todos vosotros. Gracias a todos los compañeros de docencia en el departamento de Ingeniería Cartográfica y del Terreno.

Y como no, gracias a los directores de Tesis, Diego en la primera etapa y Susana y Pablo en todo el proceso.

Por último, gracias a mi familia, en especial a T, ya que sin su apoyo esta Tesis no hubiese sido posible.

Contenido

Capitulo I

Introducción

1. Introducción

En 2015 las Naciones Unidas adoptaron la Agenda 2030 para alcanzar los Objetivos de Desarrollo Sostenible (ODS) [1], a partir de los cuales se fijaron los objetivos de la Unión Europea, en términos de descarbonización de la red eléctrica (10% de energías renovables en la red eléctrica en España en 2050) y del transporte (impulso al hidrógeno verde), así como en términos de seguridad en la movilidad y transporte para el plan de Visión 0 (0 muertos, 0 heridos, 0 atascos y 0 emisiones) para 2050 [2]. Las dos iniciativas están alineadas, ya que el plan de Visión 0 encaja en los ODS *3-Salud y bienestar* y *9-Industria, innovación e infraestructura*. En concreto, las dos iniciativas tienen retos en la misma dirección: *3.2- Reducir a la mitad el número de muertes y lesiones causadas por accidentes de tráfico en el mundo*; *9.1 Desarrollar infraestructuras fiables, sostenibles, resilientes y de calidad, incluidas infraestructuras regionales y transfronterizas, para apoyar el desarrollo económico y el bienestar humano, haciendo especial hincapié en el acceso asequible y equitativo para todos*. Para contribuir a investigar en la consecución de estos retos, la presente Tesis Doctoral se alinea, además de con los ODS ya mencionados, con: 7-Energía asequible y no contaminante, 11-Ciudades y comunidades sostenibles, y 13-Acción por el clima. Se busca contribuir a avanzar en la consecución de las metas:

> 3.6. Reducir a la mitad el número de muertes y lesiones causadas por accidentes de tráfico en el mundo.
>
> 9.1. Desarrollar infraestructuras fiables, sostenibles, resilientes y de calidad, incluidas infraestructuras regionales y transfronterizas, para apoyar el desarrollo económico y el bienestar humano, haciendo especial hincapié en el acceso asequible y equitativo para todos.
>
> 7.1. Garantizar el acceso universal a servicios energéticos asequibles, fiables y modernos.
>
> 7.2. Aumentar considerablemente la proporción de energía renovable en el conjunto de fuentes energéticas.
>
> 11.2. Proporcionar acceso a sistemas de transporte seguros, asequibles, accesibles y sostenibles para todos y mejorar la seguridad vial, en particular mediante la ampliación del transporte público, prestando especial atención a las necesidades de las personas en situación de vulnerabilidad, las mujeres, los niños, las personas con discapacidad y las personas de edad.
>
> 13.2. Incorporar medidas relativas al cambio climático en las políticas, estrategias y planes nacionales.

Para lograr alcanzar los retos de la Agenda 2030 se necesita la participación de toda la sociedad: administraciones, universidades, centros de investigación, empresas, y usuarios. Fruto del apoyo a la investigación en estas líneas, la presente Tesis Doctoral se enmarca dentro del proyecto *InRoad: Sistema integral para la prevención y la asistencia al rescate en accidentes de tráfico*, financiado por el Ministerio de Energía, Turismo y Agenda Digital y del proyecto *Ener3DMap: Gestión de Clientes y recursos energéticos distribuidos: mapeado energético* dentro de la Cátedra Iberdrola-USAL VIII Centenario. Los dos proyectos requieren la automatización en el procesamiento de datos procedentes de la digitalización de infraestructuras.

Durante las últimas décadas, los avances geomáticos en el Sistema Global de Navegación por Satélite (GNSS), sistemas de detección de luz y rango (Light Detection And Ranging, LiDAR), adquisición de información radiométrica, junto con la hibridación entre sensores, técnicas y plataformas móviles pasivo-activas, han revolucionado la adquisición masiva de datos geoespaciales, así como las técnicas de inspección y seguimiento. Todos estos avances permiten la digitalización en 3D de infraestructuras con precisión, tanto de infraestructuras viarias con un sistema de cartografiado LiDAR móvil (MLS) [3-6], como la de edificios y parcelas a partir tanto de MLS como de sistemas aerotransportados [7, 8].

Además de la mejora en las prestaciones de los equipos de adquisición de datos, la directiva INSPIRE (INfrastructure for Spatial InfoRmation in Europe) [9] establece las reglas que deben cumplir los datos geoespaciales, que los Estados miembros de la Unión Europea, ponen a disposición de los ciudadanos, para que la información espacial y geográfica contribuya al desarrollo sostenible. En el caso de España, el Instituto Geográfico Nacional (IGN) facilita estos datos en abierto para su descarga libre [10]. Los datos disponibles son nubes de puntos LiDAR, adquiridas con un sistema aerotransportado, con el relieve del territorio y de las infraestructuras y edificios existentes sobre él, y ortoimágenes aéreas procedentes de vuelos fotogramétricos. Otros países de la Unión Europea, tienen organismos similares para facilitar la investigación hacia el desarrollo sostenible [11]. La utilización de esta información libre supone un avance en los estudios científicos, ya que pone a disposición de los investigadores, información geoespacial sin necesidad de dedicar recursos propios para su captura.

Tanto si se realiza la adquisición de datos de las infraestructuras con medios propios, como si se utilizan datos de las mismas ofrecidos en abierto, el volumen de información a tratar es muy grande. Y por ello es necesario avanzar en la automatización innovadora del procesado de este Big Data geoespacial, para permitir dotar a los datos de propiedades semánticas, de un modo eficiente y eficaz. Los equipos informáticos con multiprocesador y las aplicaciones informáticas con algoritmos de procesado multihilo, proporcionan capacidad suficiente para realizar este procesado [12].

1.1. Sistemas de Cartografiado Móvil para captura de infraestructuras viarias.

Anualmente mueren 1,20 millones de personas y entre 20 y 50 millones de personas resultan heridas en accidentes de tráfico [13]; además, los accidentes de tráfico implican costes sociales importantes y elevados [14]. La influencia del factor humano en estos accidentes es conocida [15], así como el hecho de que el diseño de la vía también es responsable de que ocurran una parte de los mismos [16]. Alrededor del 30% de los accidentes de tráfico son atribuibles a infraestructuras y el suceso de estos siniestros se centra en tramos de carretera específicos [13,17].

En las infraestructuras ya en servicio, el planteamiento de seguridad seguido es el expuesto en el European Road Assessment Programme (EuroRAP). En él se obtiene un Índice de Riesgo (IR) de la carretera, en base a la estadística de accidentalidad y la intensidad de tráfico (IMD-Intensidad Media Diaria), que se complementa mediante un protocolo conjunto de inspección (con imágenes) y de puntuación de la seguridad (mediante estrellas), que valora la seguridad pasiva de la carretera [18]. En este planteamiento, los parámetros geométricos del trazado involucrados en la puntuación de la seguridad de la vía son seleccionados de una forma muy genérica. Por tanto, se plantea la mejora en la selección de los parámetros geométricos de influencia y en su determinación, con el objetivo de obtener información geométrica precisa de las vías.

Poniendo el foco en los objetivos de movilidad y transporte del plan de Visión 0, es necesario desarrollar metodologías y algoritmos para automatizar el procesado de información de infraestructuras viarias para avanzar en la construcción y mantenimiento de las carreteras, buscando carreteras seguras que ayuden a reducir los accidentes. Además de analizar las características orográficas de las mismas para estudiar la viabilidad del uso de vehículos eléctricos que permitan reducir las emisiones.

Para digitalizar estas vías, el dispositivo más adecuado, debido a la posibilidad de montarlo en un vehículo convencional, a la velocidad en la adquisición de datos, y a que permite obtener precisiones centimétricas, es un Sistema de cartografiado móvil (MLS), que se compone principalmente de un sistema de navegación y uno o más sensores LiDAR. El MLS puede completarse con sistemas fotográficos (cámaras digitales RGB) y / u otros sensores como una cámara termográfica, un radar de penetración terrestre (GPR) y un perfilómetro. El sistema de navegación está integrado por un conjunto de sensores, tales como el GNSS, la Unidad de Medición Inercial (IMU - Inertial Measurement Unit) y el Indicador de Medición de Distancia (DMI - Distance Measuring Instrument) [19]. Además de eso, el MLS es un sistema preciso y eficiente para la adquisición de datos en entornos complejos, como corredores urbanos y viales en grandes áreas; lo que proporciona una importante reducción del tiempo de recopilación y procesamiento de datos [19-21].

En el campo de la ingeniería civil, el MLS comienza a consolidarse entre investigadores e ingenieros. Recientemente, se han realizado considerables esfuerzos y avances principalmente en tareas como la segmentación de nubes de puntos, la extracción de alineamientos de carreteras y la detección automática de objetos. En [20] se propuso un método novedoso para la eliminación automática de vehículos de conjuntos de datos LiDAR móviles. En [19] se desarrolló un método basado en la segmentación, parametrización y filtrado de nubes de puntos LiDAR de MLS para extraer, semiautomáticamente, la línea central de la carretera y determinar los parámetros de la carretera horizontal y su alineación (es decir, líneas rectas, arcos circulares y clotoides). Por su parte, en [22] se describe un algoritmo para la detección automática de pasos de cebra mediante la Transformada Hough estándar, aplicada sobre imágenes de intensidad. Más recientemente, se ha propuesto un método eficiente de inventario de señales de tráfico en [23]. Paralelamente, en [24] se aplicó un algoritmo automático para detectar los límites del borde del asfalto de la carretera para la evaluación de la seguridad del mantenimiento de la carretera.

Por otro lado, la extracción automática de la alineación / señalización vial y la consistencia del diseño geométrico son líneas de investigación activas para la seguridad vial. En este sentido, en [26] se compararon varias metodologías y estrategias para obtener alineaciones viales y parámetros geométricos de vías existentes. En [27] se presentó un algoritmo automatizado para extraer marcas viales del conjunto de datos MLS. Por su parte, en [28] se proporcionó un nuevo enfoque para la extracción de bordes de carreteras basado en un modelo de contorno activo paramétrico. En [13] se presentó un nuevo modelo de consistencia basado en perfiles

de velocidad de operación continua; y en [29] se demostró claramente el papel que juegan las inconsistencias de diseño en la seguridad vial. En todos estos campos de investigación, el MLS ya ha realizado aportes importantes. Sin embargo, debido a su gran potencial investigador, aún por explorar, la aplicación del MLS podría suponer un cambio de paradigma, especialmente en la evaluación geométrica y las auditorías de seguridad vial [30].

1.2. Sistemas LiDAR aerotransportado y ortofotografía aérea para la captura de edificios

En la búsqueda de avanzar en la consecución de los ODS, el uso de energía solar se ha incrementado en los últimos años, proporcionando el 6,6% de la producción total de energía primaria de energías renovables en Europa en 2017 [31]. Con la creciente preocupación por el problema del cambio climático, los consumidores son más proclives a convertirse en prosumidores (productores y consumidores de energía simultáneamente), opción que es también promovida por los gobiernos nacionales para fomentar el consumo de energía renovable [32].

En este contexto, y gracias a la reducción de los costes de las instalaciones solares fotovoltaicas (FV) debido a la madurez de esta tecnología [33], la energía solar fotovoltaica es una opción que despierta gran interés para los usuarios, especialmente por su explotación en sus principales lugares de consumo (hogares e industrias). El lugar ideal para la instalación de sistemas fotovoltaicos es en los propios edificios, de modo que se potencia la generación de energía distribuida y el autoconsumo evitando pérdidas de energía por transporte. En concreto, de manera habitual se eligen las cubiertas como ubicación principal para instalaciones fotovoltaicas por su posición favorable con respecto al Sol [7]. Estas cubiertas garantizan la mínima pérdida de energía [34], optimizan los costes de instalación y maximizan la recepción de la radiación solar. Dado que los tejados son los lugares más comunes para ubicar estas instalaciones, conocer sus características es fundamental para analizar la viabilidad, producción y rentabilidad de las instalaciones fotovoltaicas [35]. Existen herramientas para estimar la producción de energía fotovoltaica de los tejados, como Google SunRoof [36]. Si bien este tipo de herramientas disponibles permite el análisis de costes y rentabilidad de las instalaciones fotovoltaicas en cubiertas, tienen limitaciones derivadas de: (i) el uso de datos bidimensionales, (ii) valores ideales de producción solar y (iii) desconocer los rendimientos reales de los paneles.

Para obtener las características de los tejados necesarias para determinar el potencial solar de los mismos (superficie, orientación azimutal, e inclinación, que se incorporan a los datos geométricos como semántica en su modelado) se pueden utilizar datos abiertos del IGN [10] que proporcionan nubes de puntos LiDAR aéreo y ortoimágenes aéreas. Estas nubes de puntos ofrecen precisión para la documentación, caracterización y reconstrucción en 3D, especialmente en términos de altimetría. Sin embargo, es fundamental analizar sus limitaciones en términos de resolución espacial para discernir qué elementos podrían caracterizarse adecuadamente y cuáles no. Estos datos cartográficos tienen una resolución espacial muy baja, en concreto, entre 0,5-1 puntos por metro cuadrado dependiendo de la superposición de la franja de vuelo, en las nuevas coberturas las resoluciones llegan a alcanzar los 14 puntos por metro cuadrado [37]. Por tanto, aunque se trata de una fuente de datos 3D de alta precisión, no es aconsejable utilizarla para calcular las dimensiones de los tejados, ya que podría dar lugar a su subestimación por la pérdida de información sobre los bordes reales de los mismos.

Por otra parte, las ortoimágenes aéreas del Programa Nacional de Ortofotografía (PNOA) [10] son el resultado de una red fotogramétrica captada por una cámara digital de gran formato (Vexcel UltraCam). Para esta red fotogramétrica se consideró una superposición longitudinal entre imágenes del 60% y una superposición transversal ≥ 25%. Como resultado, se obtuvo una ortoimagen con un tamaño de píxel sobre el terreno (GSD - Ground Sample Distance) de 25 cm, lo que resulta una gran ventaja en términos de resolución espacial. Esta resolución permite obtener una estimación precisa de los bordes externos del tejado mediante procedimientos de visión artificial como algoritmos de detección de bordes y procedimientos de segmentación. A pesar de su alta resolución espacial, este producto bidimensional no es útil para definir, por sí solo, el ángulo de inclinación cada agua del tejado. En cuanto a su precisión planimétrica, este producto cartográfico ofrece una precisión de ± 50 cm.

A pesar de que las ortoimágenes proceden de la fotogrametría, técnica que surge para proporcionar información de la tercera dimensión a partir de imágenes, la limitación que presentan para determinar el ángulo de inclinación hace necesaria la combinación con datos LiDAR para el cálculo simultáneo de los parámetros del tejado para varios edificios [8, 38, 39]. Sin embargo, la mayoría de los enfoques existentes provienen de software comercial [38] o calculan el potencial solar de los tejados sin discretizar entre las diferentes aguas del tejado y sus diferentes orientaciones azimutales y ángulos de

inclinación [8]. Otros enfoques hacen uso de imágenes aéreas para la reconstrucción de edificios en 3D, de tal manera que parte de la precisión se puede perder al generar el Modelo Digital de Superficie [39]. Dada la complejidad y variedad de cubiertas existentes, algunos trabajos [40] presentan una metodología centrada en cubiertas inclinadas. Sin embargo, este enfoque está orientado a objetos y calcula el potencial de energía solar para edificios individuales. Si bien este es un enfoque válido en regiones con viviendas unifamiliares, no se puede aplicar en áreas urbanas donde la radiación solar de los edificios se ve afectada por la radiación de su entorno [41]. En este sentido, [42] calculó el potencial solar del tejado en base a datos LiDAR y para vecindarios completos, teniendo en cuenta cómo los edificios del entorno afectan en términos de sombras proyectadas. Sin embargo, el uso de datos LiDAR por sí solo presenta limitaciones en términos de oclusiones, resolución espacial y precisión horizontal que pueden resolverse mediante su combinación con imágenes aéreas [43].

1.3. Automatización en el procesamiento de datos geoespaciales para la transferencia de conocimiento.

El conocimiento que se genera en las investigaciones científicas tiene que revertir a la sociedad. Cuando estos avances se aplican al mundo empresarial, se puede obtener información de retorno, que ayuda a fortalecer y perfeccionar las investigaciones realizadas. Se detallan aquí algunos proyectos de transferencia en los que el doctorando ha desarrollado aplicaciones informáticas que procesan datos geoespaciales.

PROYECTO	INROAD: Sistema integral para la prevención y la asistencia al rescate en accidentes de tráfico.
Fecha	2017-2018
Financiado por:	Acción Estratégica: Ministerio de Energía, Turismo y Agenda Digital
Socios	MovilData: Sistemas de gestión de flotas y control GPS
Aporte en el procesamiento de datos geoespaciales	Desarrollo de aplicación *inRoad inAlert* para: procesar nubes de puntos LiDAR MLS, extraer de forma automática el eje de la carretera, obtener la alineación en planta con los tramos de curvas circulares, tramos rectos, tramos de curvas de transición, y asignación de índice de riesgo según un triple criterio de estabilidad.

PROYECTO	Ener3DMap: Gestión de Clientes y recursos energéticos distribuidos: mapeado energético
Fecha	2018- 2021
Tipo	Cátedra Iberdrola VIII Centenario USAL
Socios	IBERDROLA
Aporte en el procesamiento de datos geospaciales	Aplicación Ener3DMap-SolarRoofs que procesa nubes de puntos LiDAR e ortoimágenes aéreas para proporcionar orientación, inclinación y superficie de las aguas del tejado, y calcular el potencial de producción solar.

PROYECTO	EPESOL. Nuevo instrumento de evaluación de la penetrabilidad solar en el mercado energético español.
Fecha	06/2018- 09/2021
Tipo	Retos Colaboración del Ministerio de Economía, Industria y Competitividad, en consorcio con Iberdrola España S.A.
Socios	IBERDROLA España SAU
Aporte en el procesamiento de datos geoespaciales	Desarrollo de la infraestructura de datos espaciales (IDE) Ener3DMap-SolarWeb Roofs para el cálculo de la producción solar fotovoltaica enlazando con la API de PVGIS. Desarrollo de la IDE Ener3DMap-SolarWeb Cities para el estudio de prospectiva de producción Solar.

PROYECTO	Análisis de la viabilidad técnica y económica de la electrificación de la flota de autobuses urbanos de Ávila
Periodo	03/2020 – 09/2020
Empresa	IBERDROLA España SAU
Aporte en el procesamiento de datos geoespaciales	Desarrollo de una aplicación que permita calcular la orografía de las rutas de autobuses y analizar sus pendientes, para determinar la compatibilidad con las pendientes máximas soportadas por los modelos eléctricos en el mercado.

PROYECTO	INROAD 4.0: Intelligent Roads for a 0 Vision (0 deceased, 0 Injured, 0 congestion and 0 emissions)
Fecha	2018-2022
Tipo	Proyecto Cien. Centro para el Desarrollo Tecnológico Industrial, Ministerio de Ciencia e Innovación.
Socios	TECOPYSA
Aporte en el procesamiento de datos geoespaciales	Actualmente en desarrollo: aplicación InRoad 4.0 que procesa nubes de puntos LiDAR MLS, para extraer de forma automática el eje de la carretera, calcular sus parámetros en planta y en alzado, realizar un análisis de la visibilidad de parada y la visibilidad de cruce, y proponer un índice de riesgo asociado a la visibilidad calculada.

1.4. Estructura de la Tesis Doctoral

La Tesis Doctoral que aquí se presenta se realiza por compendio de artículos científicos publicados en revistas internacionales de alto impacto, según la normativa al respecto de la Universidad de Salamanca.

Se incorporan tres artículos científicos y dos registros de propiedad intelectual. La estructura se organiza en seis capítulos para exponer la investigación realizada. Se añaden además dos apéndices finales para incluir información del impacto de las publicaciones y detalles de los registros de la propiedad intelectual de los programas desarrollados.

Dicha estructura se detalla a continuación:

Capítulo I. Introducción: En este apartado se proporciona la descripción del contexto bajo el que se desarrolla esta Tesis Doctoral. Se incluyen además secciones con las tecnologías utilizadas para su realización, una apartado con proyectos de transferencia de conocimiento, en los que se ha participado, y se termina con el presente apartado dónde se refleja la estructura del documento.

Capítulo II. Hipótesis de trabajo y objetivos: En este capítulo se muestran las hipótesis de trabajo contempladas al comienzo de esta investigación, y los objetivos que se formularon al inicio.

Capítulo III. Automatización de procesado de nubes de puntos en infraestructuras viarias: Se incorpora aquí el artículo *"Road safety evaluation through automatic extraction of road horizontal alignments from Mobile LiDAR System and inductive reasoning based on decision tree"*, que ilustra la automatización desarrollada para la extracción del eje de una carretera, la alineación en planta y la seguridad de los tramos, a partir de nubes de puntos obtenidas con un sistema de cartografiado móvil.

Capítulo IV. Automatización de procesado de nubes de puntos y ortoimágenes en edificios: Muestra el artículo *"Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity"* que ilustra el desarrollo realizado para procesar de forma combinada nubes de puntos de LiDAR aéreo y ortoimágenes, y obtener la geometría de los tejados para poder calcular el potencial solar fotovoltaico de los mismos.

Capítulo V. Automatización en el cálculo de instalaciones fotovoltaicas: Se recoge la tercera publicación *"A geospatial web-based platform for the photovoltaic potential computation: Ener3DMap-SolarWeb Roofs"* que automatiza el cálculo de la producción solar de una instalación solar fotovoltaica, a partir de la geometría de tejados generada según la investigación detallada en el artículo anterior, y con la ubicación seleccionada en el mapa y el modelo solar PVGIS.

Capítulo VI. Conclusiones y líneas futuras: Como capítulo final se analizan los resultados obtenidos, se reflejan las conclusiones, y se detallan qué frentes de investigación quedan abiertos para avanzar con la investigación.

Apéndice A. Indexación y factor de impacto de las revistas: Se recopila la información sobre la relevancia de las publicaciones que componen el compendio de artículos en los que se basa esta Tesis Doctoral.

Apéndice B. Software desarrollado: se muestran los detalles de las aplicaciones desarrolladas, y también de los registros de la propiedad intelectual.

CAPITULO II

Hipótesis de trabajo y objetivos

2. Hipótesis de trabajo y objetivos

2.1. Hipótesis de trabajo

La línea de investigación, que se sigue en esta Tesis Doctoral, se centra en la búsqueda de metodologías y el desarrollo de algoritmos y aplicaciones informáticas, para automatizar el procesado de grandes volúmenes de datos: nubes de puntos 3D y ortoimágenes georreferenciadas. Estos datos, se adquieren con los dispositivos móviles de captura de información geoespacial de escenarios reales complejos, analizados en el capítulo anterior.

Esta automatización en el procesado y la dotación de propiedades semánticas para esa información se realiza en infraestructuras y edificios, en el ámbito de la seguridad de carreteras, y en el ámbito de autoconsumo con energía solar fotovoltaica.

Con el aumento de precisión de los sistemas de digitalización 3D: Sistema de cartografiado móvil (MLS) y Sistemas LiDAR y ortoimágenes de vuelos fotogramétricos, el volumen de información generada ha crecido mucho, y se hace necesaria la investigación en el desarrollo de herramientas de tratamiento de este Big Data geoespacial, para agilizar el procesado de dicha información, la automatización del mismo, y la integración de datos procedentes de distintas fuentes.

Los trabajos realizados previamente en automatización del procesamiento de datos geoespaciales de infraestructuras, no han generado herramientas que permitan dotar de las propiedades semánticas buscadas para la categorización del riesgo de los tramos de carreteras, y para obtener la geometría de tejados.

Para conseguir avanzar en estas cuestiones, se plantean las siguientes hipótesis:

- Los sistemas de cartografiado móvil (MLS) permiten obtener la digitalización en 3D de infraestructuras viarias, y capturan la geometría, intensidades y colores para cada punto.
- El Instituto Geográfico Nacional de España, ofrece nubes de puntos en 3D, y ortoimágenes georreferenciadas procedentes de vuelos fotogramétricos, información que está disponible para su descarga libre.
- Es posible plantear metodologías para identificar y extraer geometrías automáticamente de las nubes de puntos 3D.
- Es posible combinar el procesado de nubes de puntos con imágenes de vuelos fotogramétricos, para obtener las geometrías de los tejados.
- Los equipos informáticos actuales permiten realizar procesamiento multihilo, para reducir el tiempo al manejar enormes nubes de puntos.

- Existen entornos de desarrollo de aplicaciones de software libre que permiten la programación con librerías que facilitan el procesado de ortoimágenes y nubes de puntos LiDAR.
- Hay disponibles librerías de código abierto que facilitan el desarrollo de aplicaciones de mapas web interactivos.
- Existen modelos con datos de radiación solar para calcular la producción fotovoltaica en cualquier punto a partir de los parámetros de las instalaciones.

2.2. Objetivos

A partir de las hipótesis planteadas en el apartado anterior, se plantea el siguiente objetivo general:

- Automatizar el procesado de nubes de puntos y de ortoimágenes digitales procedentes de la digitalización 3D de infraestructuras y edificios para avanzar en los objetivos de desarrollo sostenible de la Agenda 2030, y en los objetivos del Plan Visión 0 de la Unión Europea.

Para ello se plantean los siguientes objetivos específicos:

- Desarrollar metodologías y algoritmos para automatizar el procesamiento de nubes de puntos 3D, procedentes de sistemas de cartografiado móvil (MLS) de infraestructuras viarias, para extraer automáticamente el eje, tramos de curvas circulares, tramos rectos y clotoides.
- Calcular de forma automática los parámetros de estabilidad de cada tramo y asignarles un índice de peligrosidad según dichos parámetros, de forma que este índice se deba exclusivamente a causas inherentes al trazado.
- Desarrollar un software para la extracción de la orientación e inclinación de los tejados a partir de nubes de puntos de LiDAR aéreo, para aislar las aguas de los tejados y para clasificarlas según su orientación.
- Identificar los bordes exteriores de cada tejado, y por tanto su superficie, a partir de ortoimágenes procedentes de vuelos fotogramétricos.
- Crear una capa vectorial con las aguas de los tejados, añadiendo en cada agua la superficie, orientación e inclinación correspondientes, para su uso en una aplicación de mapas web, que proporcione datos de producción solar fotovoltaica.
- Enlazar la herramienta con la API de PVGIS para calcular la producción solar a partir de los parámetros proporcionados con la selección del tejado.

Capítulo III

Automatización de procesado de nubes de puntos en infraestructuras viarias.

3. Automatización de procesado de nubes de puntos en infraestructuras viarias

Se incorpora aquí el artículo *"Road safety evaluation through automatic extraction of road horizontal alignments from Mobile LiDAR System and inductive reasoning based on decision tree"*, que ha sido publicado en la revista de alto impacto: ***ISPRS Journal of Photogrammetry and Remote Sensing***, en Octubre de 2018.

3.1. Resumen

La seguridad vial es una necesidad para cualquier sociedad, debido a los altos costes en daños personales que producen los accidentes de tráfico. El plan Vision 0 de la Unión Europea, busca 0 muertes, 0 heridos, 0 atascos y 0 emisiones. La mayoría de los indicadores del riesgo potencial de un tramo de carretera se basan en estadísticas de accidentes previos, pero no analizan si hay alguna característica del trazado que sea responsable de su alta peligrosidad. En esta primera publicación científica se recoge el desarrollo de una metodología novedosa que permite evaluar la seguridad de una carretera a partir de los parámetros de la alineación en planta de sus tramos.

Se utilizó un sistema de cartografiado móvil (o MLS) para obtener las nubes de puntos 3D correspondientes a carreteras nacionales y secundarias puesto que la geometría de su trazado es más susceptible de relacionarse con una alta peligrosidad. Estos millones de puntos obtenidos necesitan ser procesados para poder extraer información útil de la vía, la alineación en planta de la misma, con sus tramos de curvas circulares, sus tramos rectos, y los tramos de curvas de transición.

La automatización se obtiene mediante un proceso de razonamiento inductivo, basado en un árbol de decisiones, que proporciona una evaluación del riesgo potencial. Para lograrlo, la nube de puntos 3D obtenida con el sistema de cartografiado móvil se clasifica mediante un algoritmo iterativo e incremental, basado en una triangulación de Delaunay 2.5D y 3D que permite identificar el eje de la carretera. Una vez obtenido el eje, mediante un algoritmo de procesamiento automático, basado en RANSAC (Random sample consensus), se extraen las curvas circulares, después, de los puntos del eje restantes, se extraen las rectas, y con los puntos pendientes de clasificar se forman los tramos de curvas de transición. Sobre estos tramos de alineación horizontal de la carretera, se obtienen los índices de consistencia geométrica, basados en un criterio de triple estabilidad conjunta. Además, este trabajo tiene como objetivo proporcionar una herramienta preventiva/predictiva sencilla y eficaz para las inspecciones de seguridad vial.

La metodología propuesta se implementó en cuatro tramos de carreteras españolas, cada uno con diferentes condiciones de tráfico que representan los tipos de vías más comunes. Esta metodología desarrollada fue validada con éxito mediante la comparación de los resultados con respecto a los proyectos de construcción de las mismas.

Estos algoritmos fueron implementados en una aplicación informática ***inRoad inAlert***, desarrollada por los autores en entorno QT creator que permite realizar el procesado con la mínima intervención del usuario (apéndice B).

Palabras clave: seguridad vial; árbol de decisión; consistencia geométrica del trazado; parámetros de alineación horizontal; sistema de cartografiado móvil.

3.2. Publicación 1: Road safety evaluation through automatic extraction of road horizontal alignments from Mobile LiDAR System and inductive reasoning based on decision tree

ISPRS Journal of Photogrammetry and Remote Sensing 146 (2018) 334–346

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Road safety evaluation through automatic extraction of road horizontal alignments from Mobile LiDAR System and inductive reasoning based on a decision tree

José Antonio Martín-Jiménez[a], Santiago Zazo[b], José Juan Arranz Justel[c], Pablo Rodríguez-González[b,d], Diego González-Aguilera[a,*]

[a] University of Salamanca, Department of Cartographic and Land Engineering, Hornos Caleros, 50, 05003 Ávila, Spain
[b] University of Salamanca, TIDOP Research Group, Hornos Caleros, 50, 05003 Ávila, Spain
[c] Technical University of Madrid, Department of Topographic Engineering and Cartography, Camino de la Arboleda, s/n Campus Sur, Autovía de Valencia Km 7, 28031 Madrid, Spain
[d] Universidad de León, Department of Mining Technology, Topography and Structures, Astorga, s/n, 24401, Ponferrada, León, Spain

ARTICLE INFO

Keywords:
Road safety
Decision tree
Geometric design consistency
Horizontal alignment parameters
Mobile LiDAR System

ABSTRACT

Safe roads are a necessity for any society because of the high social costs of traffic accidents. This challenge is addressed by a novel methodology that allows us to evaluate road safety from Mobile LiDAR System data, taking advantage of the road alignment due to its influence on the accident rate. Automation is obtained through an inductive reasoning process based on a decision tree that provides a potential risk assessment. To achieve this, a 3D point cloud is classified by an iterative and incremental algorithm based on a 2.5D and 3D Delaunay triangulation, which apply different algorithms sequentially. Next, an automatic extraction process of road horizontal alignment parameters is developed to obtain geometric consistency indexes, based on a joint triple stability criterion. Likewise, this work aims to provide a powerful and effective preventive and/or predictive tool for road safety inspections. The proposed methodology was implemented on three stretches of Spanish roads, each with different traffic conditions that represent the most common road types. The developed methodology was successfully validated through as-built road projects, which were considered as "ground truth."

1. Introduction

According to Camacho-Torregrosa et al. (2013), annually, 1.20 million people die and another 20–50 million people are injured in traffic accidents; furthermore, road crashes involve important and high social costs (da Costa et al., 2016). Although the influence of the human factor is well known (Siskind et al., 2011), roadway design also contributes to the occurrence of accidents (Garach et al., 2016). It should be noted that approximately 30% of road accidents are attributable to infrastructures, and this collision trend is focused on specific road segments (Camacho-Torregrosa et al., 2013; López et al., 2016). For that, the binomial human-road factor plays a crucial role (López et al., 2016). Thus, obtaining safe roads and reducing accidents are challenges that any society should address.

In this framework, the European Union adopted the Directive 2008/96/EC, which is based on the principle of prevention (EU, 2008). This directive establishes different procedures with the final goal of detecting roadway deficiencies and reducing Trans-European Transport Network (TETN) crashes along all its phases: from planning and design to operation of the road infrastructure (Sitran et al., 2016). This directive introduces relevant aspects, such as (a) Road Safety Impact Assessments (RSIA), (b) Road Safety Audits (RSA), (c) Safety Ranking and Management (SRM) and (d) Road Safety Inspections (RSI).

RSIA introduces road safety considerations in the initial planning stage. Through RSA, the road characteristics are checked in the design stage. These two procedures are carried out during the stage of planning. Already during the in-service road stage, SRM provides the ranking of high accident concentration sections and establishes the road infrastructure safety management; by means of RSI, road hazards and safety issues are detected. In this sense, RSI can be understood as an effective preventive tool for the road network. EU (2008) is mandatory on TETN; however, it can also be applied to any national road transport infrastructure as a "good practice guide."

Another approach to road safety, applicable to infrastructures

* Corresponding author.
E-mail address: dagulera@usal.es (D. González-Aguilera).

https://doi.org/10.1016/j.isprsjprs.2018.10.004
Received 19 May 2018; Received in revised form 1 October 2018; Accepted 10 October 2018
Available online 21 October 2018

already in service, is the one presented in the European Road Assessment Program (EuroRAP). In this program, a Risk Index (IR) of the road is obtained, based on the accident statistics and traffic intensity (Average Daily Traffic-ADT), which is complemented through an inspection protocol with images and a safety score using stars (EuroRAP, 2018). It should be noted that, in this EuroRAP approach, the geometrical aspects of the road are considered in a very generic manner.

From a dual research-engineering perspective, the concept of geometric consistency in road design has a direct influence on road safety (Ng and Sayed, 2004). The studies developed have been based fundamentally on aspects such as purely geometric, models of the speed of operation, vehicle stability and the workload of the driver (Andrasik and Bil, 2016, Eftekharzadeh and Khodabakhshi, 2014, Pérez-Zuriaga et al., 2013). However, we must highlight the works developed by Lamm et al. (1991, 1995, 1999, 2001) due to their proposal of the simultaneous triple criterion of stability: (i) in the design (Criterion I), (ii) in the speed of operation (Criterion II) and (iii) in the driving dynamics (Criterion III), which today continue to be a benchmark in the field of road safety.

With respect to data acquisition, over the last few decades, geomatic advances in the Global Navigation Satellite System (GNSS), Light Detection and Ranging (LiDAR), the acquisition of radiometric information, and the hybridization between passive-active sensors, techniques and platforms have revolutionized the massive acquisition of survey data, as well as the inspection and monitoring techniques. A clear example of this is Mobile LiDAR System (MLS) (Bitenc et al., 2011; Gonzalez-Jorge et al., 2013; Mc Elhinney et al., 2010; Puente et al., 2013a).

MLS is principally composed of a navigation system and one or more LiDAR sensors. In this manner, positional data and intensity information of the environment are acquired. Additionally, MLS can be completed by photographic systems (RGB digital cameras) and/or other sensors, such as thermal cameras, ground penetrating radar (GPR) and profilometers. A navigation system is integrated via a set of sensors, such as GNSS, Inertial Measurement Unit (IMU), and Distance Measuring Indicator (DMI) (Holgado-Barco et al., 2015). Moreover, MLS is an accurate and efficient system for data acquisition in complex environments, such as urban and road corridors over large areas, which provides important time reduction for the collection and processing of data (Castro et al., 2016; Holgado-Barco et al., 2015; Varela-Gonzalez et al., 2014).

In the civil engineering field, MLS is beginning to be consolidated among researchers and engineers. Recently, considerable efforts and progress have been mainly made in tasks such as point cloud segmentation, road alignment extraction and automatic object detection. Varela-Gonzalez et al. (2014) proposed a novel method to automatically remove vehicles from mobile LiDAR datasets. Holgado-Barco et al. (2015) developed a method based on segmentation, parameterization and filtering LiDAR point clouds from MLS to extract, semi-automatically, road centrelines and determine horizontal road parameters and their alignment (i.e., straight lines, circular arcs and clothoids). For their part, Riveiro et al. (2015) described an algorithm for the automatic detection of zebra crossings by means of the standard Hough Transform, which is applied over intensity images. More recently, a cost-effective traffic sign inventory method was proposed in Ai and Tsai (2015). In parallel, Cabo et al. (2016) applied an automatic algorithm to detect road asphalt edge limits for road maintenance and safety assessment. An adequate review of the scientific literature can be found in Yang et al. (2013).

Alternatively, the automatic extraction of road alignment/markings and geometric design consistency are active research lines for traffic safety. In this sense, Marinelli et al. (2017) compared several methodologies and strategies to obtain the road alignment and geometric parameters of existing roads. Kumar et al. (2014) presented an automated algorithm for extracting road markings from MLS data sets. For their part, Kumar et al. (2013) provided a new approach to road edge extraction based on a parametric active contour model. In Camacho-Torregrosa et al. (2013), a new consistency model was presented based on continuous operating speed profiles, and Montella and Imbriani (2015) clearly demonstrated the role that design inconsistencies play in road safety. In all these research fields, MLS has already produced significant contributions; however, due to its large research potential, still to be explored, the application of MLS could suppose a paradigm shift, especially over geometric assessment and road safety audits (Gargoum and El-Basyouny, 2017).

On the other hand, it is also worth noting other approaches, applied to in the road field to extract road geometries, which do not employ fundamentally terrestrial LiDAR sensors to data acquisition. Hatger and Brenner (2003) showed, through the combination of existing databases with aerial laser scanner (ALS) data, that it was possible to derive geometrical properties of roads such as height, longitudinal and transversal slope, curvature, and width. In Clode et al. (2004) is presented an approach to extract roads from ALS point clouds with a point density of 1 point m^{-2}, based on a progressive hierarchical classification technique using a digital terrain model created from the last pulse and intensity information of LiDAR. After that, in Clode et al. (2005) this approach to extract roads was improved by a building detection technique that also allowed the detection of existing bridges within the road network, improving the extraction of longitudinal and transverse road profiles. For its part, Alexander et al. (2010) suggested to apply backscatter coefficient versus discrete return data to classify roads from ALS data. In urban areas, Zhou and Vosselman (2012) addressed the problem of detection of road edges by detecting curbstones in three steps. They found very similar values between ALS and MLS techniques. For instance, in Marinelli et al. (2017) it is shown a novel mobile mapping (MM) vehicle that integrated set of low-cost sensors (GNSS receivers, IMU system and high definition webcams), and where is found that reliability of terrestrial MM is highly dependent on the accuracy of GNSS sensors. For their part, Javanmardi et al. (2017) propose an automatic methodology to extract road features from high resolution airborne images using adaptive thresholding, and whose accuracy is decametric. In Azimi et al. (2018) a novel pixel-wise method is developed that semantically segments high resolution aerial images in order to detect lane markings. This was done through a combination of fully convolutional neural networks with discrete wavelet transform. The images were acquired by three low cost cameras Canon Eos 1Ds Mark III model, ground sampling distance was 13 cm approximately to a flight height of about 1000 m above ground level. This method reported high pixel classification accuracy, around 99%. Regarding high-resolution satellite images, recent studies have focused on road-centerline extraction. Sujatha and Selvathi (2015) present an algorithm to segment and connect road region and remove non-road pixels using morphological operation, with a high average value of completeness-correctness-quality (90%, 96%, and 87%, respectively). Alshehhi and Marpu (2017) presented a new approach based on hierarchical graph-based image segmentation to extract roads, indicated in urban areas especially, which displayed over 90% effectiveness in road network detection.

Last but not least, there are issues related to road safety and risks. Traditionally, safety studies have focused on factors such as the probability of crashes, types of drivers and roads, etc.; decision tree (DT) techniques have been successfully proved, individually or in combination with decision rules, either as predictive models or as a tool for searching patterns that can explain accident causes. This is also due to their simplicity, the hierarchical structure and the ease of interpretation of results (de Oña et al., 2013; López and de Oña, 2017). For that reason, DTs can be characterized as an effective and adequate tool for the decision-making process. Some examples are the studies developed by Chang and Wang (2006), Chang and Chien (2013), Jung et al. (2016), Kwon et al. (2015), López et al. (2016). However, thus far, there are no studies that address how the road horizontal geometric

alignment contributes to road safety by means of a categorization of its inherent geometric risk.

To this end, this work aims to provide a novel and efficient method to assess road safety by means of MLS and the estimation of a potential risk assessment (PRA). This PRA is exclusively derived from a coarse-to-fine approach using point clouds as input data: from the automatic segmentation of roads and the extraction of its horizontal alignment parameters; to the estimation of PRA and the road safety classification based on a decision tree that is an inductive reasoning applied for the first time on geometric parameters exclusively.

The proposed methodology has been successfully implemented and validated on three road stretches that represent the most common types of existing roads in Spain and that present different traffic conditions.

The remainder of this paper is organized as follows: after this introduction, a description of the study cases, the MLS technique and the proposed methodology to assess road risk are shown in Section 2. Section 3 presents the main experimental results drawn from the research. Lastly, in Section 4, the proposed algorithms and the decision support tool are discussed, and the general conclusions from the study are shown.

2. Materials and methods

2.1. Case studies

90% of Spanish roads are secondary roads (MFOM, 2018) which connect population centres of minor importance in urban and rural areas. These roads present a greater accident rate, despite their density of traffic being less than that of highways and multilane roads (DGT, 2017). For this reason, in this research, three of the four real case studies are secondary roads, whereas the other case is a main road. According with the current Spanish geometric design standards (MFOM, 2016), there are two road groups: (i) highways/multilane roads and (ii) other roads, classified as conventional roads. Within conventional roads, the difference between main and secondary roads refers to the importance of the population centres that the road crosses or connects.

The first, second and fourth case studies are located on the road LU-722, which is situated in the northwest of Spain (Lugo province, in the region of Galicia). These non-consecutive case studies comprise a horizontal sinuous stretch, between kilometre points 4.0 and 121. Their lengths are 1,776.6, 3,160.6 m and 1,427.6 m, respectively. The average cross-section consists of a 6.50-m-wide roadway and a 0.75-m hard shoulder on either side. The Average Annual Daily Traffic (AADT) is low at 918 vehicles per day (veh/d), (Xunta de Galicia, 2016).

N-640 is the third case. This is also located in Lugo province, and it is characterized by a succession of linked curves. The stretch covers a distance of 2.6 km between kilometre points 84 and 86. The roadway has a single road with two 3.50-m lanes with outer hard shoulders of 1.00 m, roughly. AADT varies between 8,261 and 8,764 veh/d (MFOM, 2017).

By means of the first, second and third case studies will be carried out both the extraction of the horizontal geometric parameters and the training process of the inductive reasoning through a decision tree. The last case study will be applied to risk validation process exclusively.

2.2. Mobile LiDAR System (MLS) and data sets

Data acquisition was carried out by means of Lynx Mobile Mapper by Optech. This system acquires a LiDAR point cloud and RGB imagery simultaneously. The system is composed of two LiDAR sensors, four RGB cameras, a GNSS system, an IMU and a DMI. In this research it did not use camera data but only point clouds. A complete description of the platform and sensors applied is provided in Holgado-Barco et al. (2015), Puente et al. (2013b). Table 1 shows the main technical characteristics.

Table 1
Lynx Mobile Mapper Optech technical characteristics. Accuracy (1).

MLS sensors	Parameter	Value
GNSS	X,Y coordinates	0.020 m (1)
	Z coordinate	0.050 m (1)
IMU	Roll-Pitch	0.005° (1)
	Yaw	0.015° (1)
LiDAR	Measuring principle	Time of Flight (ToF)
	Maximum range	200 m
	Precision range	8 mm (1 σ)
	Ranging accuracy	± 10 mm (1 σ)
	Laser measurement rate	75–500 kHz
	Measurement per laser pulse	Up to 4 simultaneous
	Scan frequency	80–200 Hz
	Laser wavelength	1,550 nm (near infrared)
	Angular resolution	0.001° (1)

Table 2
Point clouds data sets.

Case study	Length (m)	Point cloud/point density
LU-722. Stretch 1	≈ 1,800	27,017,955 points/ ≈ 121 points m^{-2}
LU-722. Stretch 2	≈ 3,200	43,705,509 points/ ≈ 110 points m^{-2}
N-640	≈ 2,200	31,017,623 points/ ≈ 35 points m^{-2}
LU-722. Stretch 3	≈ 1,430	18,373,715 points/ ≈ 185 points m^{-2}

To avoid hiding areas, LiDAR sensors were set up at a 45° angle to the platform's trajectory. Regarding three data sets, Table 2 presents a summary of acquired point clouds of this research.

2.3. Methodology

The methodology developed comprises four main steps once the data have been acquired by MLS (Fig. 1). First, an alternative approach was implemented for Mobile LiDAR point cloud classification based on a hierarchical geometric and radiometric analysis of the original 3D MLS point cloud. Second, the horizontal alignment and its main road parameters were automatically extracted together with the computation of geometric design consistency indexes. Third, the Potential Risk Assessment of the road was estimated by a new predictive tool based on a tree induction algorithm. Fourth, the results obtained were compared and verified with those obtained through a road safety and surveyor expert, which were considered as "ground truth."

2.3.1. Step 1. Point cloud classification

The acquisition of information from MLS of a roadway is characterized as being highly dense; usually, methods are required to turn the original point cloud into a surface or volume (Arranz Justel, 2013). Typically, this process has been approached fundamentally by means of: (i) taking advantage geometric criteria (based on thresholds for the scan angles of the laser sweep, or extracted points by height difference between trajectory data and road surface, among others) or (ii) from the radiometric characteristics of the points (fixed or adaptative thresholds for the intensity values) (Diaz-Vilarino et al., 2016; Yan et al., 2016; Holgado-Barco et al., 2014, 2015; Kumar et al., 2014; Riveiro et al., 2015). However, this process is addressed here by an alternative, incremental and sequential algorithm (Arranz Justel, 2013; see Fig. 1 Step-1), following a threefold approach: (i) a first phase of detection of points belonging to the bare ground, (ii) a second phase where the road surface together with the road marks are determined from bare ground points and (iii) a third phase where all remaining elements around road surface and its environment (e.g. vertical signals, protection elements, vegetation, etc.) are classified.

The first phase is essential because establishes the reference from which the road surface and road marks (phase 2), as well as its

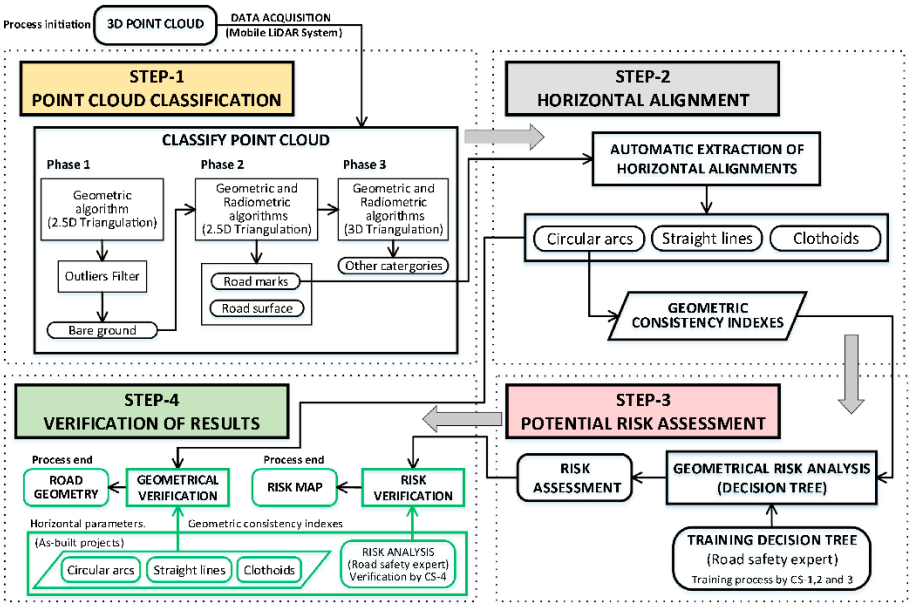


Fig. 1. General methodology developed for the automatic evaluation of road safety based on the road alignment. Note: CS: Case study.

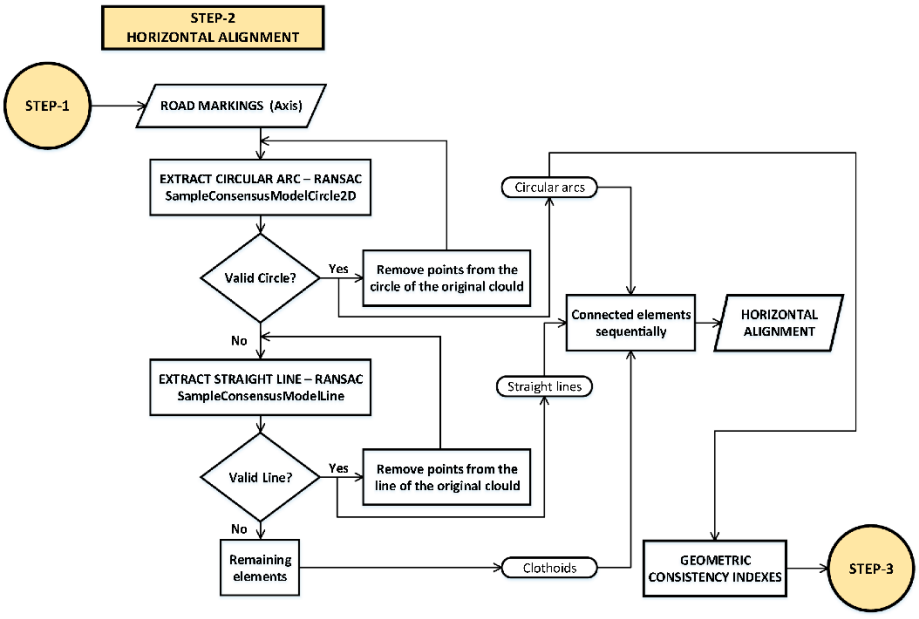


Fig. 2. Algorithm applied to extract the horizontal alignment parameters.

remaining elements (phase 3) are detected. In this first phase, a 2.5D Delaunay triangulation is used (Isenburg et al., 2006). Considering the huge number of points, the process is based on a sequential algorithm, the so-called "divide and conquer" strategy (Isenburg et al., 2006), where the cloud is divided by zones following a quadtree scheme, allowing an efficient use of the computer memory.

The second phase determines which points, of those previously classified as bare ground, are considered as road surface and road marks. To determine the points belonging to the road surface, a geometrical approach based on slopes and height differences allows us to find the edges of the asphalt and thus to determine the road surface. Regarding road marks, a radiometric algorithm based on intensity

Table 3
Road safety based on horizontal geometry consistency. Applied criteria. $V_{85(i)}$ for an individual element. $V_{85(i)}$ - $V_{85(i+1)}$ between consecutive elements.

Criterion I. Stability in designs		
$C\bar{C}R_S(gon/km)$	$\Delta V(km/h)$	Feature
$\lvert CCR_i - C\bar{C}R_S\rvert \leq 180$	$\lvert V_{85(i)} - V_D\rvert \leq 10$	Correct
$180 < \lvert CCR_i - C\bar{C}R_S\rvert \leq 360$	$10 < \lvert V_{85(i)} - V_D\rvert \leq 20$	Acceptable
$360 < \lvert CCR_i - C\bar{C}R_S\rvert$	$\lvert V_{85(i)} - V_D\rvert > 20$	Incorrect
Criterion II. Stability in operating speed		
$C\bar{C}R_S(gon/km)$	$\Delta V(km/h)$	Feature
$\lvert CCR_i - C\bar{C}R_S\rvert \leq 180$	$\lvert V_{85(i)} - V_{85(i+1)}\rvert \leq 10$	Correct
$180 < \lvert CCR_i - C\bar{C}R_S\rvert \leq 360$	$10 < \lvert V_{85(i)} - V_{85(i+1)}\rvert \leq 20$	Acceptable
$360 < \lvert CCR_i - C\bar{C}R_S\rvert$	$\lvert V_{85(i)} - V_{85(i+1)}\rvert > 20$	Incorrect
Criterion III. Stability in driving dynamics		
$C\bar{C}R_S(gon/km)$	$\Delta f_R = f_{RA} - f_{RD}$	Feature
$\lvert CCR_i - C\bar{C}R_S\rvert \leq 180$	$\Delta f_R \geq +0.01$	Correct
$180 < \lvert CCR_i - C\bar{C}R_S\rvert \leq 360$	$+0.01 > \Delta f_R \geq -0.04$	Acceptable
$360 < \lvert CCR_i - C\bar{C}R_S\rvert$	$\Delta f_R < -0.04$	Incorrect

values is applied to classify these features of the road surface.

Finally, the third phase classifies the remaining points (e.g. protection elements, vertical signals, vegetation, etc.). In order to obtain optimal results and considering the complexity of the road environment, a 3D Delaunay triangulation (Cavendish et al., 1985) is used to classify these elements, since points having the same planimetric location could have different height. In particular, this 3D approach considers the 3D coordinates of a point and thus takes advantage of the geometric relationships of the objects in the space.

It should be noted, that this third phase is not required for extracting road marks (see step 1 in Fig. 1); however we perform a whole classification of the road and its environment for other road safety studies related with visibility and protection elements that goes beyond the scope of this paper.

2.3.2. Step 2. Horizontal alignment and geometric consistency indexes

Once the road marks were classified, an iterative procedure using central lines exclusively, based on the Random Sample Consensus algorithm (RANSAC) (Fischler and Bolles, 1981) is utilized to automatically extract the horizontal geometric road elements/parameters (i.e., curves, straight lines and clothoids). It should be noted that in the case of conventional roads the geometric road axis is defined by the central horizontal road mark, which delimits each driving direction (MFOM, 2016).

Alternatively, a joint adaptive thresholding of the RANSAC algorithm is performed based on the geometric features of the road marks and the horizontal alignment parameters according to the type of road and the current regulation norm in Spain (MFOM, 1987, 2016). This

Table 4
Summary of road axis segmentation.

Case study	3D initial point cloud		2D road mark clasification		Road axis		
	Points	Density (point m^{-2})	Points	(%)	Points	Spacing (point m^{-1})	(%)
LU-722. Stretch 1	27,017,955	≈121	320,858	1.2	17,782	≈10	5.5
LU-722. Stretch 2	43,705,509	≈110	447,295	1.0	31,624	≈10	7.1
N-640	31,017,623	≈35	216,546	0.7	18,894	≈9	8.7
LU-722. Stretch 3	18,373,715	≈185	297,913	1.6	14,293	≈10	4.8

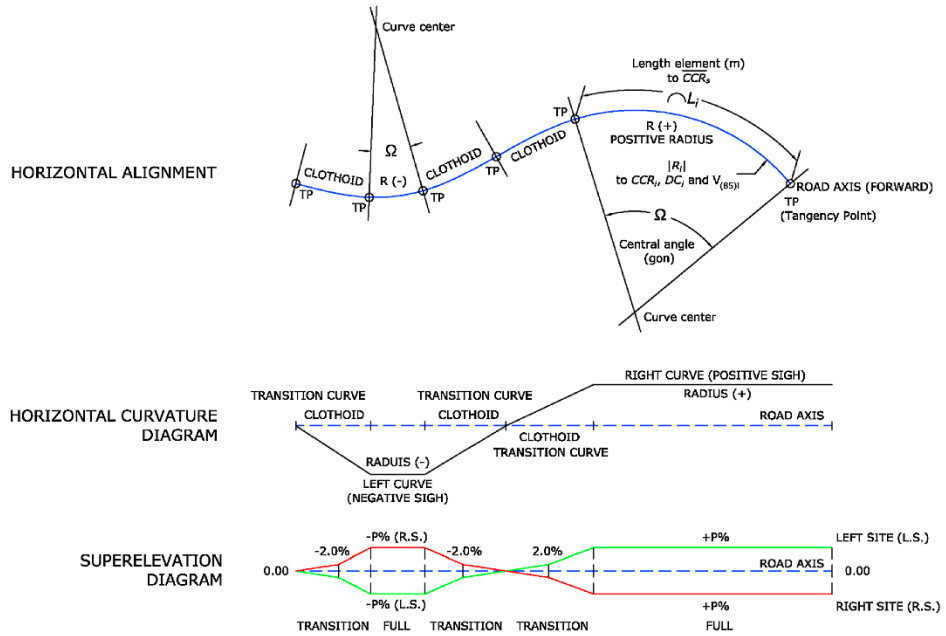


Fig. 3. Parameterization scheme considered for the formulation.

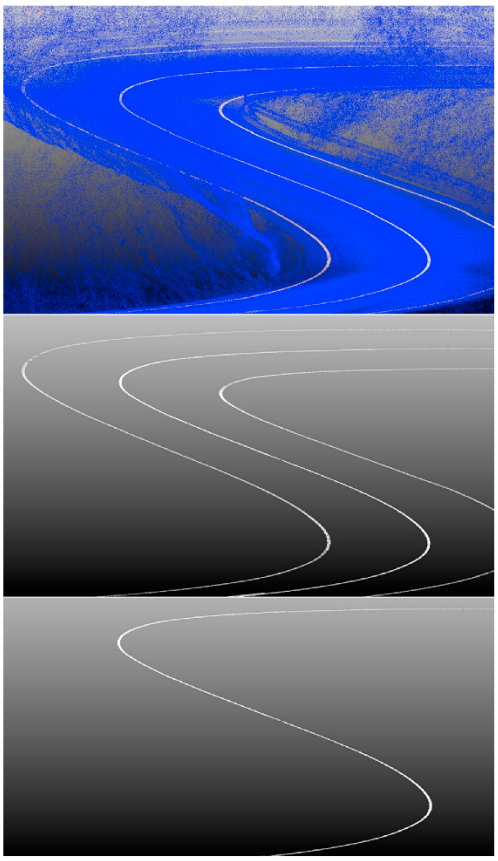


Fig. 4. Scheme resulting from the process of classification using the algorithm developed.

parameterization process is performed in three sub-steps (Fig. 2), which are supported by an open source point cloud library (PCL) based on the C++ language. First, each circular arc is obtained. Second, every straight line is found, and lastly, the remaining elements are classified as clothoids. Finally, all parametric elements are connected sequentially.

Regarding geometric parameters, the lack of consistency in road design has a negative direct impact on the increase in accident rate (Lamm, et al. 2001). In this sense, geometric consistency comprises three joint stability indexes that allow road safety to be assessed. These are (i) design consistency (Criterion I), (ii) operating speed consistency (Criterion II) and (iii) consistency in driving dynamics (Criterion III) (Lamm et al., 1999, 2001).

Furthermore, considering that the risk of accidents increases when the radius (R) of the curve decreases (Rasdorf et al., 2012; You et al., 2012) and that the Curvature Change Rate (CCR) is a key parameter due to its influence in the operation speed (Lamm et al., 2001), this study is focused on just circular alignments (Andrasik and Bil, 2016; Misaghi and Hassan, 2005). In this sense, it is noteworthy that the geometric consistency indexes are obtained exclusively from R and CCR parameters, highlighting the relevance that they have in road safety (Montella and Imbriani, 2015).

First, according to the Spanish geometric design standard (MFOM, 2016), the geometric consistency indexes are determined individually for each curve (Eq. (1)), as well as globally for the entire stretch analysed, excluding those straight lines (Eq. (2)) (Criterion I):

$$CCR_i = \frac{400^g}{2\cdot\pi\cdot R_i} = \frac{63.6620}{R_i}[gon/m] \approx \frac{63,700}{R_i}[gon/km] CCR_i = \frac{63,700}{R_i} \quad (1)$$

$$C\bar{C}R_S = \frac{\sum_{i=1}^{n}(CCR_i\cdot L_i)}{\sum_{i=1}^{n} L_i} \quad (2)$$

where CCR_i is the curvature change rate corresponding to the i curve and expressed in gons/km, R_i is the radius of the i curve in metres, and L_i is the length of the element analysed in metres. CCR_S is the global curvature change rate (in gons/km) as a function of the weighted average of the elements considered. It should be noted that $CCR = 0$ in straight lines because $R = \infty$ (Lamm et al., 1999).

Next, the operating speed of each curve is estimated (Eqs. (3) and (4)) via the eighty-fifth percentile of the speed (V_{85}) (Criterion II). This parameter represents the speed at which 85% of the drivers operate on a road in service, which is internationally accepted as a suitable measure of the operating speed (Fitzpatrick et al., 2000):

$$V_{85} = e^{(4.561-0.0058\cdot DC_i)} \quad (3)$$

$$DC_i = \frac{360\hat{A}^\circ}{2\cdot\pi\cdot R} = \frac{57.295}{R} \approx \frac{5730}{R}[\hat{A}^\circ/100m] \quad (4)$$

yielding V_{85} in km/h (Morrall and Talarico, 1994), with DCi being the degree of curvature expressed in degrees for every 100 m.

Frequently, the value of the design speed (V_D) is unknown. For this reason, V_D has been estimated (Eq. (5)) according to the procedure described in Lamm et al. (1999) for roads in service.

$$V_D \approx V_{85CCR_S} = \frac{10^6}{8270 + 8.01\cdot C\bar{C}R_S} \quad (5)$$

with V_D expressed in km/h.

Finally, the coefficients of lateral friction (Montella and Imbriani, 2015) considered in the design step (f_{RA}) (Eqs. (6) and (7)) and demanded (f_{RD}) (Eq. (8)) according to the operating speed V_{85} (Lamm et al., 2001) are calculated (Criterion III):

$$f_{RA} = 0.925\cdot n\cdot f_T \quad (6)$$

$$f_T = 0.59 - 4.85\cdot 10^{-3}\cdot V_D + 1.51\cdot 10^{-5}\cdot (V_D)^2 \quad (7)$$

$$f_{RD} = \frac{{V_{85}}^2}{127\cdot R} - e \quad (8)$$

where 0.925 is a reduction coefficient related to tires, n is the utilizing factor (0.60 for roads in service), f_T is the tangential friction factor, and e is the superelevation expressed in %/100 (Lamm et al., 2001); in this case, the theoretical superelevation is considered according to the Spanish regulation norm (MFOM, 2016). Table 3 shows the three stability criteria considered together with their thresholds.

For a better understanding of the geometrical parameters considered, the reader is referred to Fig. 3.

2.3.3. Step 3. Potential risk assessment by decision tree

This step is crucial in the proposed methodology. The potential risk assessment (PRA) of the road stretch is determined by means of a triple stability criterion because of the influence that the lack of consistency has on the increase in the accident rate on roads (Lamm et al., 2001). This triple criterion comprises geometric consistency indexes that will define the three attributes/variables of inductive process. To achieve this, a data mining process is performed based on geometrical parameters exclusively, which is supported by the decision tree (DT) inductive algorithm.

DTs, also known as identification trees, are among the nonparametric methods more widely applied to supervise inductive learning (Soler Flores, 2014). Moreover, the implementation of an approach by DT has the advantage that it does not require prior probabilistic knowledge of the study phenomena (de Oña et al., 2013). In this case

Table 5
Geometric verification. LU-722 stretch 1 (Case study 1). (2) GC: Geometric consistency. (3) R: Absolute value of radius (m). (4)Ω: Central angle (gon). Please note that central angle refers to the azimuth variation between the ends of the circular alignment exclusively.

Ground truth General geometric parameters Length: 1,776.60 m Maximum/minimum radius: 177.40/29.08 (m) Number of curves: 14 $C\bar{C}R_S$:789 (gon/km) V_D: 68.5 ≈ 70 km/h Horizontal alignment. Circular arc				Obtained values by the RANSAC algorithm General geometric parameters Length: 1,776.35 m Maximum/minimum radius: 180.32/29.78 (m) Number of curves: 14 $C\bar{C}R_S$:788 (gon/km) V_D: 68.6 ≈ 70 km/h Horizontal alignment. Circular arc			
Curve	GC(2) indexes			Curve	GC(2) indexes		
$R^{(3)}/\Omega^{(4)}$	*CCRi*	*DCi*	*V85*	$R^{(3)}/\Omega^{(4)}$	*CCRi*	DCi	*V85*
91.45/33.19	697	63	66	91.67/39.56	695	63	66
127.64/70.18	499	45	74	127.81/73.65	498	45	74
115.50/33.86	552	50	72	120.19/26.47	530	48	72
69.99/51.86	1044	94	56	75.72/56.30	841	76	62
115.00/35.01	554	50	72	174.75/13.84	365	33	79
170.00/79.39	375	34	79	170.84/83.43	373	34	79
31.50/59.69	2022	182	33	29.78/55.49	2139	192	31
109.00/74.78	584	53	70	109.89/74.09	580	52	71
29.08/58.44	2191	197	31	30.17/73.72	2111	190	32
29.92/140.84	2129	192	31	29.94/152.75	2128	191	32
41.84/20.34	1522	137	43	41.22/32.37	1545	139	43
49.56/75.17	1285	116	49	49.77/79.21	1280	115	49
29.60/90.41	2152	194	31	30.04/124.70	2120	191	32
177.40/69.17	359	32	80	180.32/88.22	353	32	79

Table 6
Geometric verification. LU-722 stretch 2 (Case study 2).

Ground truth General geometric parameters Length: 3,160.62 m Maximum/minimum radius: 852.60 / 30.00 (m) Number of curves: 23 $C\bar{C}R_S$:713 (gon/km) V_D: 71.5 ≈ 70 km/h Horizontal alignment. Circular arc				Obtained values by the RANSAC algorithm General geometric parameters Length: 3,160.73 m Maximum/minimum radius: 947.42/30.25 (m) Number of curves: 23 $C\bar{C}R_S$:737 (gon/km) V_D: 70.6 ≈ 70 km/h Horizontal alignment. Circular arc			
Curve	GC(2) indexes			Curve	GC(2) indexes		
$R^{(3)}/\Omega^{(4)}$	*CCRi*	*DCi*	*V85*	$R^{(3)}/\Omega^{(4)}$	*CCRi*	*DCi*	*V85*
367.60/10.37	156	16	87	385.38/11.05	165	15	88
40.30/14.65	1422	142	42	45.21/29.57	1409	127	46
309.90/18.74	185	18	86	310.53/23.57	205	18	86
30.00/97.42	1910	191	32	30.25/128.21	2106	189	32
852.60/5.11	67	7	92	947.92/5.57	67	6	92
81.25/23.50	705	71	63	82.70/23.08	770	69	64
43.93/102.45	1304	130	45	44.02/119.92	1447	130	45
178.00/30.84	322	32	79	177.44/43.04	359	32	79
798.50/8.64	72	7	92	900.44/7.92	71	6	92
150.60/61.83	380	38	77	151.00/60.70	422	38	77
99.30/32.87	577	58	68	103.31/38.81	617	55	70
51.50/84.92	1113	111	50	51.41/94.67	1239	111	50
66.39/37.04	863	86	58	62.16/33.28	1025	92	56
113.05/34.53	507	51	71	98.89/40.54	644	58	68
73.64/27.99	778	78	61	75.02/28.84	849	76	62
40.45/43.75	1417	142	42	42.13/51.35	1512	136	43
59.80/24.37	958	96	55	60.22/42.26	1058	95	55
44.57/38.17	1286	129	45	45.63/78.09	1396	126	46
217.14/15.59	264	26	82	246.63/ 16.77	258	23	84
45.00/21.68	1273	127	46	40.36/23.64	1578	142	42
30.40/63.97	1885	188	32	33.11/71.03	1924	173	35
56.60/55.75	1012	101	53	56.29/85.90	1132	102	53
104.38/21.09	549	55	70	105.08/30.89	606	55	70

and according to Information Theory (Quinlan, 1996), a process for categorizing the analysed attributes/variables is performed.

In this framework, the key is to establish a classification model that minimizes the uncertainty regarding the risk predictions. To this end, the uncertainty for the information content of a discrete and random variable X, or self-information, can be adequately measured by the entropy function $H(X)$, as an appropriate indicator of the associated average uncertainty of a process (Cover and Thomas, 1991; Molina et al., 2016; Pearl, 1988), which is expressed as:

Table 7
Geometric verification. N-640 (Case study 3).

Ground truth				Obtained values by the RANSAC algorithm			
General geometric parameters Length: 2,149.41 m Maximum/minimum radius: 599.33 / 171.27 (m) Number of curves: 3 $C\bar{C}R_S$:183 (gon/km) V_D: 102.7 ≃ 100 km/h Horizontal alignment. Circular arc				General geometric parameters Length: 2,149.55 m Maximum/minimum radius: 601.20 / 171.49 (m) Number of curves: 3 $C\bar{C}R_S$:181 (gon/km) V_D: 102.9 ≃ 100 km/h Horizontal alignment. Circular arc			
Curve	$GC^{(2)}$ indexes			Curve	$GC^{(2)}$ indexes		
$R^{(3)}/\Omega^{(4)}$	*CCRi*	*DCi*	*V85*	$R^{(3)}/\Omega^{(4)}$	*CCRi*	*DCi*	*V85*
368.51/ 45.18	173	16	87	367.07/ 56.09	174	16	87
171.27/ 36.61	372	33	79	171.49/ 35.20	371	33	79
599.33/ 22.12	106	10	90	601.20/ 22.53	106	10	90

$$H(X) = -\sum_{x} p(x)\cdot log_2 p(x) \tag{9}$$

where P is the probability mass function of X. The entropy measure enables an assessment of the additional information required to specify a particular alternative (Barton et al., 2008), and therefore, reducing $H(X)$ by acquiring information is interpreted as reducing the uncertainty regarding X (Molina and Zazo, 2018; Molina et al., 2016).

On the other hand, to build a DT, it is necessary: (i) to establish a node sequences using attributes/variables (de Oña et al., 2013), in this case the three stability criterion and (ii) a node splitting criterion to form a tree. This latter condition leads on the one hand, to reduce complexity through removing the sections that provide little power to classify instances (Galathiya et al., 2012), and on the other hand to reduce classification errors, due to specialization in the training set (Körting, 2006). In this manner, node uncertainties are reduced (Singh and Gupta, 2014), overfitting phenomenon is avoided (Breiman et al., 2017; Kang and Choi, 2000) and a better predictions are achieved (Galathiya et al., 2012).

Overfitting the training data is a negative phenomenon of machine learning process as consequence of an excessive adaptation of the algorithm to the training data (Chicco, 2017; Domingos, 2012). This leads to erroneous classifications on unseen data, although DT may correctly perform on the training data (Kang and Choi, 2000). Here, as splitting or pruning criteria, it is essentially applied information gain, which is defined by the difference between the entropy of the node before (parent) and after (child) splitting respectively (de Oña et al., 2013; Singh and Gupta, 2014). It is worth highlighting that the highest information gain involves the highest reduction in entropy (Zhang et al., 2004).

In essence, this general framework (based on node sequences and splitting criterion) define the decision rules by which a DT, by itself and automatically, decides data split and draws its boundaries. For a complete background on DT theoretical construction process, please refer to benchmark works such as Breiman et al., (2017) and Quinlan (1993).

Alternatively, the inductive reasoning process is performed by open-source freeware WEKA data mining through the iterative J4.8 algorithm (de Oña et al., 2013; WEKA, 2018; Witten and Frank, 2005). J4.8 is WEKA's implementation of a decision tree learner (Witten and Frank, 2005), which is based on C4.5 algorithm (Al-Turaiki et al., 2016; Quinlan, 1993). Please note that these algorithms are inspired by entropy (de Oña et al., 2013), and the gain ratio "normalizes" the information gain (Quinlan, 1993), and besides, the overfitting is avoided by post-pruning process after the tree-creation, because of this is more effective method than pre-pruning to address overfitting problems; in this way the training data are suitably classified (Kang and Choi, 2000). J4.8 algorithm divides the dataset according to the best informative attribute/variable, selecting in every iteration the attribute/variable with the maximum gain ratio or highest reduction in entropy (Al-Turaiki et al., 2016; de Oña et al., 2013; Witten and Frank, 2005). This classification approach has easily interpretable results and comparable accuracy to other classification models as its main advantages (Al-Turaiki et al., 2016; Quinlan, 1993).

Finally, the inductive process of DT is exclusively trained with the circular alignments of the first three case studies (please see Section 3.3; Fig. 6). Previously, these circular alignments were categorized by a road safety expert into three level risks (high, medium and low). It is worth to highlight that safety expert provides the reference data that will be used into the DT training.

2.3.4. Step 4. Verification of results

A twofold process of verification, both in terms of geometric results and road safety is carried out. First geometrically through "as-built" horizontal alignment and secondly by risk validation, both provided by an expert road surveyor using manual delineation and design works from airborne images. Note the use of term "as-built", because it is very common that some changes can affect the original road design project during the construction phase (e.g. unpredicted and specific terrain conditions, etc.).

Therefore, we evaluate the accuracy of the geometry by comparing the horizontal alignment detected by the RANSAC algorithm in this work with the as-built horizontal alignment obtained manually by an expert road surveyor.

Secondly, risk validation is performed via a road safety expert, who also discretizes the risk levels of the circular alignments of fourth case study according to three levels (high, medium, low). In order to validate the achieved risk levels through the DT inductive process they will be compared with the safety expert ground truth.

On the other hand, DT process verification is carried out by means of: (1) Overall Accuracy (*OA*) and (2) Kappa concordance coefficient (*K*). *OA* is defined as the probability that an instance will be correctly classified according to the following expression:

$$OA_{(\%)} = \left(\frac{TP + TN}{N}\right)\cdot 100 \tag{10}$$

where *TP* and *TN* and the true positives and true negatives respectively, and *N* is the total number of instances considered. For its part, *K* coefficient is a statistic that measures pairwise agreement between a set of categorized data, correcting for expected chance agreement (Carletta, 1996; Garcia-Rodenas et al., 2017). The kappa coefficient is expressed as:

$$K = \frac{P(A) - P(E)}{1 - P(E)} \tag{11}$$

where $P(A)$ is the observed concordance proportion and $P(E)$ is the expected concordance proportion. In this equation, the numerator is the observed proportion, while the denominator is the maximum value that the numerator can take. K is defined in the range $[-1, 1]$. $K = 1$ is produced only when there exists concordance in 100% of the observations. $K = 0$ implies no agreement. Negative values indicate no agreement, but they are unlikely in practice.

This final step provides a full comprehensive reliability assessment, for both the defined geometric algorithms and the methodology developed as the decision-making process on road safety.

3. Results

3.1. Road segmentation

The effectiveness of the developed algorithm for the classification of the data acquired by MLS is shown in Table 4 and Fig. 4. The high level of reduction obtained from the original 3D point cloud acquired

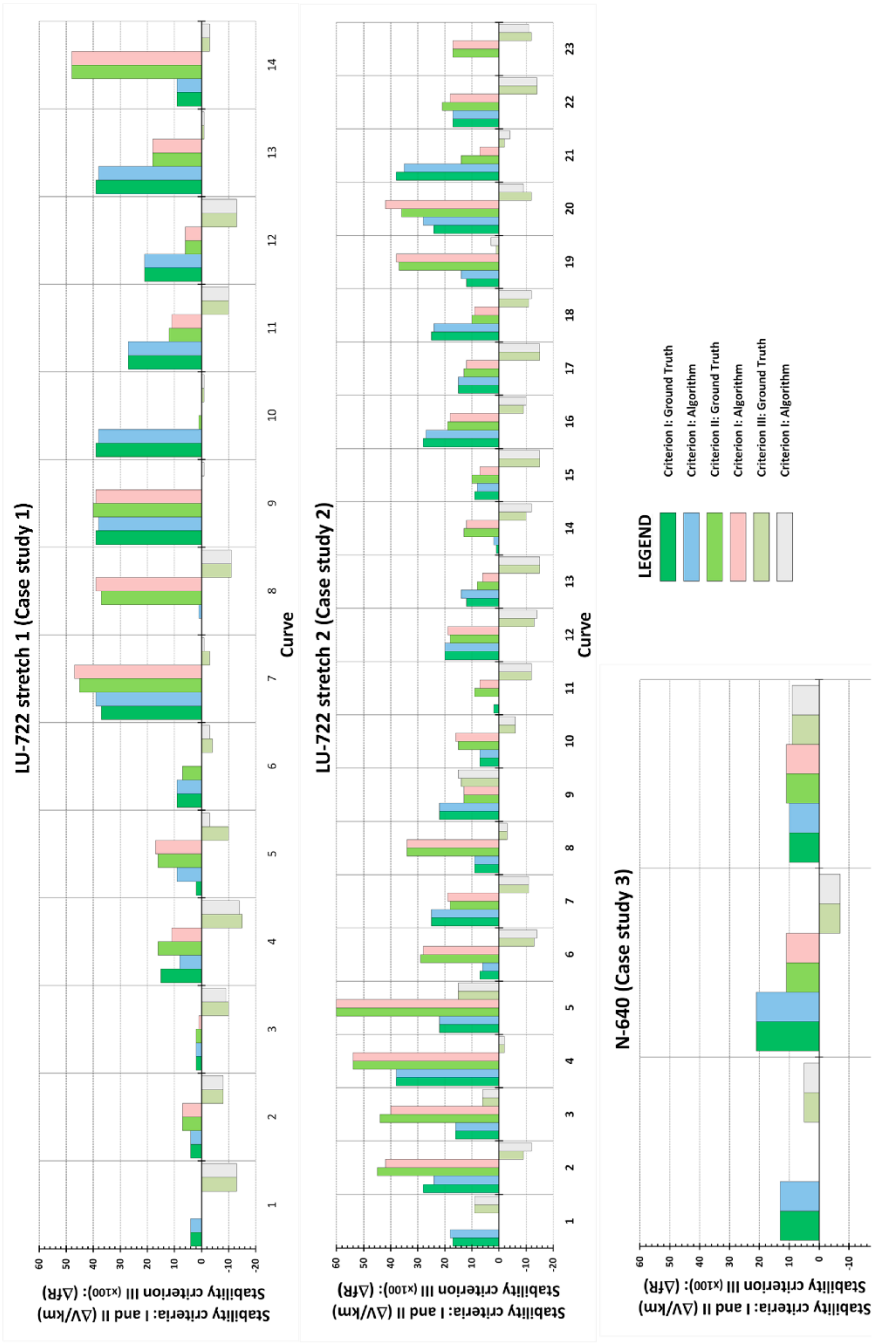


Fig. 5. Stability criteria. Ground truth versus algorithm developed.

(indicated in points m^{-2}) can be observed. In particular, a first classification of 2D points (expressed in points per metre or spacing m^{-1}) corresponding to road marks is obtained, which represents approximately 1% of the original 3D point cloud. Next, a second classification applied over road marks is applied to extract only the road axis that represents between 4.8% and 8.7% of the road mark points with a spacing between points of 10 cm. As commented in Section 2.3.1, to increase the quality of classification process an outliers filter is applied

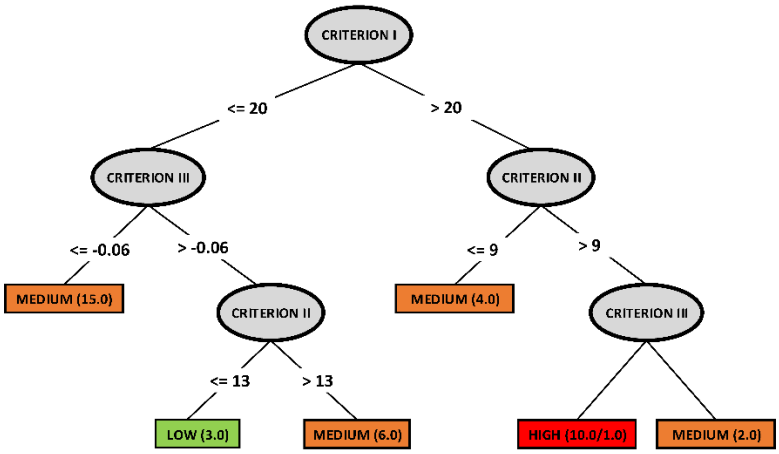


Fig. 6. Decision tree. Hierarchical structure developed. K = 0.946. Note: (A) number format indicates the total number of instances that reached the leaf. In the (A/B) format case, B is the total number of misclassified instances.

Table 8
Risk validation results. LU-722 (Case study 4).

		Classification (DT process)		
		Low	Medium	High
Ground truth (road safety expert)	Low	2	0	0
	Medium	1	5	0
	High	0	0	3

to the point cloud, in this case Statistical Outlier Removal (SOR) filter (PCL, 2018). The SOR parameters were a 20 points neighbourhood and a standard deviation multiplier threshold of 1.0. These values were determined on the basis of empirical tests. The high spacing achieved in the classification of points has allowed a more efficient development of the subsequent process of automatic extraction of the route (Fig. 2). This road classification approach has allowed us to apply a more efficient process in the automatic extraction of the horizontal alignment (Fig. 2).

3.2. Geometric verification

This step involves the first verification of the proposed methodology by as-built road projects, which were considered as "ground truth." Applying Eqs. (1)–(8) (see Step 2.3.2), the geometric consistency indexes and values of three joint criteria are calculated. The following tables (Tables 5–7) and Fig. 5 summarize the main results obtained, which are strictly focused on circular alignments.

Generally, the main geometric parameters (R and CCR_i), did not present significant discrepancies, exhibiting discrepancies of approximately 3.6% and 4.9%, respectively. From a quantitative point of view, these discrepancies are the average absolute value of the percentage of variation of the analysed geometrical parameter with respect to the ground truth (in these cases as-built projects). In this sense, it is worth mentioning that the minimum absolute discrepancies obtained for the lengths of the sections were between 0.11 m and 0.25 m, for a total of 7,086.6 m, as well as a practical coincidence for the global parameter $\overline{CCR}_S$ (Tables 5 and 7). In addition, the correct identification of the total number of circular alignments for each study case should be noted.

Alternatively, the maximum radius (R_{max}) and the minimum radius (R_{min}) exhibit values similar to those obtained during the verification, with relative discrepancies of 4% for R_{max} and 1% for R_{min}, and with absolute discrepancies between 1.87 m and 2.92 m for R_{max} and between 0.22 m and 0.70 m for R_{min}. Although an absolute maximum discrepancy between radiuses of 94.82 m was observed (852.60 m versus 947.42 m), this type of discrepancy is common in the case of circular alignments with central angles lower than 6 gons (MFOM, 2016). In addition, this absolute difference has not provided significant changes in the geometric consistency indexes and thus in the stability criteria, as seen in Table 6. It is also necessary to outline the correct detection of the different circular curves and even circular curves linked consecutively.

In more detail and from a statistical point of view, for a maximum relative discrepancy of 10%, the success rates for the geometric parameters R, CCR_i and DC_i were 89%, 78% and 87%, respectively; even if the admissible relative discrepancy is reduced by up to 5%, the success rates for each parameter could be perfectly acceptable, reaching success rates of 81%, 67% and 78%, respectively. In contrast, if the maximum relative discrepancy is fixed at 15%, the success rates increase to 98%, 91% and 95%, respectively. In particular, analysed parameter success rate was calculated as the average percentage of the values that comply with the fixed threshold discrepancy.

3.3. Risk validation

To establish the relationships, thresholds and hierarchy among the different criteria established, as well as the expert classification, an inductive learning process based on a decision tree was applied, as described in Section 2.3.3. Forty (40) instances were used to train the DT (total number of circular alignments for the first three case studies considered). Fig. 6 shows the results obtained.

The DT automatically establishes the discretization of attributes/variables (in this case the three stability criterion) based on the highest information gain (highest reduction in entropy) as a node splitting criterion to form a tree. In particular, determining the information gain and according to the gain ratio "normalizes," Criterion I of stability was identified as the most determining of the three criteria considered (gain = 0.415 bits). This criterion is the one that provides more information about the process and therefore the central node (root node) of the tree.

From the identified central node (Criterion I), the instances are split into child nodes, and recursively in each iteration, the attribute/variable with the maximum gain ratio is selected as node splitting criterion. It should be noted that, in the next levels, both criteria II and III are

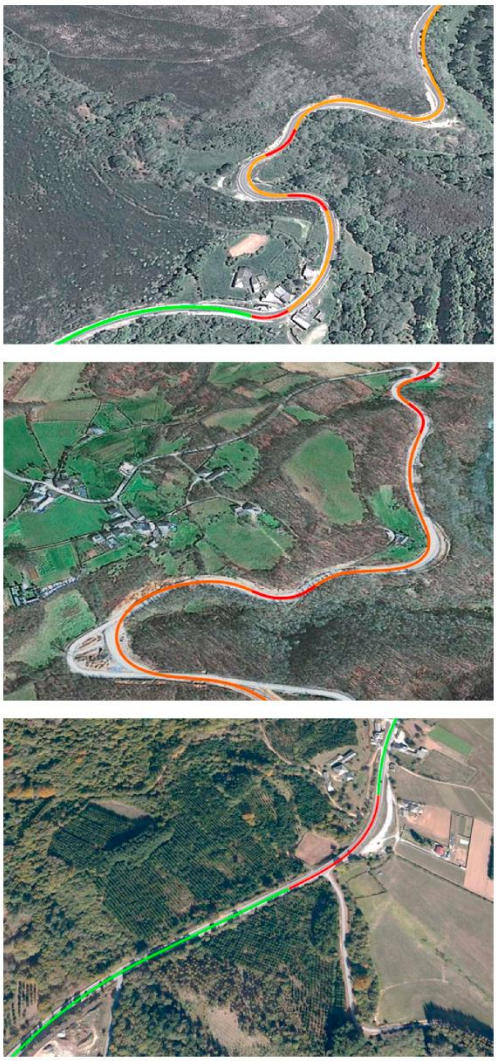


Fig. 7. Risk mapping for the different geometric elements of the horizontal alignment. Case study 1 LU-722 Stretch 1. Red line: High risk. Orange line: Medium risk. Green line: Low risk. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

applied, which implies that there is not a significant entropy improvement between them. Alternatively, the final tree may not necessarily be symmetric because of the application of pruning algorithms that reduce the complexity of the tree, keeping the final accuracy. By the post-pruning process the tree subsections that do not improve the classification results are removed. In this sense, it is worth to highlight that a tree subsection is reconverted to a leaf, whose output is defined only if this operation does not get worsen the prediction accuracy. In the case of J4.8 algorithm this is done in post-processing, once the final structure has been obtained, the full training data is back-fitted against the structure.

The final kappa coefficient obtained from the trained DT was 0.946, which entails a high degree of agreement. Then, this trained DT was applied to circular alignments of fourth case study, as a form of validating the classification obtained through the DT. The achieved overall accuracy (*OA*) and kappa coefficient (*K*) were 90.9% and 0.8553 respectively. These values show a high degree of agreement, and they validate the DT and its geometrical risk classification results. Table 8 shows confusion matrix results achieved. In this sense, the DT provided a suitable parameterization of the expert informal knowledge for road safety inspections based on the geometric parameters obtained.

Fig. 7 shows the evaluation risk resulting from the DT and corresponding to each one of the circular elements that define the horizontal alignment. It should be noted that the transition curves were categorized according to the risk level resulting from the circular elements that compose these curves.

4. Discussion and conclusions

The main contribution of this research, related to remote sensing, is to show the application possibilities that the MLS technology can offer to road infrastructure risk assessment by means of: (i) an automatic risk mapping of the road based on geometrical consistency indexes and (ii) accurate stability criteria derived from these indexes. This is done through the integral approach presented, which combines geometry and risk, using 3D MLS point clouds. This approach represents a novel method of evaluating security in roads for the engineering community and/or road managers.

Even though the achieved quality by the alternative approaches to MLS outlined in the introduction (e.g. airborne laser scanner or aerial imagery) is highly promising (around 90% and above), these quality values are not yet comparable with those values required in road projects (centimetric accuracy). Therefore, datasets coming from MLS are the suitable source for obtaining the horizontal road alignment from which to derive the geometric consistency indices.

Ultimately, a high degree of reliability was obtained in the extraction of the geometrical elements of the road horizontal alignment. This is evident considering the total number of circular elements detected, as well as the low discrepancies obtained for the main geometric parameters (R and CCR_i). The minimum differences observed in the geometric consistency indexes and in the stability criteria are also noteworthy.

The methodology presented in this paper for the evaluation of road safety, through the twofold approach (geometry plus risk) and supported by artificial intelligence techniques (inductive process through a decision tree), can suitably complement other safety approaches, such as EuroRAP, which is based on a statistical-qualitative approach, sensitive to the distorting effect of road accidents. In this sense, the methodology developed allows us to objectify the intrinsic risk that the geometric design confers to the road. Furthermore, the scalability of this approach is also a significant and additional advantage, since it could be applied to other road typologies or countries by means of adapting the formulations and the specific models to the road design standards in each country.

On the other hand, the majority of studies based on road safety are focused on external causes to the road geometry, such as probability of crashes, types of drivers, level of use and existing conditions of the pavement surface, among others, or searching patterns that can explain accident causes in a particular section of the road. However, the inherent risk that the geometric parameters themselves provide have received less attention. In this sense, it is worth to highlight that the exposed approach in this paper represents a novelty regarding the manner to assess road safety from MLS data and supported by an inductive reasoning process, based on a decision tree, which provides a potential risk assessment based on geometric parameters exclusively.

In future research, an extension of the categorization and evaluation of the road risk by DT would be desirable, incorporating those geometric constraints related to the vertical alignment (vertical agreement, *Kv* parameter and real superelevation) and completing it with the

available sight distance. Furthermore, the idea of this approach is to offer a future alternative to those roads where road marks are not available (situation that can exist in Spanish secondary roads) or roads with road marks in bad state or partially removed.

Finally, and in the framework of prevention and roadway deficiency detection, the exposed methodology together with the data acquisition from MLS can be an adequate and effective tool for road safety inspections. Through this approach, and taking into account the service life of road infrastructures, it would be possible to have a complete risk evolution according to the changes that may have been necessary to make to the original designs.

Acknowledgements

Authors thank especially, INSITU Ingeniería S.L. for the mobile laser scanner data and the equipment used for this work. This research has been partially supported by the INROAD project (TSI-100505-2016-019) Energy, Tourism and Digital Society Ministry (National projects: Strategical action/Call 2016).

References

Ai, C., Tsai, Y.J., 2015. Critical assessment of an enhanced traffic sign detection method using mobile LiDAR and INS technologies. J. Transp. Eng. 141 (5), 04014096.

Alexander, C., Tansey, K., Kaduk, J., Holland, D., Tate, N.J., 2010. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. ISPRS J. Photogramm. Remote Sens. 65 (5), 423–432.

Al-Turaiki, I., Alshahrani, M., Almutairi, T., 2016. Building predictive models for MERS-CoV infections using data mining techniques. J. Infect. Public Health 9 (6), 744–748.

Alshehhi, R., Marpu, P.R., 2017. Hierarchical graph-based segmentation for extracting road networks from high-resolution satellite images. ISPRS J. Photogramm. Remote Sens. 126, 245–260.

Andrasik, R., Bil, M., 2016. Efficient road geometry identification from digital vector data. J. Geogr. Syst. 18 (3), 249–264.

Arranz Justel, J.J., 2013. Diseño, optimización y análisis de sistemas basados en técnicas láser, para el modelado geométrico, registro y documentación, aplicados a entidades de interés patrimonial. PhD. Thesis. Technical University of Madrid, Madrid, Spain, 499 pp.

Azimi, S.M., Fischer, P., Körner, M., Reinartz, P., 2018. Aerial LaneNet: Lane marking semantic segmentation in aerial imagery using wavelet-enhanced cost-sensitive symmetric fully convolutional neural networks. arXiv preprint arXiv:1803.06904.

Barton, D.N., Saloranta, T., Moe, S.J., Eggestad, H.O., Kuikka, S., 2008. Bayesian belief networks as a meta-modelling tool in integrated river basin management - Pros and cons in evaluating nutrient abatement decisions under uncertainty in a Norwegian river basin. Ecol. Econ. 66 (1), 91–104.

Bitenc, M., Lindenbergh, R., Khoshelham, K., van Waarden, A.P., 2011. Evaluation of a LIDAR land-based mobile mapping system for monitoring sandy coasts. Remote Sens. 3 (7), 1472–1491.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 2017. Classification and Regression Trees. Chapman & Hall/CRC. Taylor & Francis Group, Washington, D.C., USA.

Cabo, C., Kukko, A., Garcia-Cortes, S., Kaartinen, H., Hyyppa, J., Ordonez, C., 2016. An algorithm for automatic road asphalt edge delineation from mobile laser scanner data using the line clouds concept. Remote Sens. 8 (9), 740.

Camacho-Torregrosa, F.J., Pérez-Zuriaga, A.M., Manuel Campoy-Ungria, J., Garcia-Garcia, A., 2013. New geometric design consistency model based on operating speed profiles for road safety evaluation. Accid. Anal. Prev. 61, 33–42.

Carletta, J., 1996. Assessing agreement on classification tasks: The kappa statistic. Comput. Linguistics 22 (2), 249–254.

Castro, M., Lopez-Cuervo, S., Parens-Gonzalez, M., de Santos-Berbel, C., 2016. LIDAR-based roadway and roadside modelling for sight distance studies. Surv. Rev. 48 (350), 309–315.

Cavendish, J.C., Field, D.A., Frey, W.H., 1985. An approach to automatic three-dimensional finite element mesh generation. Int. J. Numer. Meth. Eng. 21 (2), 329–347.

Chang, L., Chien, J., 2013. Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. Saf. Sci. 51 (1), 17–22.

Chang, L., Wang, H., 2006. Analysis of traffic injury severity: An application of non-parametric classification tree techniques. Accid. Anal. Prev. 38 (5), 1019–1027.

Chicco, D., 2017. Ten quick tips for machine learning in computational biology. Biodata Mining 10, 35.

Clode, S., Rottensteiner, F., Kootsookos, P.J., 2005. Improving city model determination by using road detection from lidar data, In: Joint Workshop of ISPRS and DAGM-CMRT05, Anonymous ISPRS, pp. 159–164.

Clode, S., Kootsookos, P.J., Rottensteiner, F., 2004. The automatic extraction of roads from LIDAR data, In: The International Society for Photogrammetry and Remote Sensing's Twentieth Annual Congress, Anonymous ISPRS, pp. 231–236.

Cover, T.M., Thomas, J.A., 1991. Chapter 1. Introduction and preview. Elements of Information Theory. John Wiley & Sons, Hoboken, New Jersey, USA.

da Costa, J.O., Prudencio Jacques, M.A., Cunha Soares, F.E., Freitas, E.F., 2016. Integration of geometric consistency contributory factors in three-leg junctions collision prediction models of Portuguese two-lane national highways. Accid. Anal. Prev. 86, 59–67.

de Oña, J., Lopez, G., Abellan, J., 2013. Extracting decision rules from police accident reports through decision trees. Accid. Anal. Prev. 50, 1151–1160.

DGT, 2017. < http://www.dgt.es/es/prensa/notas-de-prensa/2016/20160509-dos-cada-tres-fallecidos-accidente-trafico-producen-carreteras-convencionales.shtml > (Accessed 10/01, 2017).

Diaz-Vilarino, L., Gonzalez-Jorge, H., Bueno, M., Arias, P., Puente, I., 2016. Automatic classification of urban pavements using mobile LiDAR data and roughness descriptors. Constr. Build. Mater. 102, 208–215.

Domingos, P., 2012. A few useful things to know about machine learning. Commun. ACM 55 (10), 78–87.

Eftekharzadeh, S.F., Khodabakhshi, A., 2014. Safety evaluation of highway geometric design criteria in horizontal curves at downgrades. Int. J. Civil Eng. 12 (3), 326–332.

EU, 2008. Directive 2008/96/EC of the European Parliament and of the Council of 19 November 2008 on Road Infrastructure Safety Management European Commission.

EuroRAP, 2018. < http://www.eurorap.org/ > (Accessed 01/22, 2018).

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus - a paradigm for model-fitting with applications to image-analysis and automated cartography. Commun. ACM 24 (6), 381–395.

Fitzpatrick, K., Elefteriadou, L., Harwood, D.W., Collins, J.M., McFadden, J., Anderson, I. B., Krammes R.A., Irizarry N., Parma K.D., Bauer K.M., Passetti K., 2000. Speed prediction for two-lane rural highways. Publication No: 99-171, Federal Highway Administration (FHWA), Georgetown Pike, USA.

Galathiya, A., Ganatra, A., Bhensdadia, C., 2012. Improved decision tree induction algorithm with feature selection, cross validation, model complexity and reduced error pruning. Int. J. Comput. Sci. Inf. Technol. 3 (2), 3427–3431.

Garach, L., de Ona, J., Lopez, G., Baena, L., 2016. Development of safety performance functions for Spanish two-lane rural highways on flat terrain. Accid. Anal. Prev. 95, 250–265.

Garcia-Rodenas, R., Lopez-Garcia, M.L., Teresa Sanchez-Rico, M., 2017. An approach to dynamical classification of daily traffic patterns. Comput.-Aided Civ. Infrastruct. Eng. 32 (3), 191–212.

Gargoum, S., El-Basyouny, K., 2017. Automated extraction of road features using LiDAR data: a review of LiDAR applications in transportation (Contributed Paper). In: 4th International Conference on Transportation Information and Safety (ICTIS). 8–10 August 2017 Banff, Alberta, Canada.

Gonzalez-Jorge, H., Puente, I., Riveiro, B., Martinez-Sanchez, J., Arias, P., 2013. Automatic segmentation of road overpasses and detection of mortar efflorescence using mobile LiDAR data. Opt. Laser Technol. 54, 353–361.

Hatger, C., Brenner, C., 2003. Extraction of road geometry parameters from laser scanning and existing databases. Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci. 34 (3/W13), 225–230.

Holgado-Barco, A., Gonzalez-Aguilera, D., Arias-Sanchez, P., Martinez-Sanchez, J., 2015. Semiautomatic extraction of road horizontal alignment from a mobile LiDAR system. Comput.-Aided Civ. Infrastruct. Eng. 30 (3), 217–228.

Holgado-Barco, A., Gonzalez-Aguilera, D., Arias-Sanchez, P., Martinez-Sanchez, J., 2014. An automated approach to vertical road characterisation using mobile LiDAR systems: Longitudinal profiles and cross-sections. ISPRS J. Photogramm. Remote Sens. 96, 28–37.

Isenburg, M., Liu, Y., Shewchuk, J., Snoeyink, J., 2006. Streaming computation of delaunay triangulations. ACM Trans. Graphics 25 (3), 1049–1056.

Javanmardi, M., Javanmardi, E., Gu, Y., Kamijo, S., 2017. Towards high-definition 3D urban mapping: road feature-based registration of mobile mapping systems and aerial imagery. Remote Sens. 9 (10), 975.

Jung, S., Qin, X., Oh, C., 2016. Improving strategic policies for pedestrian safety enhancement using classification tree modeling. Transportation Res. Part A-Policy Practice 85, 53–64.

Kang, B., Choi, K., 2000. Automatic transliteration and back-transliteration by decision tree learning. In: Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC), pp. 1135–1141.

Körting, T. S., 2006. C4. 5 algorithm and multivariate decision trees. Image Processing Division, National Institute for Space Research–INPE Sao Jose dos Campos–SP, Brazil.

Kumar, P., McElhinney, C.P., Lewis, P., McCarthy, T., 2014. Automated road markings extraction from mobile laser scanning data. Int. J. Appl. Earth Observation Geoinformation 32, 125–137.

Kumar, P., McElhinney, C.P., Lewis, P., McCarthy, T., 2013. An automated algorithm for extracting road edges from terrestrial mobile LiDAR data. ISPRS J. Photogramm. Remote Sens. 85, 44–55.

Kwon, O.H., Rhee, W., Yoon, Y., 2015. Application of classification algorithms for analysis of road safety risk factor dependencies. Accid. Anal. Prev. 75, 1–15.

Lamm, R., Wolhuter, K.M., Beck, A., Rusher, T., 2001. Introduction of a new approach to geometric design and road safety (Contributed Paper). In: 20th South African Transport Conference 'Meeting the Transport Challenges in Southern Africa'. 16 – 20 July 2001 South Africa.

Lamm, R., Psarianos, B., Mailaender, T., 1999. Highway Design and Traffic Safety Engineering Handbook. MacGraw-Hill, New York, USA.

Lamm, R., Guenther, A.K., Choueiri, E.M., 1995. Safety module for highway geometric design. Transp. Res. Rec. 1512 (9), 7–15.

Lamm, R., Choueiri, E., Mailaender, T., 1991. Side friction demand versus side friction assumed for curve design on two-lane rural highways. Transp. Res. Rec. 1303, 11–21.

López, G., de Oña, J., 2017. Extracting crash patterns involving vulnerable users on two-lane rural highways. Securitas Vialis 9 (1), 1–13.

López, G., de Oña, J., Garach, L., Baena, L., 2016. Influence of deficiencies in traffic control devices in crashes on two-lane rural roads. Accid. Anal. Prev. 96, 130–139.

Marinelli, G., Bassani, M., Piras, M., Lingua, A.M., 2017. Mobile mapping systems and

spatial data collection strategies assessment in the identification of horizontal alignment of highways. Transportation Res. Part C-Emerging Technol. 79, 257–273.
Mc Elhinney, C., Kumar, P., Cahalane, C., McCarthy, T., 2010. Initial results from european road safety inspection (Eursi) mobile mapping project. In: Proceedings of the ISPRS Commission V Mid-Term Symposium Close Range Image Measurement Techniques 38, pp. 440–445.
MFOM, 2018. < https://www.fomento.gob.es > (Accessed 01/09, 2018).
MFOM, 2017. Mapa de Tráfico año 2016. Provincia de Lugo. Red de Carreteras del Estado.
MFOM, 2016. Ministerio de Fomento. Instrucción de Carreteras. Norma 3.1-IC.
MFOM, 1987. Ministerio de Fomento. Normas de Carreteras. 8.2-IC. Marca Viales.
Misaghi, P., Hassan, Y., 2005. Modeling operating speed and speed differential on two-lane rural roads. J. Transportation Eng.-ASCE 131 (6), 408–417.
Molina, J.-L., Zazo, S., 2018. Assessment of temporally conditioned runoff fractions in unregulated rivers. J. Hydrol. Eng. 23 (5), 04018015.
Molina, J.-L., Zazo, S., Rodriguez-Gonzalvez, P., Gonzalez-Aguilera, D., 2016. Innovative analysis of runoff temporal behavior through bayesian networks. Water 8 (11), 484.
Montella, A., Imbriani, L.L., 2015. Safety performance functions incorporating design consistency variables. Accid. Anal. Prev. 74, 133–144.
Morrall, J., Talarico, R., 1994. Side friction demanded and margins of safety on horizontal curves. Transp. Res. Rec. 1435, 145.
Ng, J.C.W., Sayed, T., 2004. Effect of geometric design consistency on road safety. Can. J. Civ. Eng. 31 (2), 218–227.
PCL, 2018. < http://pointclouds.org/ > (Accessed 07/15, 2018).
Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Francisco, USA.
Pérez-Zuriaga, A.M., Camacho-Torregrosa, F.J., García, A., 2013. Tangent-to-curve transition on two-lane rural roads based on continuous speed profiles. J. Transp. Eng. 139 (11), 1048–1057.
Puente, I., Gonzalez-Jorge, H., Martinez-Sanchez, J., Arias, P., 2013a. Review of mobile mapping and surveying technologies. Measurement 46 (7), 2127–2145.
Puente, I., Gonzalez-Jorge, H., Riveiro, B., Arias, P., 2013b. Accuracy verification of the lynx mobile mapper system. Opt. Laser Technol. 45, 578–586.
Quinlan, J.R., 1993. C4. 5: Programs for Machine Learning, Volume 1 of Morgan Kaufmann series in Machine Learning, Morgan Kaufmann, San Mateo, USA.
Quinlan, J.R., 1996. Learning decision tree classifiers. ACM Comput. Surv. 28 (1), 71–72.
Rasdorf, W., Findley, D.J., Zegeer, C.V., Sundstrom, C.A., Hummer, J.E., 2012. Evaluation of GIS applications for horizontal curve data collection. J. Comput. Civil Eng. 26 (2), 191–203.
Riveiro, B., Gonzalez-Jorge, H., Martinez-Sanchez, J., Diaz-Vilarino, L., Arias, P., 2015. Automatic detection of zebra crossings from mobile LiDAR data. Opt. Laser Technol. 70, 63–70.
Singh, S., Gupta, P., 2014. Comparative study ID3, Cart and C4. 5 decision tree algorithm: a survey. Int. J. Adv. Inf. Sci. Technol. 3 (7), 97–103.
Siskind, V., Steinhardt, D., Sheehan, M., O'Connor, T., Hanks, H., 2011. Risk factors for fatal crashes in rural Australia. Accid. Anal. Prev. 43 (3), 1082–1088.
Sitran, A., Delhaye, E., Uccelli, I., 2016. Directive 2008/96/EC on road infrastructure safety management: an ex-post assessment 5 years after its adoption. Transport Res. Arena Tra 2016 (14), 3312–3321.
Soler Flores, F., 2014. Estimación de sucesos poco probables mediante redes bayesianas. PhD. Thesis. University of Castilla-La Mancha, Ciudad Real, Spain, 271 pp.
Sujatha, C., Selvathi, D., 2015. Connected component-based technique for automatic extraction of road centerline in high resolution satellite images. Eurasip J. Image Video Process. 8.
Varela-Gonzalez, M., Gonzalez-Jorge, H., Riveiro, B., Arias, P., 2014. Automatic filtering of vehicles from mobile LiDAR datasets. Measurement 53, 215–223.
WEKA, 2018. Weka 3: Data Mining Software in Java. Machine learning Group at the University of Waikato. < https://www.cs.waikato.ac.nz/ml/weka/ > (Accessed 01/10, 2018).
Witten, I.H., Frank, E., 2005. Data mining. Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco, USA.
Xunta de Galicia, 2016. < http://civ.xunta.gal/seccion-organizacion/c/CIV_Axencia_Galega_de_Infraestruturas?content=Direccion_Xeral_Infraestruturas/Plan_aforos/seccion.html&std=plan-aforos.html > (Accessed 01/08, 2018).
Yang, B., Fang, L., Li, J., 2013. Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. ISPRS J. Photogramm. Remote Sens. 79, 80–93.
Yan, L., Liu, H., Tan, J., Li, Z., Xie, H., Chen, C., 2016. Scan line based road marking extraction from mobile LiDAR point clouds. Sensors 16 (6), 903.
You, K., Sun, L., Gu, W., 2012. Reliability-based risk analysis of roadway horizontal curves. J. Transportation Eng.-Asce 138 (8), 1071–1081.
Zhang, L.V., Wong, S.L., King, O.D., Roth, F.P., 2004. Predicting co-complexed protein pairs using genomic and proteomic data integration. Bmc Bioinf. 5, 38.
Zhou, L., Vosselman, G., 2012. Mapping curbstones in airborne and mobile laser scanning data. Int. J. Appl. Earth Observation Geoinf. 18, 293–304.

Capítulo IV

Automatización de procesado de nubes de puntos y ortoimágenes en edificios.

4. Automatización de procesado de nubes de puntos y ortoimágenes en edificios.

Este apartado contiene el artículo *"Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity"*, que se publicó en la revista de alto impacto ***Automation in Construction***, en Septiembre de 2019.

4.1. Resumen

El autoconsumo de energía solar fotovoltaica en edificios hace necesaria la instalación de sistemas fotovoltaicos (FV) mayoritariamente en tejados, aprovechando estas ubicaciones de los edificios tanto por la superficie disponible, como por la certeza de una gran cantidad de radiación solar anual recibida.

Puesto que los parámetros necesarios para un diseño adecuado de un sistema fotovoltaico son principalmente geométricos: orientación azimutal, ángulo de inclinación y dimensiones efectivas de las diferentes pendientes del tejado; los datos aéreos LiDAR y las ortoimágenes ofrecidas por el Instituto Geográfico Nacional, son las fuentes de datos más adecuadas para este propósito.

Este artículo presenta una metodología automática novedosa que combina LiDAR y procesamiento de ortoimágenes para caracterizar geométricamente tejados a nivel de pendiente y calcular su potencial solar fotovoltaico. En concreto, se utilizan los datos LiDAR para calcular la orientación y la inclinación de cada agua. Las nubes de puntos del Instituto Geográfico Nacional, tienen una resolución de 0,5 a 14 puntos/m^2, que es insuficiente para obtener la superficie de cada agua, pero sí permite extraer el resto de parámetros. Se utilizan las ortoimágenes para detectar los bordes externos de los tejados y combinando las nubes de puntos con estos bordes, se extraen los planos para identificar las aguas de los tejados y sus superficies. De este modo se genera una capa con las aguas de los tejados, con la orientación, inclinación y superficie de cada una de ellas.

La metodología desarrollada ha sido validada con resultados obtenidos de una nube de puntos 3D aérea de mayor resolución de las cubiertas en estudio. Se han probado diferentes ubicaciones y tipos de tejados con el fin de confirmar el funcionamiento de la metodología en diferentes condiciones, pudiendo caracterizar con precisión la geometría de la mayoría de los tipos de tejados, como tejados planos, cubiertas a un agua, tejados a dos aguas y tejados piramidales, a nivel de ciudad, barrio y edificio.

Se ha desarrollado una aplicación informática ***Ener3DMap-Solar Roofs***, en QT creator para realizar este procesado (apéndice B).

Palabras clave: caracterización geométrica, tejados, energía fotovoltaica, LiDAR, ortoimagen, potencial solar, automatización.

4.2. Publicación 2: Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity

Automation in Construction 109 (2020) 102965

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity

J. Martín-Jiménez*, S. Del Pozo, M. Sánchez-Aparicio, S. Lagüela

Department of Cartographic and Land Engineering, University of Salamanca, Hornos Caleros, 50, Ávila 05003, Spain

ARTICLE INFO

Keywords:
Geometric characterization
Roofs
Photovoltaic energy
LiDAR
Orthoimagery
Solar potential
Automation

ABSTRACT

Photovoltaic self-consumption in buildings requires the installation of photovoltaic (PV) systems mostly on roofs, taking the advantage of these building locations for both the available surface area and the certainty of a big amount of annual incoming solar radiation. Since the parameters required for a proper PV system design are mainly geometric: azimuth-orientation, tilt angle and effective dimensions of the different roof slopes; aerial LiDAR data and orthoimagery offered by the National Geographic Agencies became suitable data sources for this purpose, ensuring its availability for any city regardless of its location. This paper presents a novel automatic methodology that combines LiDAR and orthoimage data processing to geometrically characterize roofs at slope level and calculate their PV solar potential. The methodology developed has been validated against results obtained from a higher-resolution aerial 3D point cloud of the roofs under study. Different locations and roof types have been tested in order to confirm the performance of the methodology under different conditions, being able to accurately characterize the geometry of most types of roofs, such as flat roofs, gable or saddle roofs, single pitched roofs and pyramid roofs at city, neighbourhood and building level.

1. Introduction

The use of solar energy has been a subject of increasing interest in the last years, supplying the 6.6% of the total primary energy production of renewable energy in Europe in 2017 [1] and with a great increase regarding the contribution of 2% to the total energy production in 2007. In addition, the growing concern about the climate change issue has led to a rising willingness of consumers to become self-consumers, also encouraged by national governments in an attempt to increase their share in renewable energy consumption [2].

Within this framework, and thanks to the reduction of investment costs of photovoltaic (PV) panels due to the maturity of this technology [3], PV solar energy constitutes an option of interest for the building users, especially for its exploitation in their main places of consumption (home and work). In a commitment towards the distributed energy generation and self-consumption to avoid energy losses due to transportation, the ideal place for installing PV systems is in the buildings themselves. Specifically, roofs are chosen as the main location for the PV installation due to their favourable position with respect to the Sun [4]. That is, roofs ensure the minimum energy-loss and optimize installation costs, as well as maximize the reception of Sun radiation. For this reason, the location, characteristics and dimensions of roofs are critical when analysing the feasibility, production and profitability of PV installations [5]. In this regard, there are tools available to estimate the PV energy production of roofs, such as Google SunRoof [6]. While this kind of tools allows the analysis of costs and profitability of PV installations on roofs, they have limitations derived from: (i) the use of two-dimensional data, (ii) ideal values of solar production and (iii) disregarding the real yields of the panels.

With the aim at avoiding the 2D limitation, LiDAR data and photogrammetry have arisen as providers of the information about the third dimension, allowing the simultaneous computation of the roof parameters for several buildings [7–11]. However, most of the existing approaches come from commercial software [7] or calculate the solar potential of roofs without discretizing between the different roof slopes and their different azimuth-orientations and tilt angles [8]. Other approaches make use of aerial imagery for 3D building reconstruction, in such way that the accuracy commonly associated to the spatial resolution of this imagery is lost in the generation of the Digital Surface Model [9]. Provided the complexity and variety of existing roofs, some works [12] present a methodology focusing on pitched-roofs. However, this approach is object-oriented, and computes the solar energy potential for individual buildings. While this is a valid approach in regions with single family houses, it cannot be applied in urban areas where the

* Corresponding author.
E-mail addresses: joscabula@usal.es (J. Martín-Jiménez), s.p.aguilera@usal.es (S. Del Pozo), mar_sanchez1410@usal.es (M. Sánchez-Aparicio), sulaguela@usal.es (S. Lagüela).

https://doi.org/10.1016/j.autcon.2019.102965
Received 5 April 2019; Received in revised form 10 September 2019; Accepted 11 September 2019
Available online 24 October 2019

Table 1
Comparison between the methodology proposed and the state of the art.

Existing works	Reference	Strengths of the proposed methodology compared to existing works
Arefi & Reinartz, 2013	[16]	Automatic computation of roof parameters, not limited to the geometric modelling but including data exploitation.
Fan et al., 2014	[17]	More robust and more accurate results due to the combination of two data sources (LiDAR and aerial orthoimagery) instead of being dependent on the resolution and precision of a single data source (LiDAR data).
Ghaffarian & Ghaffarian, 2014	[18]	Buildings are not only detected but parametrized, providing characterization data useful for further purposes such as for PV potential computation.
Martín et al., 2015	[8]	Lower computation requirements thanks to the processing at roof level.
Wang et al., 2015	[19]	Symmetric and asymmetric roof structures can be analysed.
Ghaffarian & Ghaffarian, 2016	[20]	Extraction of roof parameters useful for further purposes such as PV potential computation: azimuth-orientation, tilt angle and dimensions of each roof slope.
Kiti et al., 2017	[14]	The methodology does not rely on third-party software for any step, including the computation of solar radiation.
Li and Liu, 2017	[12]	The 3D modelling regarding the extraction of parameters of interest (azimuth-orientation, tilt angle and dimensions of each roof slope) is fully automatic.
Palmer et al., 2018	[7]	Self-developed methodology not relying on third-party software. No manual procedures are required.
Zhao et al., 2018	[21]	Alternative methodology with which not only extract roofs but also obtain their azimuth-orientation, tilt angle and effective dimensions.
Pirotti et al., 2019	[22]	Façades are automatically discarded on the first steps of the algorithm. Roofs are detected and modelled with the parameters required for the determination of solar radiation.

solar radiation of buildings is affected by the radiation of their surroundings [13]. In this sense, [14] computed roof solar potential based on LiDAR data and for complete neighbourhoods, taking into account how the buildings of the surroundings affect in terms of projected shadows. However, the use of LiDAR data alone presents limitations in terms of occlusions, spatial resolution and horizontal accuracy that can be solved by its combination with aerial imagery [15]. A summary of the contribution of the methodology proposed regarding the state of the art is presented in Table 1.

Accordingly, the proposed methodology allows an accurate estimation of solar production in roofs based on their 3D characterization in terms of location, azimuth-orientation, morphology and dimensions for most existing roof types (Fig. 1). This is possible due to the optimization of each data source for their combined processing in order to compute such geometric parameters. While LiDAR data is used for the computation of the angular parameters of the roofs, aerial orthoimagery provides accurate information about the contours of the roofs due to its high spatial resolution. The main contribution of this paper is the full automation of the roof characterization, discriminating between the different roof slopes per building, and processing several buildings simultaneously (going from city to roof scale). In this way, it is possible to perform studies at city, neighbourhood or building level obtaining high precision results with a great level of detail for all cases.

In order to describe the proposed methodology, the paper is organized as follows. After this introduction, both the data sources and the workflow of the methodology proposed are described in detail in Section 2. The evaluation system and experimental results at city, neighbourhood and building level are firstly described in Section 3 and then discussed in Section 4. Finally, the conclusions and future works are summarized in Section 5.

2. Data sources

Given that the main purpose of this research work is to present a standard methodology for the decision-making regarding PV installations in roofs, the focus is set on the use of available free geomatic data that guarantees precision in the results. The methodology makes use of LiDAR data and aerial orthoimagery in a combined way (Fig. 2) to automatically obtain the location, azimuth-orientation, tilt angle and dimensions of each of the slopes of the roof analysed. These are the main parameters required to perform an appropriate query to any solar radiation database in order to obtain PV potential of the roof. In addition, the importance of an accurate estimation of these parameters when performing feasibility and profitability studies of PV installations on roofs should be highlighted.

2.1. Data description

As already mentioned, the methodology developed uses two data sources as input: i) 3D point clouds from aerial LiDAR technology and ii) aerial orthoimagery, both widely used and available data sources in most countries (especially in Europe, through the INSPIRE standard [23]). In this case, both data sources are in the same reference system (ETRS89) and are available at the download centre of the Spanish National Geographic Agency [24]. This national agency analyses changes in the Spanish territory based on aerial data acquired every 2 or 3 years.

While, at first instance LiDAR technology seemed to be ideal and enough to perform 3D geometric analysis, it is limited by its spatial resolution, horizontal accuracy and the concavity/convexity in some building boundaries [15,25]. Since one of the key parameters when calculating PV solar potential is the total available area, the detection of

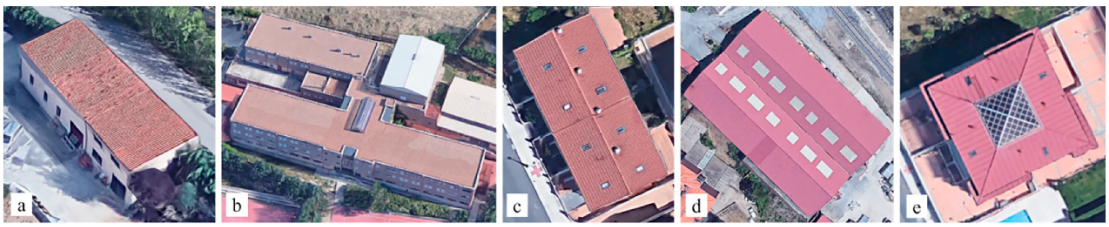


Fig. 1. Typology of roofs solved satisfactorily with this novel methodology: single pitched roofs (a), flat roofs (b), gable or saddle roofs (c and d) and pyramid roofs (e). Source of the images: Google Maps ©.

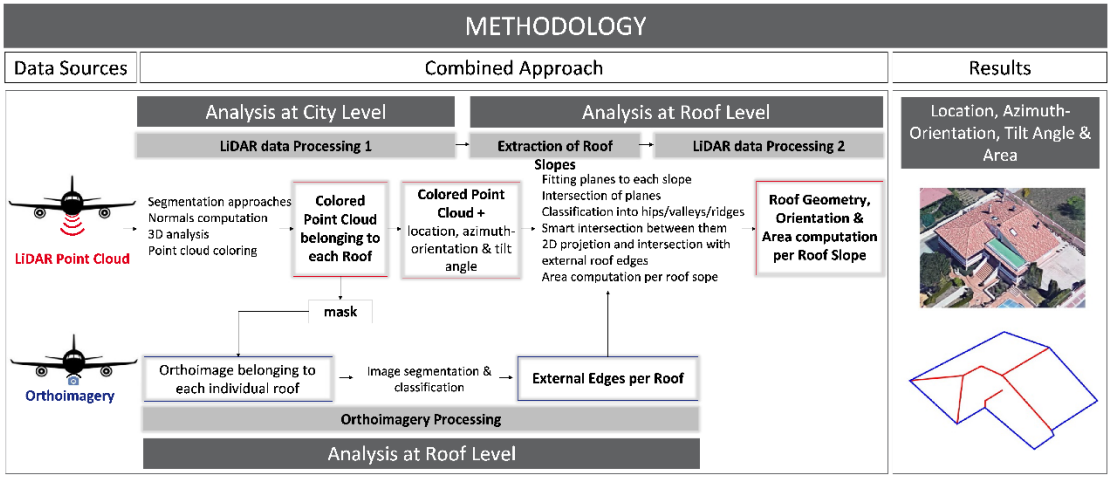


Fig. 2. Workflow of the methodology proposed for the 3D characterization of roofs in order to assess their PV potential.

Table 2
Main features of the input data used by the proposed methodology.

Data source		
Aerial LiDAR data	Accuracy	± 20 cm (Z), ± 40 cm (XY)
	Spatial resolution	0.5–1 p/m^2
	Geodesic reference	ETRS89
	Wavelength	1064 nm (infrared)
	Coverage	4 km^2
	Information stored in the point cloud	XYZ coordinates, RGB data, point source id, user data, scan angle, flight line edge, scan direction, returns, return number, GPS time, classification and infrared intensity
	Type of LiDAR	Optech ALTM 3025 (time-of-flight)
	Format	LAZ 1.1
Aerial orthoimages	Accuracy	± 50 cm
	Spatial resolution	25 cm
	Geodesic reference	ETRS89
	Type of camera	Vexcel UltraCam (large format aerial digital camera)
	Format	ECW

exterior roof edges is required when evaluating solar potential at roof level. In this sense, LiDAR data is limited and should be hybridized with another data source that allows boundaries detection with guarantees of accuracy, such as aerial orthoimagery (Table 2).

Therefore, in the proposed methodology, aerial LiDAR data is used to: (i) segment the roofs from the whole LiDAR point cloud, (ii) extract their location, (iii) detect the number of slopes per roof, (iv) calculate their azimuth-orientation and (v) tilt angle per roof slope, (vi) define the roof structure (ridge, rips and valleys) and (vii) serve as a mask for the identification of roofs in the orthoimagery processing. Then, orthoimagery is applied to determine the exterior roof edges which serve to delimit the area belonging to each specific roof slope. In this way, both technologies offer their advantages and their combination ensures accuracy in terms of planimetric, altimetric and angular characterization.

2.2. Aerial LiDAR data

This data source offers ranging accuracy for 3D documentation, characterization and reconstruction, especially in terms of altimetry (Table 2). However, it is crucial to analyse its limitations in terms of spatial resolution to discern which elements could be properly characterized and which could not. The aerial LiDAR data used in this study is available in .LAZ files at the IGN data repository [24]. This cartographic data has very low spatial resolution, specifically, between 0.5 and 1 points per square meter depending on the flight strip overlap. Therefore, even though it is a high precision 3D data source, it is inadvisable to use it for calculating roof dimensions since it would result in their underestimation.

2.3. Aerial orthoimagery

The aerial orthoimages from the Spanish National Orthophotography Program (PNOA, Spanish acronym) [24] are the result of a photogrammetric network captured by a large format digital camera (Vexcel UltraCam). A longitudinal overlap between images of 60% and a transversal overlap ≥25% was considered for this photogrammetric network. As a result, an orthoimage with a GSD of 25 cm (Table 2) is obtained, resulting in a great advantage in terms of spatial resolution. This resolution allows one to obtain an accurate estimation of the roof's external edges by means of digital image procedures such as edge detector algorithms and segmentation procedures. Despite its high spatial resolution, this bidimensional product is not useful to define the angular roof parameters on its own, i.e. azimuth-orientation, tilt angle and real area of each roof slope. Regarding its planimetric accuracy, this cartographic product offers a precision of ± 50 cm.

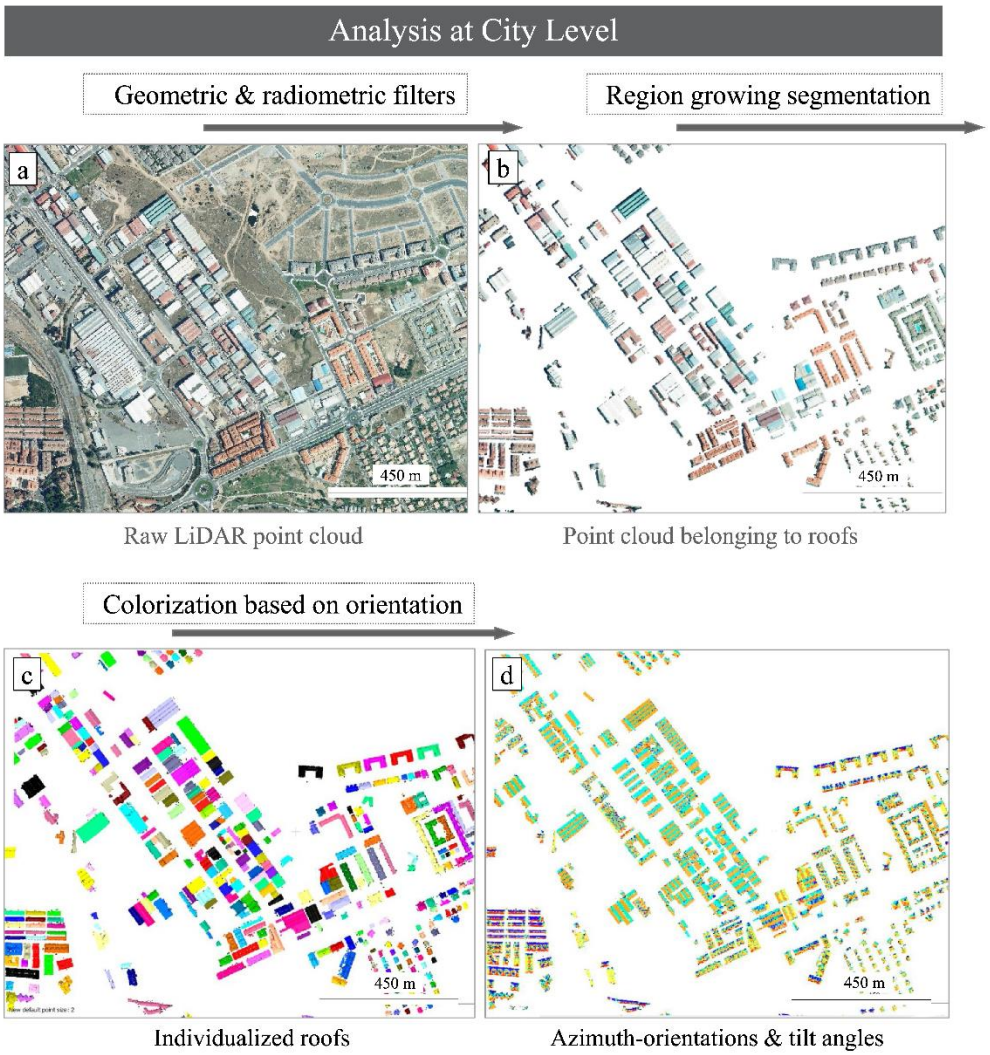


Fig. 3. LiDAR data processing at city level. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

3. Methodology

As shown in the workflow diagram (Fig. 2), the procedure begins with the LiDAR data processing, from which the layer corresponding to individualized roofs with information about their location, azimuth-orientation and tilt angle is obtained. Then, the point cloud corresponding to each individual roof serves as a mask and input data to the orthoimagery processing where it is used to crop the matching orthoimage regarding the buildings under study. After the appropriate image processing (Section 2.3), the exterior roof edges are obtained. Finally, these edges serve as external delimitation to accurately quantify the area of each roof slope identified through the LiDAR processing. Below, all the phases and algorithms implemented in the proposed methodology are described in detail.

3.1. LiDAR data processing (LDP)

This approach integrates several steps and two correlative working scales: one at city level (CL) (Fig. 3) and the other at individual roof level (RL) (Fig. 6). To operate at individual RL, the proposed methodology integrates a sequential segmentation procedure based on both geometric and radiometric criteria. All the algorithms and methods implemented are described below in their order of application.

3.1.1. LDP at city level: location, azimuth-orientation and tilt angle

The result is the point cloud corresponding only to roof areas coloured according to the azimuth-orientation (Fig. 4b) and with information regarding both roof azimuth-orientation and tilt angle after going through the following 8 steps:

- CL1-Removal of points belonging to the ground. Only those points

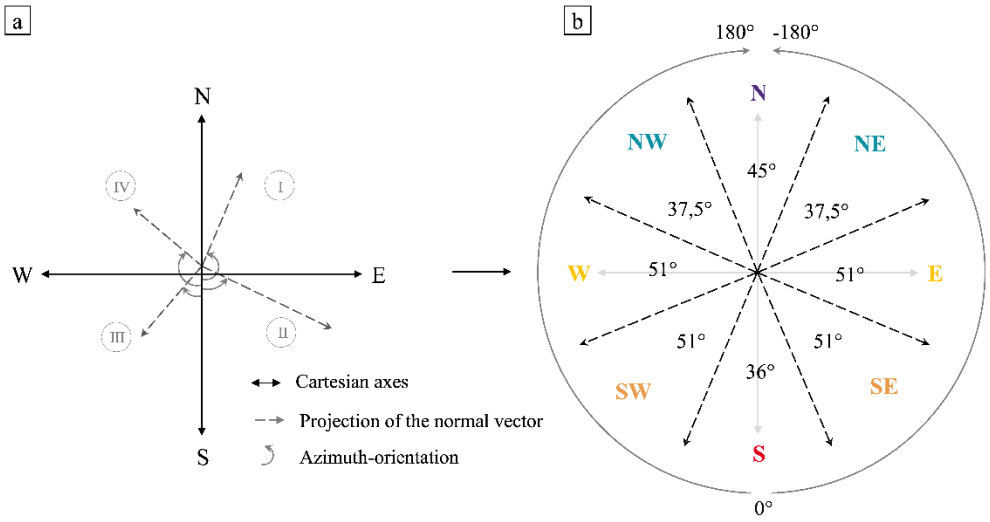


Fig. 4. (a) The four quadrants in which the Cartesian axes are divided as well as the projection of the normal vector on them. (b) The colour assigned to each point according to the 8 orientations established in the Spanish Technical Building Code (CTE [30]). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with a single LiDAR return (Table 2) continue in the process, in order to ensure that they correspond to solid surfaces (Fig. 3a–b).

- CL2-Removal of points belonging to vegetation areas. A previous computation of the Normalized Vegetation Index (NDVI) [26] is required (Fig. 3a–b). It is quantified thanks to the red and near infrared information collected by the LiDAR (Table 1) and by applying Eq. (1).

$$NDVI = \frac{NIR - RED}{NIR + RED}, \; where(-1 < NDVI < 1) \tag{1}$$

A NDVI threshold of 0.3 [26,27] was established in such a way that points with higher NDVI value are considered vegetation and consequently removed.

- CL3-Removal of points belonging to building façades. This segmentation is based on the Z-component of the normal vector (**Nz**) of each 3D point. Each normal vector is estimated through the adjustment of a 2D subspace that is tangent at the point of interest based on pairwise point relationships [28]. Specifically, a surface is adjusted to the 8 nearest neighbour points of each 3D point analysed (search resolved by k-d tree) within a maximum radius of 4 m. In this calculation, the covariance matrix of the 8 points is analysed and the normal vector of the surface adjusted is determined as the normal vector of the analysed point. A threshold of 0.15 was established in such a way that any point with a **Nz** value below this threshold is considered a point corresponding to façades due to the verticality of the surface adjusted.
- CL4-Removal of remaining noise points. Possible residual points from the previous steps are deleted by applying a Statistical Outlier Removal (SOR) filter [29]. A threshold of 1-m deviation in 30 neighbour points was established. After this step, the resulting point cloud corresponds only to roof areas (Fig. 3b).
- CL5-Individualization of roofs. Each roof is individualized by means of a region growing segmentation establishing a cluster size of 20 cm in the 20 nearby neighbours (Fig. 3c).
- CL6-Azimuth-orientation computation. The azimuth-orientation and tilt angle of the roofs are calculated by analysing the X-Y-Z-components of the normal vector (**Nx**, **Ny** and **Nz**) of each point (Fig. 3c–d). This requires projecting the normal vector to the horizontal plane and evaluating its corresponding quadrant (Fig. 4a). The azimuth-orientation ($-180° < Azi < 0 < Azi < 180°$) is calculated by Eqs. (2), (3) and (4), depending on the normal vector, so if it belongs to the 1st, 2nd and 3rd, or to the 4th quadrant respectively (Fig. 4).

$$Azi(°) = \tan^{-1}\frac{Nx}{Ny} - 180 \quad (1^{st} \text{ quadrant}) \tag{2}$$

$$Azi(°) = \tan^{-1}\frac{Nx}{Ny} \quad (2^{nd} \text{ and } 3^{rd} \text{ quadrant}) \tag{3}$$

$$Azi(°) = 180 + \tan^{-1}\frac{Nx}{Ny} \quad (4^{th} \text{ quadrant}) \tag{4}$$

- CL7-Point cloud colorization. Once the azimuth-orientation is computed, the point cloud belonging to roofs is coloured (Fig. 3d) following the criteria of Fig. 4b.
- CL8-Tilt angle computation. This is the last step of the LiDAR data processing at CL. It is calculated by a simple trigonometric process outlined in Fig. 5 and Eq. (5).

$$Tilt(°) = 90 - \left[\tan^{-1}\left(\frac{\boldsymbol{Nz}}{\sqrt[2]{\boldsymbol{Nx}^2 + \boldsymbol{Ny}^2}}\right)\right] \tag{5}$$

3.1.2. LDP at roof level: internal roof edges and dimensions of roof slopes

In order to calculate the azimuth-orientation, tilt angle and dimensions of each roof slope it is necessary to determine the number of roof slopes per building under study. However, before starting this analysis, each roof has to be classified as flat or non-flat. Specifically, a threshold of 0.5% is established to highlight the difference between the vertical and the horizontal component, in such a way that if the difference (in absolute value) between the mean values of the **Nz** and **Nx** and between the mean values of the **Nz** and **Ny** is < 0.5% the roof is classified as flat roof. It should be noted that these roofs are characterized by a point cloud with mixed colours in reference to the azimuth-orientation colorization (Fig. 11-2d). In such cases, the azimuth-orientation is established as "flat", the tilt angle is the mean/average

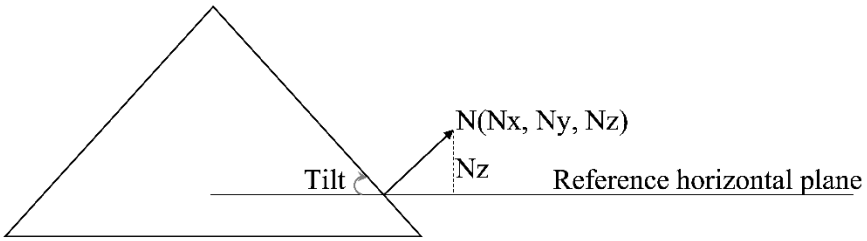


Fig. 5. Tilt angle of a roof slope based on the analysis of the normal vector.

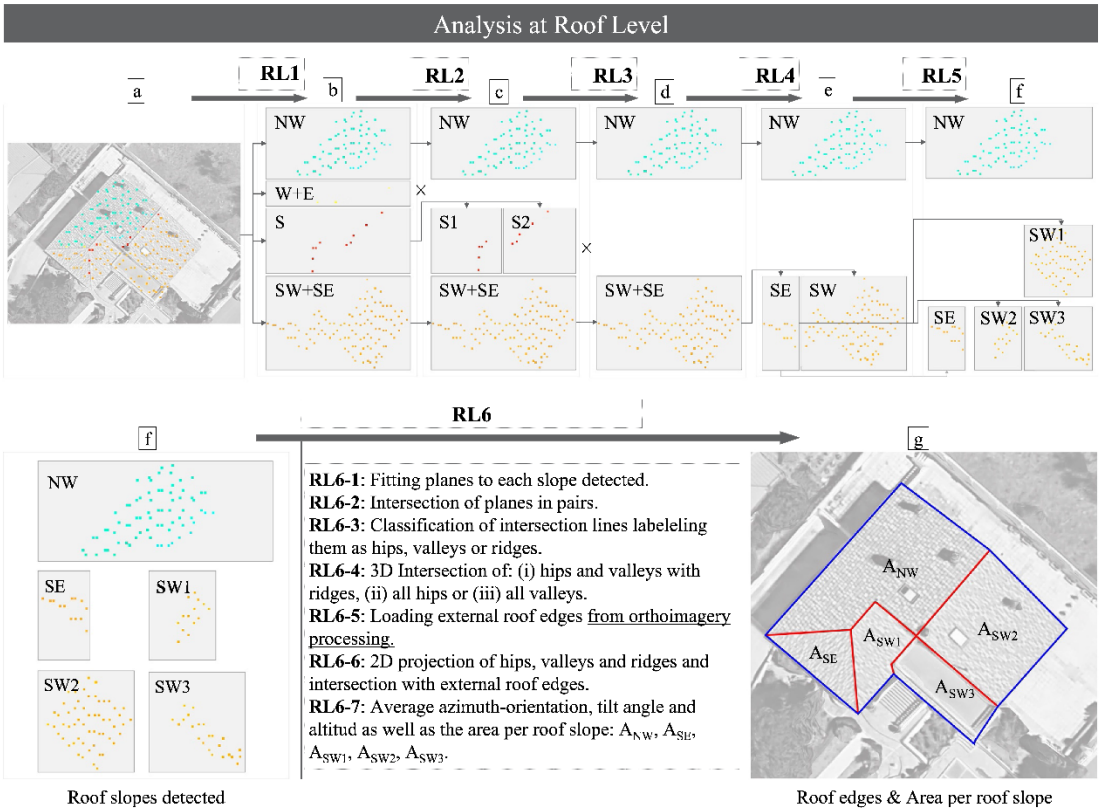


Fig. 6. Workflow of the LiDAR data processing at roof level for non-flat roofs. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

value of all points of the roof and the area is obtained with the external roof edges from the orthoimagery processing.

For non-flat roofs, the number of roof slopes and their main geometric parameters are obtained after applying the following 6 steps outlined in Fig. 6, where each colour indicates a different orientation of the roof (step CL7):

- RL1-Colour Segmentation. Taking as basis the colours assigned to each point in the step CL7, the point cloud of each roof is segmented into as many groups as colours.
- RL2-Region growing segmentation. Each of the point groups after RL1 is segmented based on region growing establishing a cluster size of 10 cm in the 10 nearby neighbours. Some groups of points disappear because they are considered noise due to their low point density. This is the case of the yellow points (West and East facing) in Fig. 6.
- RL3-Spatial dispersion filter. At this step, spatially dispersed groups of points are removed, even if they have a good point density. The criterion to consider spatial dispersion was the ratio between the variance (in both X and Y dimensions) and the number of points. A threshold of 1.3 was established in such a way that groups of points with higher dispersion value, are removed, given that it is highly probable that those points do not belong to any roof slope. They should belong to roof edges due to the change of curvature in those areas. This is the case of red points (South facing) in Fig. 6.
- RL4-Positive/Negative azimuth segmentation. Those groups of

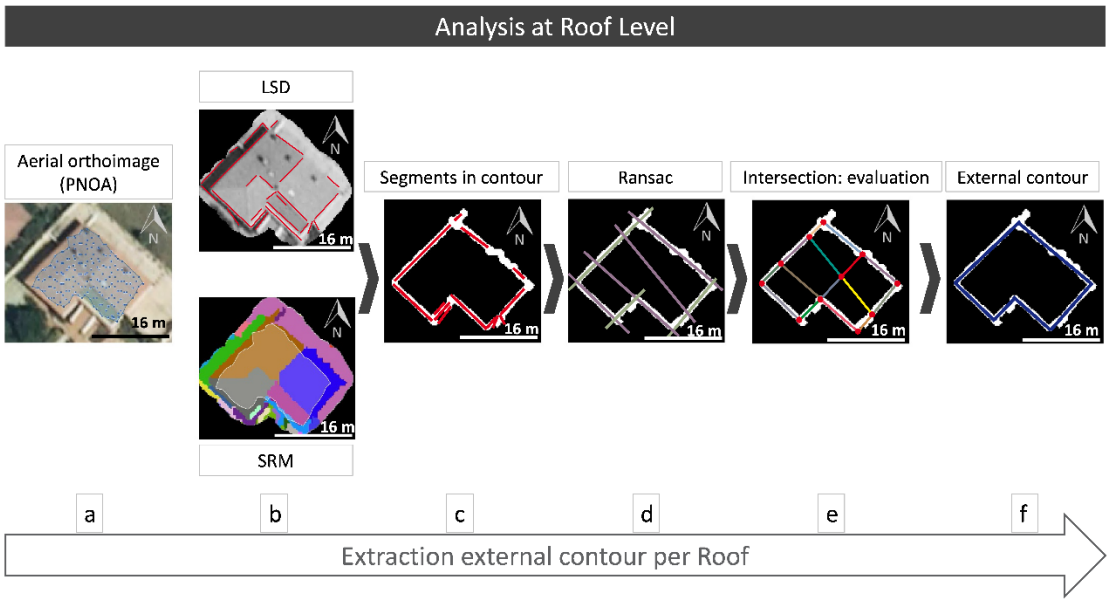


Fig. 7. Methodological workflow to get the external edges of each roof.

points that have passed the RL1, RL2 and RL3 segmentations are divided into groups of points with positive and negative azimuth in case of having both. This is the case of orange points in Fig. 6 (South-East and South-West facing).

- RL5-Inclination dispersion filter. In this step, the tilt angle value of all the points belonging to each group is analysed in terms of dispersion. If a high change of tilt angle is computed, the group is divided into as many subgroups as tilt angles are computed. This is the case of the South-West orange group in Fig. 6, which was divided into 3 different groups (SW1, SW2 and SW3) belonging to 3 different roof slopes.
- RL6-Roof edge parametrization. This is the final step of the LiDAR processing where each point group, representative of each roof slope, is subjected to 7 sequential steps after which the internal roof edges are parametrized. Thanks to the detection of these boundaries and those coming from the orthoimagery processing (Section 2.3) it is possible to accurately estimate the dimensions of each roof slope as well as the average value of the corresponding azimuth-orientation and tilt angle:

  RL6-1. Fitting planes to each of the 3D point groups throughout the M-estimator Sample and Consensus (MSAC) algorithm [31]. The maximum distance threshold established for a point to be considered by the MSAC algorithm is 0.4 m, compatible with the planimetric precision of LiDAR data (Table 2).

  RL6-2. Intersection of planes in pairs based on a neighbourhood criterion. Given the low resolution offered by the LiDAR data used in this study and the loss of points through the previous processing steps, an average distance adaptive threshold of 8 m (adaptation regarding the size of the building) is established to determine the neighbourhood between roof slopes.

  RL6-3. Classification of the intersection lines into (i) ridges, (ii) valleys and/or (iii) hips depending on their slope in Z axis. If the slope is < 2% it is considered as a ridge. The remaining lines are classified as valleys if the 3D points of the surroundings have higher Z-coordinate and as hips if they have lower Z-coordinate.

  RL6-4. In case there are the three types of lines, lines classified as valleys and/or hips are 3D intersected with those classified as ridges, and in case there are only valleys or hips, they are intersected in pairs.

  RL6-5. In this step, the orthoimagery processing is linked to the LiDAR processing through the incorporation of the outer edges defined in the orthoimagery process (in blue). These external boundaries will serve as outer limits to the 2D intersection of ridges, hips and valleys in order to define precisely each roof slope.

  RL6-6. The 2D projection of ridges, hips and valleys and the intersection with outer limits is performed. At this point, each roof slope is accurately defined.

  RL6-7. Finally, the area of each roof slope defined is computed, in 2D firstly and in 3D later thanks to the tilt angle value.

3.2. Orthoimagery processing: external roof edges and total roof area

The orthoimagery processing for the external roof edges detection and the computation of the total roof area consists on a triple-stage procedure (Fig. 7).

3.2.1. Roof mask

A combined processing between the individual roofs (RI.) from the LiDAR procedure and the orthoimagery is carried out to go from the CI. to the RI. in the orthoimagery. The result of this combination allows one to know the area of the orthoimage in which the roof is located. Making a mask from the contour of the point cloud corresponding to each individual roof, a positive 2 m buffer is created so that the roof and its immediate surroundings fit within. This process is required given the low spatial resolution of the LiDAR data and in order to solve the overestimation of the external boundaries of roofs by using this data source. As a result, an image mask is obtained, and the following stages are performed per roof (Fig. 7a).

3.2.2. Boundaries detection strategies

In order to define the external roofs edges from the orthoimage, a

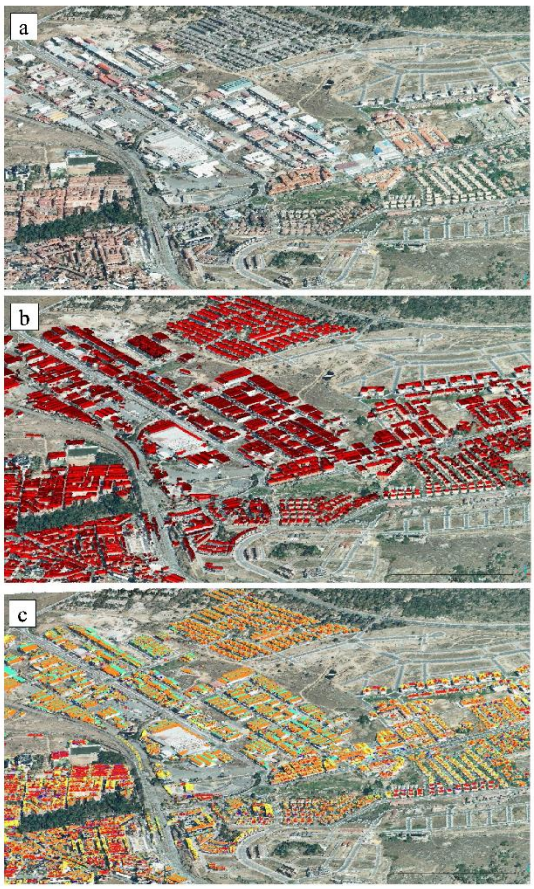


Fig. 8. Perspective view of part of the city of Ávila. (a) Raw LiDAR point cloud with true RGB values. (b) Incorporation of the roof layer in red. (c) Roof layer colourized according to the azimuth-orientation of roofs and with information about their tilt angles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

procedure consisting on the parallel application of an image segmentation and an identification of straight contours is applied (Fig. 7b). Image segmentation is performed using the Statistical Region Merging (SRM) algorithm [32], which is based on region growing and merging techniques. The identification of straight contours on the orthoimagery is performed with the Line Segment Detector (LSD) algorithm [33,34]. This algorithm detects lines based on the image gradient orientation of each pixel, so that lines are composed of adjacent pixels with the same image gradient orientation, which is different from the image gradient orientation of the surrounding pixels.

The image segmentation is applied to detect those pixels strictly corresponding to roof and non-roof areas. Pixels corresponding with roof areas are put together and the contour of the final "roof group" is determined.

The LSD is applied as a double validation of the pixels corresponding to the external roof edges. Since the LSD algorithm results in the extraction of lines from the exterior contour of the roof but also from the interior of the roof (i.e. lines that define the different roof slopes), the double validation by the SRM algorithm allows the strict definition of the external roof edges (Fig. 7c).

In the case that more than one line-segment are detected for one external roof edge, RANSAC algorithm [35] is applied to combine them in one line (Fig. 7d).

3.2.3. External edges detection

Lines from Stage 2 are intersected (Fig. 7e) with the aim at obtaining the real external contour of the roof (Fig. 7f). The contour obtained after the application of the SRM algorithm is used as reference, especially for the case of irregular roofs (for example, U-shaped). In the case the intersection point between lines is within the reference contour, it is considered as a real intersection point. If there are more than one intersection points in the same edge, the most external ones are considered as the real boundaries of the roof.

4. Results

This section deals with the application of the proposed methodology to several buildings and case studies in order to check its scalability and suitability to accurately characterize roofs at city, neighbourhood and building level of those countries with geospatial data within the INSPIRE standards. The case studies were a sample of buildings in the city of Ávila (Spain) and in the city of Vaihingen an der Enz (Germany). Specifically, the data of the city of Ávila corresponded to:

- Free aerial data offered by the Spanish National Geographic Agency at the IGN data repository [24] that was used geometrically

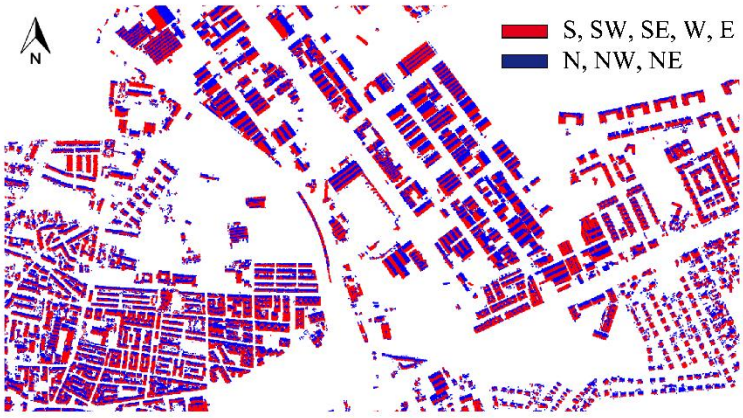


Fig. 9. Plan view of part of the roofs of the city of Ávila where the favourable orientations for the installation of PV systems have been coloured in red (S, SW, SE, W and E) and those unfavourable in blue (N, NW and NE). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

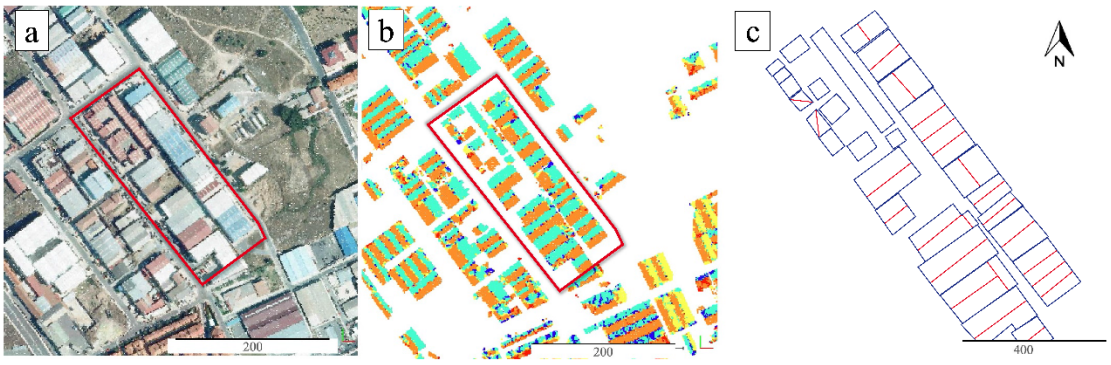
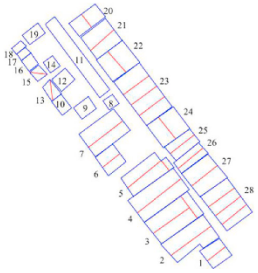


Fig. 10. Roof characterization of the chosen neighbourhood (red polygon) in the industrial estate of Ávila. Plan view of: (a) the raw LiDAR point cloud, (b) processed point cloud corresponding to roofs coloured according to the azimuth-orientation and with tilt angle information, and (c) the final vector result with information about the azimuth-orientation, tilt angle and area per roof slope. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Roof characterization parameters obtained for the 28 buildings analysed at neighbourhood scale in the city of Ávila.



characterize roofs at city, neighbourhood and building level.
- A higher resolution aerial data commissioned by the City Council of Ávila that was used to compare both results.

Regarding the data of the city of Vaihingen an der Enz it corresponded to:

- Aerial data offered by the German Society of Photogrammetry, Remote Sensing and Geoinformation (DGPF).
- Reference data available within the framework of an ISPRS test project [36] that served as ground truth to validate the proposed methodology for the automatic characterization of roofs.

4.1. Roof characterization at city level

For this purpose, six LiDAR data sheets of the city of Ávila from the IGN data repository were used covering 24 km^2. Fig. 8 shows the results obtained after applying the proposed methodology to this aerial LiDAR data (Table 2).

It can be concluded that roofs of the city of Ávila have a predominant Northeast orientation with 18.75%. Other orientations are distributed as follows: 13.07% of roof slopes oriented to the North, 13.04% to the Southeast, 12.75% to the Southwest, 10.86% to the West, 10.61% to the South, 10.60% to the East and 10.32% to the Northwest (Fig. 8c). On the other hand, it can be figured that the average tilt angle of the roof slopes of the city is 21.61°.

Facing the installation of solar panels on roofs, arguably 57.86% of the roofs of Ávila have a favourable orientation for the direct installation of integrated PV panels, without changing their orientation on the roof (Fig. 9). This percentage corresponds to the roof slopes oriented to the South, Southwest, Southeast, West and East.

4.2. Roof characterization at neighbourhood level

To test the methodology at neighbourhood level, a set of 28 buildings from the industrial estate of Ávila (Fig. 10), located at the Northeast of the city, was selected. This area was chosen because most of the roofs of industrial buildings have any of the typologies (Fig. 1) for which the automatic characterization has been resolved through the proposed methodology.

Then, in Table 3, all the parameters obtained after applying the methodology proposed in the chosen study area are summarized.

Under the assumption of a perfect symmetry of the roof slopes of the buildings 1, 4, 6, 7, 13, 15, 20, 21, 22, 23, 24, 25, 26, 27 and 28, which were identified as symmetrical through the analysis of the orthoimages, a RMSE of ± 5.25% is derived from the area estimation per roof slope when using low resolution LiDAR data (0.5 p/m^2, Table 2). For the case study and given that the average area per roof slope for those buildings supposed symmetrical was 313.63 m^2, the area of each roof slope has been overestimated or underestimated by an average of 16.47 m^2.

4.3. Roof characterization at building level

For this analysis 4 buildings of the city of Ávila were chosen based on their roof typology: a single pitched roof, a flat roof, a slope gable roof and a pyramid roof were analysed (Fig. 11). Results obtained after applying the proposed methodology are summarized in Table 4.

5. Validation and discussion

In order to validate the methodology, two analysis have been performed in different cities and datasets features. For the city of Ávila, the validation consisted in comparing results obtained for the 28 industrial buildings of Fig. 10 when processing the free available aerial data offered by the Spanish National Geographic Agency and when processing a 4 times higher spatial resolution aerial data. The latest corresponds to a 2 points per square meter LiDAR point cloud from a flight commissioned by the City Council of Ávila in 2010. For the city of Vaihingen an der Enz, the validation involved the comparison of the results obtained after processing orthoimagery and LiDAR data provided by the DGPF with the reference values offered by the ISPRS test project [36]. In this case, the residential area called "Area 3" was chosen and the spatial

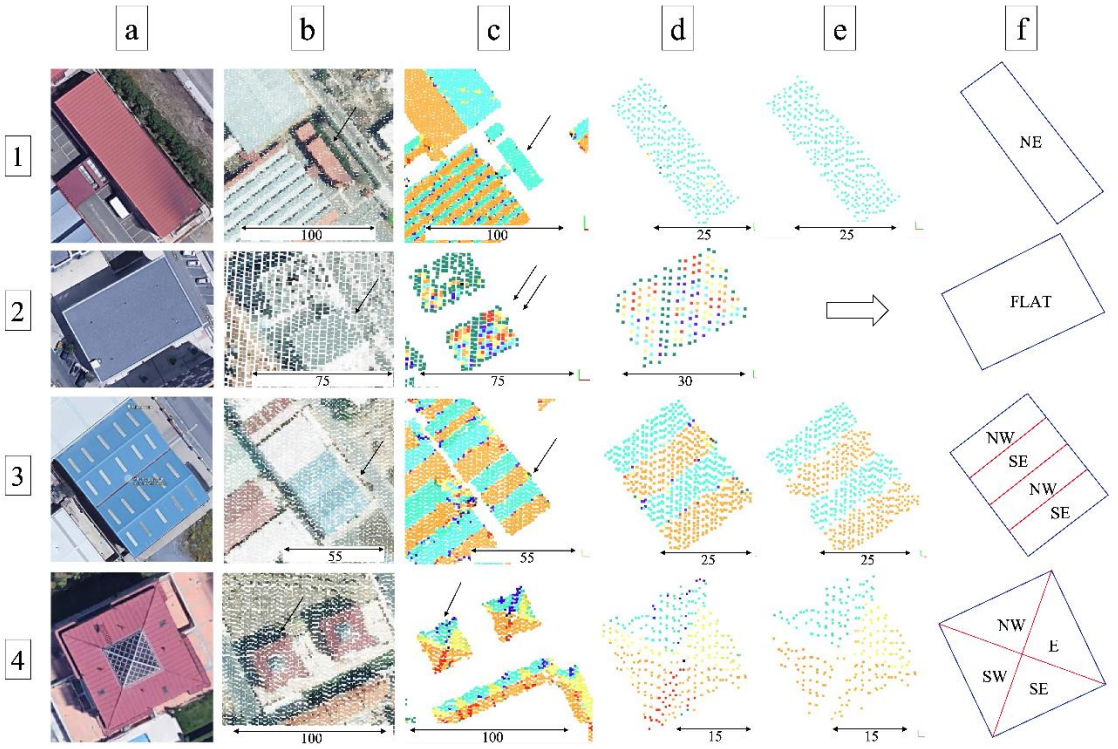


Fig. 11. Roof characterization of the four chosen buildings of the city of Ávila. (a) Plan view of: a (1) single pitched roof, a (2) flat roof, a (3) 4-slope gable roof and a (4) pyramid roof; source: Google Maps ©. Being (b) the raw LiDAR point cloud, (c) the processed point cloud corresponding to roofs coloured according to the azimuth-orientation and information regarding tilt angles after applying the CL8 step of the methodology, (d) the individual roof of interest, (e) points belonging to each roof slope detected after applying the step RL5, note that in case of flat roofs, the processing finishes in "d" when they are identified as flat roofs (see Section 3.1.2); and (f) the final vector result obtained after applying the step RL6 of the methodology that contains information regarding the azimuth-orientation, tilt angle and area per roof slope.

Table 4
Roof characterization parameters obtained for the 4 buildings analysed at building scale.

Building	Roof slope	Azimuth (0°–360°)	Orientation	Tilt angle (°)	Area (m^2)	Altitude (m)
1	1	52.96°	NE	10.13°	483.89	1141.50
2	1	FLAT	FLAT	FLAT	769.07	1151.80
3	1	322.28°	NW	10.92°	309.17	1149.60
	2	143.57°	SE	10.16°	323.96	1149.60
	3	323.19°	NW	10.24°	263.95	1149.60
	4	142.44°	SE	10.70°	224.31	1149.60
4	1	334.34°	NW	18.34°	136.71	1177.20
	2	65.65°	E	18.14°	129.77	1177.10
	3	153.65°	SE	17.72°	134.97	1177.00
	4	243.98°	SW	17.87°	135.99	1177.10

resolution of data was 8 cm for the orthoimages and 4 points per square meter for the LiDAR data.

5.1. Validation for the data set of Ávila (Spain)

The azimuth-orientation, tilt angle, area and altitude of the 28 buildings analysed in Section 3.2 are here updated (Table 5) with results obtained after applying the methodology to the higher spatial resolution LiDAR data of the same area (Fig. 12).

As in Section 3.2, under the assumption of perfect symmetry between the roof slopes of buildings 1, 4, 6, 7, 13, 15, 20, 21, 22, 23, 24, 25, 26, 27 and 28, a RMSE of 2.63% is derived from the area estimation per roof slope for this case study when using a 2 p/m^2 point cloud LiDAR data, and given that the average area per roof slope for those groups of symmetrical buildings was 313.63 m^2.

By comparing results obtained from both LiDAR data sets (Table 6) the following is concluded:

- Angular parameters (azimuth-orientation and tilt angle) presented an average error below 1°: 0.19% and 0.45% respectively considering 360° and 90° for azimuth-orientation and tilt angle, which are below the adjustment accuracies offered by solar panel supports [37].
- The altitude estimation highlighted for its low error, 0.02 m in this case study. This is due the precision offered by LiDAR technology in altimetry measurements (around 20 cm, Table 2).
- Area per roof slope obtained an average error of 14.51 m^2 which corresponds to an error in percentage of 4.63% taking as reference the average area of the roof slopes analysed (313.63 m^2). It would suppose an underestimation or overestimation of 7 panels (of 2 m^2) in roofs with capacity for 156 panels. For residential use, where 10 panels are usually installed per roof, this would lead to underestimating or overestimating the installation in 1 or zero panels (0.45 panels of 2 m^2).

Table 5
Roof characterization parameters obtained for the 28 buildings analysed at neighbourhood scale when using the 2 p/m^2 LiDAR data.

Building	Roof slope	Azimuth (0°–360°)	Orientation	Tilt angle (°)	Area (m^2)	Altitude (m)
1	1	142.36°	SE	10.88°	203.67	1143.60
	2	322.44°	NW	11.14°	203.12	1143.60
2	1	143.77°	SE	6.20°	515.02	1148.00
	2	322.33°	NW	6.39°	511.56	1147.80
3	1	142.04°	SE	15.31°	493.70	1147.40
	2	321.84°	NW	15.42°	497.78	1147.20
	3	FLAT	FLAT	3.14°	250.85	1142.60
4	1	144.08°	SE	5.67°	620.65	1145.70
	2	322.24°	NW	5.65°	610.34	1145.70
5	1	144.26°	SE	4.74°	420.00	1147.40
	2	323.24°	NW	4.68°	408.04	1147.40
	3	325.44°	NW	3.46°	42.21	1143.50
6	1	143.38°	SE	16.43°	197.30	1148.00
	2	323.62°	NW	15.77°	216.01	1147.60
7	1	143.00°	SE	5.55°	494.37	1149.30
	2	322.76°	NW	5.52°	494.41	1149.20
8	1	55.77°	NE	11.29°	106.12	1150.30
9	1	143.19°	SE	8.86°	217.17	1150.40
10	1	227.49°	SW	10.48°	167.46	1150.50
11	1	52.17°	NE	11.11°	831.74	1150.50
12	1	320.98°	NW	9.16°	203.65	1150.30
13	1	270.95°	W	7.78°	67.53	1152.10
	2	92.32°	E	5.76°	69.46	1152.10
14		143.22°	SE	10.24°	118.18	1150.30
15	1	8.27°	N	9.39°	56.84	1152.40
	2	187.91°	S	8.39°	60.51	1152.10
16	1	233.62°	SW	10.57°	128.25	1150.30
17	1	233.17°	SW	11.44°	78.60	1150.30
18	1	322.83°	NW	9.84°	68.74	1147.60
19	1	321.49°	NW	11.19°	217.87	1150.30
20	1	231.77°	SW	5.85°	221.26	1153.90
	2	53.27°	NE	5.85°	220.04	1153.80
21	1	322.91°	NW	5.61°	373.08	1155.70
	2	143.44°	SE	5.60°	330.11	1153.70
22	1	232.97°	SW	5.66°	365.70	1155.70
	2	52.70°	NE	5.65°	369.01	1155.70
23	1	323.32°	NW	5.48°	378.1	1153.30
	2	143.08°	SE	5.45°	296.6	1153.10
	3	323.36°	NW	5.54°	356.82	1153.20
	4	142.88°	SE	5.53°	299.56	1153.30
24	1	232.64°	SW	5.60°	375.49	1153.80
	2	53.24°	NE	5.63°	367.25	1153.70
25	1	324.11°	NW	12.46°	205.86	1151.80
	2	143.91°	SE	10.81°	207.69	1151.70
26	1	323.81°	NW	11.98°	221.96	1152.10
	2	142.95°	SE	8.48°	224.61	1152.40
27	1	323.15°	NW	3.34°	399.59	1153.60
	2	143.39°	SE	3.42°	405.06	1153.50
28	1	322.51°	NW	10.21°	292.48	1149.60
	2	143.43°	SE	10.35°	274.65	1149.60
	3	322.56°	NW	10.35°	288.46	1149.60
	4	142.91°	SE	10.77°	265.80	1149.60

- Assuming perfect symmetry between the roof slopes of the buildings 1,4, 6, 7, 13, 15, 20, 21, 22, 23, 24, 25, 26, 27 and 28, the improvement on the results of area calculation using the higher spatial resolution LiDAR data validates the methodology, since the error is dependent on the data source.

5.2. Validation for the data set of Vaihingen an der Enz (Germany)

Finally, in order to validate the performance of the proposed methodology with ground-truth reference data, a sample of buildings from a residential neighbourhood of the city of Vaihingen an der Enz was analysed. Specifically, the neighbourhood called "Area 3" in the ISPRS test project [36] that has 47 buildings and for which reference data is available. "Area 3" was chosen from among the three areas offered because most of the roofs (42 of the 47) have a morphology that can be characterized by the proposed methodology. However, the analysis was carried out in 41 buildings since there was a scarce LiDAR data recorded for the building located to the right of the building 2 in Fig. 13. This is due to the fact that slate roofs absorb the infrared radiation [38], spectral range for which the LiDAR used by the DGPF operates (Leica ALS50).

After segmenting the point cloud by eliminating the 6 buildings that were not analysed for the aforementioned reasons, the methodology was applied for the point cloud with the 41 remaining buildings. The geometrical parameters obtained for each roof slope are summarized in Table 7.

Since the average altitude of the ground of such neighbourhood (Z coordinate of the LiDAR data) was 266.50 m and that of the roofs analysed was 273.42, it was concluded that it is a low-rise residential neighbourhood with buildings of an average height of 6.92 m ± 3.55 m.

In order to check the results obtained and validate the methodology, each geometric parameter was compared with its corresponding reference value offered by the ISPRS test project [36]. Table 8 shows the comparative analysis after which the following conclusions were derived:

- Angular parameters (azimuth-orientation and tilt angle) presented an average error around 1°, matching the results obtained in the test site of the city of Ávila. These uncertainty values are below the adjustment accuracies offered by solar panel supports [37].
- An average error of 6.97 m^2 was obtained when evaluating the area per roof slope. It corresponds to an error in percentage of 2.93% taking as reference an area of 238.29 m^2. This error supposes an underestimation or overestimation of 3 PV panels (of 2 m^2). Better results have been obtained compared to those of the city of Ávila due to the greater spatial resolution offered by the aerial data provided by the DGPF.
- The altitude estimation had a 0.10 m average error, a slightly higher error than the one obtained in the test site of the city of Ávila. However, in both cases, errors are insignificant for the solar radiation computation.

6. Conclusions

The proposed methodology effectively solves the automatic parameterization of roofs at city, neighbourhood and building level from the combined use of discrete aerial LiDAR data and aerial orthoimagery and for the vast majority of existing roof types: single pitched roofs, flat roofs, gable or saddle roofs and pyramid roofs. In addition, the possibility of processing at city level allows the performance of prospective studies, towards the analysis of possibilities and establishment of regulations and incentives by the urban and regional Administrations.

The methodology involves not only the automatic extraction of roof boundaries but also the parametrization of each roof slope in terms of azimuth-orientation, tilt angle, altitude and effective dimensions, which are the parameters required for the computation of the PV capacity of roofs. However, it should be highlighted that the methodology is sensitive to changes in the resolution of the input data, in such a way that important improvements in accuracy are obtained with higher resolution data sets. Specifically, significant improvements have been obtained regarding the delimitation of the roof slopes and the computation of their dimensions when comparing results from higher spatial resolution orthoimagery and LiDAR data. This is because a higher density point cloud provides a better definition of the elements and therefore, in the case of the building roofs, allows a better differentiation of each roof slope and a better fit of planes. As for the rest of the parameters, very precise results are obtained in the angular measurements for both higher and lower spatial resolution aerial data. Specifically, an average error of 0.20% was obtained in the azimuth-orientation estimation, 1.29% in the tilt angle estimation, 0.63% in the

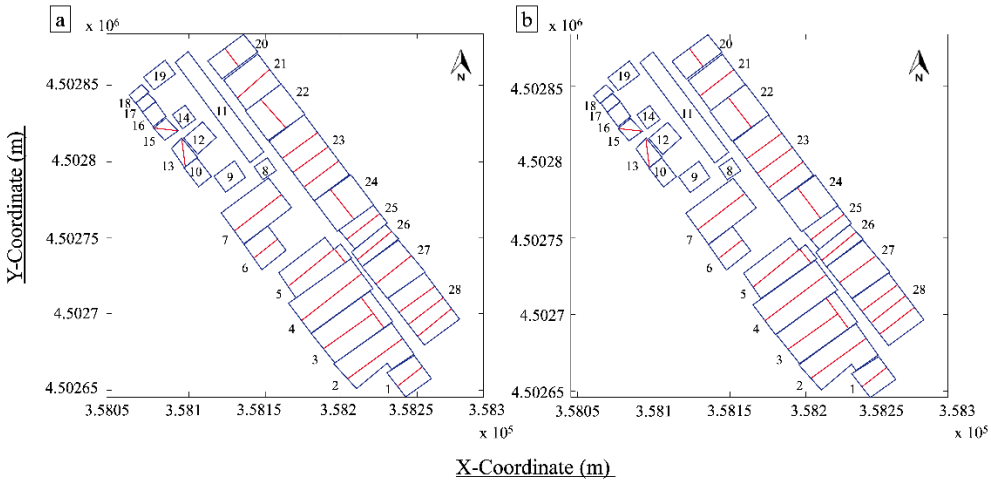


Fig. 12. Roof characterization parameters obtained for the 28 buildings analysed at neighbourhood scale when using (a) the LiDAR data of 0.5 p/m^2 resolution and when using (b) the LiDAR data of 2 p/m^2 resolution. The abscissa and ordinate represent the X and Y coordinates respectively in EPSG 25830 projection (ETRS89/UTM zone 30N).

Table 6
Uncertainty estimation taking as real values those obtained after processing the higher spatial resolution LiDAR data set (2 p/m^2). Values in bold indicate the relative average error in percentage.

	Azimuth (°)	Tilt angle (°)	Area per roof slope (m^2)	Altitude (m)
Min error	0.01°	0.01°	0.45	0.00
Max error	4.27°	1.08°	76.83	0.10
Average error	0.70°	0.40°	14.51	0.02
Average error (%)	**0.19%**	**0.45%**	**4.63%**	**0.15%**
(Reference)	(360°)	(90°)	(313.63 m^2)	(15 m)

altitude estimation and 2.93% in the area per roof slope estimation.

Future works will present a three-fold approach: on the one hand, focusing on the improvement of the characterization of the roofs, the authors will deal with the detailed study and analysis of hipped roofs to include them in the methodology in order to make it applicable to all typologies of existing roofs. In addition, procedures to automatically detect the projected shadows on roofs will be designed based on the altitude value and making a study of the sight line of each building. Finally, the analysis of the PV potential will be performed by integrating the study of the effect of the geomorphology of the soil on the incident solar radiation to PV surfaces.

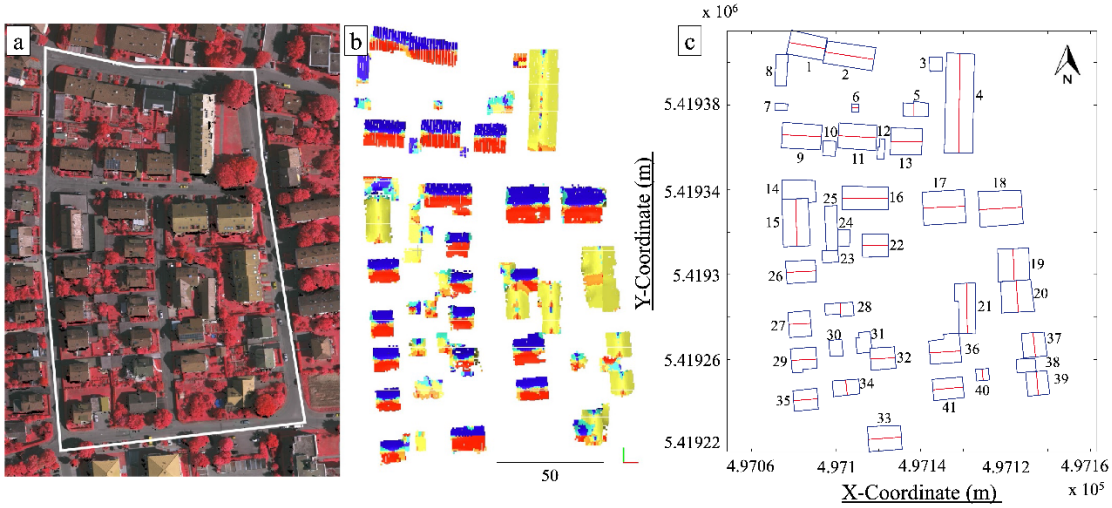


Fig. 13. Roof characterization of the chosen neighbourhood in the city of Vaihingen an der Enz. (a) Orthoimage of the neighbourhood in false colour, (b) processed point cloud corresponding to roofs coloured according to the azimuth-orientation and with tilt angle information, and (c) the final result with information about the azimuth-orientation, tilt angle and area per roof slope. Note that in (c) the abscissa and ordinate represent the X and Y coordinates respectively in EPSG 25832 projection (ETRS89/UTM zone 32N).

Table 7
Roof characterization parameters obtained for the 41 buildings analysed at neighbourhood scale when using 4 p/m^2 LiDAR data and an orthoimage resolution of 8 cm.

Building	Roof slope	Azimuth (0°–360°)	Orientation	Tilt angle (°)	Area (m^2)	Altitude (m)
1	1	10.99°	N	23.36°	100.74	270.00
	2	191.39°	S	21.29°	102.86	270.25
2	1	9.28°	N	24.60°	138.13	270.49
	2	189.05°	S	24.82°	133.98	270.55
3	1	FLAT	FLAT	1.18°	39.20	263.60
4	1	91.29°	E	26.74°	346.68	275.97
	2	271.10°	W	27.01°	352.07	275.91
5	1	FLAT	FLAT	2.01°	43.08	266.77
	2	FLAT	FLAT	1.95°	31.48	266.09
6	1	7.75°	N	7.70°	5.88	267.40
	2	188.75°	S	12.33°	6.01	267.44
7	1	FLAT	FLAT	2.96°	19.73	266.79
8	1	0.84°	N	4.55°	81.01	265.59
9	1	3.57°	N	35.47°	133.30	273.85
	2	183.59°	S	35.31°	133.85	273.75
10	1	2.33°	N	2.56°	43.22	267.99
11	1	3.49°	N	28.16°	122.61	272.89
	2	182.50°	S	27.88°	121.11	272.95
12	1	2.74°	N	2.71°	27.43	267.71
13	1	1.93°	N	31.14°	104.28	273.73
	2	180.37°	S	31.07°	109.02	273.51
14	1	FLAT	FLAT	0.92°	144.68	270.58
15	1	268.08°	W	41.45°	161.49	276.16
	2	88.15°	E	41.84°	172.91	276.07
16	1	2.32°	N	34.26°	142.21	273.50
	2	180.72°	S	33.32°	139.50	273.61
17	1	356.52°	N	32.20°	184.63	276.64
	2	176.83°	S	32.19°	182.52	276.61
18	1	356.85°	N	31.75°	186.69	276.26
	2	177.18°	S	32.30°	195.96	276.41
19	1	267.28°	W	33.16°	134.60	277.10
	2	87.12°	E	33.08°	130.76	276.84
20	1	266.67°	W	33.44°	127.24	275.36
	2	87.06°	E	33.21°	133.30	275.51
21	1	268.94°	W	30.05°	115.04	273.02
	2	88.55°	E	31.79°	124.98	273.29
22	1	355.13°	N	35.94°	77.29	273.45
	2	178.80°	S	37.07°	80.71	273.48
23	1	267.11°	W	5.14°	45.38	270.76
24	1	272.08°	W	3.87°	44.13	269.47
25	1	87.56°	E	6.04°	118.84	269.56
26	1	353.01°	N	30.51°	84.75	275.72
	2	178.20°	S	31.57°	85.69	275.53
27	1	356.04°	N	29.57°	71.07	277.49
	2	176.46°	S	28.85°	70.05	277.50
28	1	32.71°	NE	7.05°	38.01	271.59
	2	199.52°	SW	6.09°	41.31	271.48
29	1	355.44°	N	28.32°	24.26	278.62
	2	174.70°	S	29.71°	43.38	278.83
30	1	104.88°	E	23.58°	51.92	272.71
31	1	281.96°	W	5.86°	64.29	271.71
32	1	355.38°	N	45.58°	83.76	275.85
	2	175.56°	S	44.70°	74.33	276.02
33	1	355.45°	N	27.75°	103.88	275.95
	2	175.20°	S	22.63°	95.75	276.19
34	1	87.49°	E	2.94°	48.82	273.27
	2	264.89°	W	4.38°	39.39	273.21
35	1	355.22°	N	30.36°	64.29	278.70
	2	175.26°	S	29.67°	62.85	278.66
36	1	354.72°	N	34.89°	95.35	274.79
	2	175.17°	S	36.17°	72.48	275.31
37	1	265.21°	W	33.39°	75.75	275.49
	2	84.91°	E	34.76°	77.14	275.53
38	1	FLAT	FLAT	1.79°	58.15	271.42
39	1	264.66°	W	32.83°	68.12	275.74
	2	85.60°	E	33.29°	69.01	275.75
40	1	266.37°	W	21.02°	16.96	272.54
	2	86.96°	E	23.15°	16.00	272.55
41	1	354.06°	N	26.07°	80.44	277.19
	2	174.74°	S	26.57°	76.84	277.21

Table 8
Uncertainty analysis when results obtained after processing the aerial data with the proposed methodology are compared with the ground truth data offered by the ISPRS Test Project. Values in bold indicate the relative average error in percentage.

	Azimuth (°)	Tilt angle (°)	Area per roof slope (m^2)	Altitude (m)
Min error	0.02°	0.09°	0.01	0.01
Max error	4.08°	3.62°	40.59	0.17
Average error	0.71°	1.16°	6.97	0.10
Average error (%)	**0.20%**	**1.29%**	**2.93%**	**0.63%**
(Reference)	(360°)	(90°)	(238.29 m^2)	(15 m)

Declaration of competing interest

Authors declare that this work is original, the information reported in the paper is accurate according to the best knowledge of the authors, the paper has not been previously published or submitted elsewhere, and the authors are aware of ELSEVIER policy on plagiarism and self-plagiarism and there is not conflict of interest.

Acknowledgements

Authors would like to thank Iberdrola S.L. and University of Salamanca for the funding provided through Cátedra Iberdrola VIII Centenary. We also express our gratitude to the Ministry of Science, Innovation and Universities for the funding given for the project RTC-2017-6291-3.

Funding

This work was supported by the Ministry of Science, Innovation and Universities of Spain [project RTC-2017-6291-3]; and Cátedra Iberdrola-USAL VIII Centenary.

References

[1] Eurostat, Community Innovation Survey of Renewable Energy Statistics of the European Communities, Office for Official Publications of the European Communities, 2017 Available at https://ec.europa.eu/eurostat/statistics-explained (last access: 27/03/2019).
[2] Spanish Official State Bulletin, BOE, BOE-A-2018-13593 About Urgent Measures for Energy Transition and Consumer Protection, Available at https://boe.es/eli/es/rdl/2018/10/05/15 (Last access: 27/03/2019).
[3] S. Castellanos, D.A. Sunter, D.M. Kammen, Rooftop solar photovoltaic potential in cities: how scalable are assessment approaches? Environ. Res. Lett. 12 (12) (2017) 125005, , https://doi.org/10.1088/1748-9326/aa7857.
[4] M.C. Brito, S. Freitas, S. Guimarães, C. Catita, P. Redweik, The importance of facades for the solar PV potential of a Mediterranean City using LiDAR data, Renew. Energy 111 (2017) 85–94, https://doi.org/10.1016/j.renene.2017.03.085.
[5] M. Fionnuala, K. McDonnell, A feasibility assessment of photovoltaic power systems in Ireland; a case study for the Dublin Region, Sustainability 9 (2) (2017) 302, https://doi.org/10.3390/su9020302.
[6] Google's Project Sunroof, Available at https://google.com/get/sunroof (Last access: 27/03/2019).
[7] D. Palmer, E. Koubli, I. Cole, R. Gottschalg, T. Betts, A GIS-based method for identification of wide area rooftop suitability for minimum size PV systems using LiDAR data and photogrammetry, Energies 11 (12) (2018) 3506, https://doi.org/10.3390/en11123506.
[8] A.M. Martín, J. Domínguez, J. Amador, Applying LIDAR datasets and GIS based model to evaluate solar potential over roofs: a review, AIMS Energy 3 (3) (2015) 326–343, https://doi.org/10.3934/energy.2015.3.326.
[9] S. Schuffert, T. Voegtle, N. Tate, A. Ramírez, Quality assessment of roof planes extracted from height data for solar energy systems by the EAGLE platform, Remote Sens. 7 (12) (2015) 17016–17034, https://doi.org/10.3390/rs7121586.
[10] K. Chen, W. Lu, F. Xue, P. Tang, L.H. Li, Automatic building information model reconstruction in high-density urban areas: augmenting multi-source data with architectural knowledge, Autom. Constr. 93 (2018) 22–34, https://doi.org/10.1016/j.autcon.2018.05.009.
[11] D. Li, M. Lu, Integrating geometric models, site images and GIS based on Google Earth and keyhole markup language, Autom. Constr. 89 (2018) 317–331, https://doi.org/10.1016/j.autcon.2018.02.002.
[12] Y. Li, C. Liu, Estimating solar energy potentials on pitched roofs, Energy Build. 139 (2017) 101–107, https://doi.org/10.1016/j.enbuild.2016.12.070.

[13] J. Sarralde, D.J. Quinn, D. Wiesmann, K. Steemers, Solar energy and urban morphology: scenarios for increasing the renewable energy potential of neighbourhoods in London, Renew. Energy 73 (2015) 10–17, https://doi.org/10.1016/j.renene.2014.06.028.
[14] S. Kiti, V. Wang, B. Sharp, Rooftop solar potential based on LiDAR data: bottom-up assessment at neighbourhood level, Renew. Energy 111 (2017) 463–475, https://doi.org/10.1016/j.renene.2017.04.025.
[15] T. Schenk, B. Csathó, Fusion of LIDAR data and aerial imagery for a more complete surface description, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 34 (3/A) (2002) 310–317, https://doi.org/10.1109/DFUA.2003.1219962.
[16] H. Arefi, P. Reinartz, Building reconstruction using DSM and orthorectified images, Remote Sens. 5 (4) (2013) 1681–1703, https://doi.org/10.3390/rs5041681.
[17] H. Fan, W. Yao, Q. Fu, Segmentation of sloped roofs from airborne LiDAR point clouds using ridge-based hierarchical decomposition, Remote Sens. 6 (4) (2014) 3284–3301, https://doi.org/10.3390/rs6043284.
[18] S. Ghaffarian, S. Ghaffarian, Automatic building detection based on Purposive FastICA (PFICA) algorithm using monocular high-resolution Google Earth images, ISPRS J. Photogramm. Remote Sens. 97 (2014) 152–159, https://doi.org/10.1016/j.isprsjprs.2014.08.017.
[19] H. Wang, W. Zhang, Y. Chen, M. Chen, K. Yan, Semantic decomposition and reconstruction of compound buildings with symmetric roofs from LiDAR data and aerial imagery, Remote Sens. 7 (10) (2015) 13945–13974, https://doi.org/10.3390/rs71013945.
[20] S. Ghaffarian, S. Ghaffarian, Z. Samir, Y. Ruicheck, Automatic building roof segmentation based on PFICA algorithm and morphological filtering from LiDAR point clouds, 37th Asian Conference on Remote Sensing, ACRS 2016: Spatial Data Infrastructure for Sustainable Development, 2016 https://ris.utwente.nl/ws/portalfiles/portal/42925411/Ghaffarian2016automatic.pdf (Last access: 21/10/2019).
[21] R. Zhao, M. Pang, M. Wei, Accurate extraction of building roofs from airborne light detection and ranging point clouds using a coarse-to-fine approach, J. Appl. Remote. Sens. 12 (2) (2018) 026011, , https://doi.org/10.1117/1.JRS.12.026011.
[22] F. Pirotti, C. Zanchetta, M. Previtali, S. Della Torre, Detection of building roofs and facades from aerial laser scanning data using deep learning, ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4211 (2019) 975–980, https://doi.org/10.5194/isprs-archives-XLII-2-W11-975-2019.
[23] INSPIRE Geoportal, http://inspire-geoportal.ec.europa.eu/ (Last access: 27/03/2019).
[24] Downloading Centre – Spanish Centre of Geographic Information, http://centrodedescargas.cnig.es/CentroDescargas/index.jsp (Last access: 27/03/2019).
[25] A. Sampath, J. Shan, Building boundary tracing and regularization from airborne LiDAR point clouds, Photogramm. Eng. Remote Sens. (7) (2007) 805–812, https://doi.org/10.14358/PERS.73.7.805.
[26] G.M. Gandhi, S. Parthiban, N. Thummalu, A. Christy, NDVI: vegetation change detection using remote sensing and GIS – a case study of Vellore District, Proc. Comput. Sci. 57 (2015) 1199–1210, https://doi.org/10.1016/j.procs.2015.07.415.
[27] A.K. Bhandari, A. Kumar, G.K. Singh, Feature extraction using Normalized Difference Vegetation Index (NDVI): a case study of Jabalpur City, Proc. Technol. 6 (2012) 612–621, https://doi.org/10.1016/j.protcy.2012.10.074.
[28] K. Jordan, P. Mordohai, A quantitative evaluation of surface normal estimation in point clouds, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2014, , https://doi.org/10.1109/IROS.2014.6943157.
[29] Point Cloud Library (PCL) Documentation, http://pointclouds.org/documentation/tutorials/statistical_outlier.php#statistical-outlier-removal (Last access: 27/03/2019).
[30] The Spanish Technical Building Code (Royal Decree 314/2006 of 17 March 2006), Available online https://www.codigotecnico.org/ (Last access: 27/03/2019).
[31] A. Nurunnabi, G. West, D. Belton, Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point cloud data, Pattern Recogn. 48 (4) (2015) 1404–1419, https://doi.org/10.1016/j.patcog.2014.10.014.
[32] R. Nock, F. Nielsen, Statistical region merging, IEEE Trans. Pattern Anal. Mach. Intell. 26 (11) (2004) 1452–1458, https://doi.org/10.1109/TPAMI.2004.110.
[33] R.G. Von Gioi, J. Jakubowicz, J.M. Morel, G. Randall, LSD: a line segment detector, Image Process. On Line 2 (2012) 35–55, https://doi.org/10.5201/ipol.2012.gjmr-lsd.
[34] R.G. Von Gioi, J. Jakubowicz, J. Morel, G. Randall, LSD: a fast line segment detector with a false detection control, IEEE Trans. Pattern Anal. Mach. Intell. 32 (4) (2010) 722–732, https://doi.org/10.1109/TPAMI.2008.300.
[35] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395, https://doi.org/10.1145/358669.358692.
[36] F. Rottensteiner, G. Sohn, M. Gerke, J.D. Wegner, ISPRS test project on urban classification and 3D building reconstruction, Commission III-Photogrammetric Computer Vision and Image Analysis, Working Group III/4-3D Scene Analysis, 2013, pp. 1–17 http://www.cvlibs.net/projects/autonomous_vision_survey/literature/Rottensteiner2013.pdf (Last access: 21/10/2019).
[37] S. Roberts, N. Guariento, Building Integrated Photovoltaics: A Handbook, Birkhäuser, Basel, 978-3764399481, 2009, https://doi.org/10.1007/978-3-0346-0486-4.
[38] G. Gaussorgues, S. Chomet, Infrared Thermography, vol. 5, Springer Science & Business Media, 978-0470682234, 2012, https://doi.org/10.1002/9780470682234.

Capítulo V

Automatización en el cálculo de instalaciones fotovoltaicas

5. Automatización en el cálculo de instalaciones fotovoltaicas

En este capítulo se incluye el artículo *"A geospatial web-based platform for the photovoltaic potential computation: Ener3DMap-SolarWeb Roofs"* que fue publicado en la revista de alto impacto ***Renewable and Sustainable Energy Reviews***, en Agosto de 2020.

5.1. Resumen

Para explotar y gestionar de un modo eficaz las energías renovables es necesario conocer la intensidad energética en la ubicación de la instalación y, además, la influencia sobre ella de la geometría y del entorno. Por esta razón, el procesamiento e interpretación eficientes de datos geoespaciales y energéticos combinados es muy importante.

Este artículo presenta el desarrollo de una herramienta web para el cálculo automático del potencial fotovoltaico en tejados y parcelas sin edificación. La aplicación denominada ***Ener3DMap-SolarWeb Roofs*** (apéndice B), utiliza la librería para mapas Leaflet y soporta formatos WMS, GeoJSON, GeoCSV y KML, entre otros. Estos formatos de datos, mapas base, datos geométricos de los tejados calculados automáticamente con algoritmos de procesamiento propios, según la metodología descrita en la segunda publicación de esta Tesis Doctoral, datos catastrales y un modelo de radiación solar se integran en la herramienta.

Las principales aportaciones de esta aplicación frente a otras soluciones existentes son los diferentes tipos de datos soportados, el alto nivel de automatización y las distintas escalas para las que se calculan los datos de producción solar (horario, mensual y anual). Las capacidades de la herramienta se prueban a través de su aplicación para analizar el potencial solar de tejados con distintas formas geométricas y para diferentes configuraciones de paneles solares. También permite realizar el cálculo para instalaciones proyectadas en parcelas sin edificios, o en edificios proyectados.

La precisión de los resultados se garantiza mediante la integración de una metodología validada para el cálculo de la geometría y de un modelo de radiación solar validado, PVGIS.

Adicionalmente se desarrolla otra infraestructura de datos espaciales que, con la misma base, se centra en estudios de prospectiva solar a nivel de ciudades o barrios, permitiendo analizar el potencial para cada orientación de los tejados existentes en el área seleccionada. La herramienta es ***Ener3DMap-SolarWeb Cities*** (apéndice B).

Palabras clave: infraestructura de datos espaciales; potencial de energía solar; tejados; mapas web; sistemas de gestión de energía; administración.

5.2. Publicación 3: A geospatial web-based platform for the photovoltaic potential computation: Ener3DMap-SolarWeb Roofs

Renewable and Sustainable Energy Reviews 135 (2021) 110203

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: http://www.elsevier.com/locate/rser

Ener3DMap-SolarWeb roofs: A geospatial web-based platform to compute photovoltaic potential

M. Sánchez-Aparicio, J. Martín-Jiménez, S. Del Pozo, E. González-González, S. Lagüela *

Department of Cartographic and Land Engineering, University of Salamanca, Spain

ARTICLE INFO

Keywords:
Spatial data infrastructure
Solar energy potential
Rooftops
Web-mapping
Energy management system
Management

ABSTRACT

The effective exploitation and management of renewable energies requires knowledge not only of the energy intensity at the exploitation site but also of the influence of the geometry of the site and its surroundings. For this reason, the efficient processing and interpretation of combined geospatial and energy data is a key issue. This paper presents the development of a web-based tool for the automatic computation of photovoltaic potential on rooftops and on parcels without buildings. The tool called Ener3DMap-SolarWeb Roofs is based on Leaflet and supports WMS, GeoJSON, GeoCSV and KML formats, among others. With these data formats, base maps, geometric data from the rooftops automatically computed from LiDAR and imagery data with self-developed processing algorithms, cadastral data and a solar radiation model are integrated in the tool. These different types of data, the high level of automation and the different scales for which energy data is calculated (hourly, monthly and annually) are the main contributions of the presented tool compared to other existing solutions. The capacities of the tool are tested through its application to analyze the solar potential of rooftops with different shapes and for different solar panel configurations. The accuracy of the results is ensured through the integration of a validated methodology for the computation of geometry and a validated solar radiation model, PVGIS.

1. Introduction

A commitment to change the global energy model is a necessary reality in today's world. In fact, the transition from fossil fuel-based systems to renewable alternatives based on sustainable practices has already been initiated in promising ways [1]. To meet global climate targets for reducing greenhouse gas emissions, energy efficiency and sustainability, one of the keys is to avoid energy losses by producing energy close to its consumption [2], i.e., in homes and industries. In this sense, photovoltaic (PV) energy is one of the most versatile and appropriate renewable energies since (i) it can be installed in urban environments directly on building rooftops and in dedicated plots in urban and rural environments [3,4] and (ii) the orientation of the solar panels can be modified to obtain higher energy peaks (South orientation) or more constant production (East-West orientation) [5]. Due to the modular nature of PV panels, they can be directly integrated on buildings, allowing individual or shared self-consumption [6], contributing to electricity cost savings. Even though PV technology has become cheaper in recent years, it has still not expanded as would be expected. One of the main reasons is that citizens and entrepreneurs are unaware of the potential and capability to generate energy from the sun on their buildings or the amortization schedule of such technology [7].

The other reason for the slow incorporation of PV technology in the energy mix is that efficient installation and management of solar energy resources requires a knowledge of the energy potential in different time periods, of the technology and of the economic indices [8]. There are different data sources that facilitate an understanding of the incoming solar radiation on a surface [9], but the design of PV installations requires knowing the solar potential at the installation position, improving the scale [10].

The need for geo-referencing all the data involved makes geographical databases the optimal support for the energy analysis, in a similar approach as that used for land use change and fire assessment [11,12]. Solar analysis has also been performed using geographical databases, with different scales and levels of computation, as reviewed in Ref. [13].

This study develops a web-based tool for the computation of the solar energy potential on building rooftops and on parcels without buildings. The strengths of the tool lie in its high automation level, as the roof geometry is automatically provided without user data input, its online nature, removing the need to install software, and its ease of use. The

* Corresponding author.
E-mail address: sulaguela@usal.es (S. Lagüela).

https://doi.org/10.1016/j.rser.2020.110203
Received 6 December 2019; Received in revised form 24 July 2020; Accepted 1 August 2020
Available online 13 August 2020
1364-0321/

List of abbreviations

2D	2 Dimensions
3D	3 Dimensions
API	Application Programming Interface
DLG	Digital Line Graphic
DOM	Digital Orthophoto Map
DSM	Digital Surface Model
GeoJSON	Geometry Javascript Object Notation
GeoCSV	Geometry Comma-Separated Values
GIS	Geographic Information System
JSON	Javascript Object Notation
KML:	Keyhole Markup Language
LiDAR	Light Detection And Ranging
PNOA	Plan National de Ortofotografía Aérea (Spanish Plan of Aerial Orthophotography)
PHP	Hypertext PreProcessor
PS	Peak Sun
PV	PhotoVoltaic
PVGIS	PhotoVoltaic Geographic Information Model
SDI	Spatial Data Infrastructure
WMS	Web Mapping Service

Units

kWh	kilowatt-hour
kWh/year	kilowatt-hour per year
kB	kilobyte
m^2	square meters
MB	Megabyte
s	seconds
W	Watts

automatic incorporation of geometric data from existing rooftops is achieved with the application of the methodology published in Ref. [14], in which LiDAR data and aerial orthophotographies are processed to compute tilt angle and orientation, useable area and building position (latitude and longitude).

To guarantee the replicability of the methodology developed, the input sources were selected such that they are available in most countries, as explained in Ref. [14]. In the case of geospatial and geometrical data, its applicability was ensured by the use of data sources following the INSPIRE directive [15].

The paper is organized as follows: Section 2 outlines all necessary layers and input data for the design of PV installations on rooftops, as well as the methodology followed to compute geometric roof data from LiDAR data, and the design and implementation of the web tool; Section 3 describes the operation of the web tool; Section 4 presents the validation tests, both experimental and theoretical. Finally, Section 5 summarizes the main conclusions.

1.1. State of the art among the scientific community

The need to understand the process of computing rooftop solar potential has led to the development of different solutions that combine geometry with solar irradiation data. Geographic Information Systems (GIS) have been widely used for the resolution of the renewable energy siting problem, as discussed in publications on offshore and inland wind energy [16,17], geothermal energy [18], hydraulic and marine energies [19,20] and bioenergy [21]. In the particular case of solar energy, several GIS tools have been developed by the scientific community to solve the siting problem for different applications, from solar farms [22,23] to the integration of solar panels in building envelopes [24–26].

Spatial Data Infrastructures (SDIs) are web-based GIS tools [27]. SDIs can store different types of geospatial data, which can be visualized and queried [28]. Presents the use of an SDI for an analysis of the solar potential in an urban environment. Most tools require user data input to perform a more complex study.

Rooftop data can be computed with different methods [29]. Presents a methodology for the computing solar energy in a city through artificial intelligence and image segmentation techniques [30]. Proposes the automatic detection of buildings from a photogrammetric digital surface model (DSM) and a digital orthophoto map (DOM) with the help of historical digital line graphic (DLG) data [31]. Presents the detection of residential buildings with a combination of neural networks and LiDAR data.

However, the solar energy computation requires not only the detection of each building but also the computation of geometric parameters: height, orientation, tilt angle and useable area. Since none of the methodologies mentioned for rooftop analysis found in literature include this additional step, the methodology presented in Refs. [14] is applied for the generation of data from rooftops automatically, as it is the only one that provides all the information required for the design of a PV system in rooftops, in an automatic way.

1.2. State of the art in solar radiation mapping and computation tools

Scientific developments in solar radiation mathematical modeling and building rooftop geometric modeling have enabled the development of tools (open access or under license) for user design of solar PV installations. However, the level of automation of the existing tools is generally low; most require manual user input of the rooftop geometric data.

One of the most advanced tools is Google SunRoof [32], which allows the user to search for a building in Google Maps and analyze the rooftop's solar potential. However, the tool is only for existing rooftops, works in 2D with satellite imagery and does not consider the variation in solar radiation with height. In addition, the geometric parameters of the roof must be manually input.

Other tools, such as Solmetric and PVSOL, provide numerical values for solar irradiation, but do not generate a map for its visual interpretation. There is also variability in the frequency in which existing tools compute solar data: although some tools compute solar energy hourly, others provide only monthly data. These last tools are consequently limited to use for PV installation design but cannot be used daily to assess the energy consumption and generation once the PV panels have been installed.

Table 1 presents more information about existing tools, but as a summary, existing tools for the computation of solar radiation have one or more of the following limitations: (i) manual input of rooftop geometry; (ii) the solar radiation is computed numerically, and no visual interpretation is provided; and (iii) the frequency of the datum of solar radiation is not enough for the management of the PV installation. Thus, Ener3DMap-Solar Web is generated with the aim at overcoming all the limitation mentioned.

2. Materials and methods

2.1. Materials

Ener3DMap-SolarWeb hosts all the data required for the rooftop PV potential computation characterized by its different nature and format. Indeed, it combines raster, vector and alphanumeric data with different roles, serving as: base-maps, descriptive information about rooftops and PV systems (geometry parameters, performance), key data for querying

Table 1
Characteristics of existing solar computation tools for comparison with ENER3DMAP-Solar Web Tool.

	LOCATION	SCALE	SOLAR POTENTIAL MAPPING	ROOF DATA PROVIDED	AUTOMATION LEVEL	MANUAL PV CUSTOMIZATION	RESULTS	TEMPORAL RESOLUTION OF SOLAR ENERGY PRODUCTION
ENER3DMAP-SOLAR WEB	Spain	Roof	Yes	Area, tilt, orientation, height	High	Yes	Per roof (built and unbuilt): longitude, latitude, orientation, area, tilt angle, panels, annual, monthly, daily or hourly PV production, equivalent hours of sun, climatic zone, cadastral data	Hour, Day, Month or Year
UVEK-GIS	Switzerland	Roof	Yes	No	High	No	Suitability level (low, medium, high and very high) for the use of solar energy + kWh estimation	Year
MAPDWELL	USA & Chile	Roof	No (installed capacity mapping based on the most favorable roof characteristics)	No	Medium	Yes	No. Of panels, PV installed capacity, investment plan, efficiency, yearly PV production and carbon footprint	Year
NYSOLARMAP	New York	Region	No	No	Medium	No	Information about installed capacity by region, investment plan	No data on PV energy production
SUNROOF	USA	Region	No (installed capacity mapping based on the most favorable roof characteristics)	No	Low	No	Maximum useable sunlight per city and year available for PV panels and general statistics on: number of roofs with higher installable power, roof orientations	Year
SUISSEENERGIE.CH	Switzerland	Generic Roof	No	No	Low	Yes	Annual production kWh, part of own consumption, solar current injected into the grid, installation costs, payback period	Month
PVSOL	International	Generic Roof	No	No	Low	Yes	Longitude, latitude, annual global irradiation, annual consumption estimation, best PV configuration	Year
EASY-PV	International	Generic Roof	No	No	Low	Yes	Number of panels, PV installed capacity, investment plan	No data on PV energy production
LGENERGY	International	Roof	No	No	Low	Yes	Number of panels, PV installed capacity	No data on PV energy production

databases and key data for computing the PV production (mainly location, orientation and tilt angles). There are two levels of layers according to their importance in the process: (i) exclusive base layers for which only one can be visible at a time, and (ii) overlays, or layers placed over the base layers to provide key information about the rooftops.

Base layers are included in the tool in order to facilitate the navigation of the user for the location of the rooftop of interest.

The overlays are two: the rooftop geometry layer, and the layer with cadastral data. The first is the core of the tool, since it includes all the information about the rooftops required for the analysis of the incoming solar radiation (height, orientation, tilt angle and useable area). This layer is generated with the procedure proposed in Ref. [14]. The selection of this procedure is justified by the fact of being the only procedure that provides, automatically, all the information needed for the solar analysis.

The layer with cadastral data, is a complementary layer, is not required for the analysis of the solar radiation received on the rooftop. This layer enables the estimation of the electrical consumption in the building, providing an overview of the benefits of installing PV panels on each rooftop.

2.1.1. Base layers

Currently, there are a variety of base maps available that can be used as a cartographic base for web mapping. These are maps from different sources, from national mapping agencies or companies to crowd sourcing initiatives [33,34] that have global or national coverage. Specifically, Ener3DMap-SolarWeb hosts three widely used web map data services:

- OpenStreetMap [35]: This collaborative project allows the visualization, creation and editing of vector maps by any user to create a global map that is as complete and current as possible (Fig. 1a).
- Google Maps [36]: This web-mapping service developed by Google offers a wide variety of cartographic resources, from satellite images and maps to 360° panoramic images. It also presents other interesting features, such as planning a route on foot, by car or public

transportation and real-time traffic conditions. Ener3DMap-SolarWeb uses its orthoimagery service (Fig. 1b).

- Spanish Plan for Aerial Orthophotography (PNOA) [37]: This plan contains the most recent orthophotos of the Spanish territory updated and offered by the Spanish Geographic Institute. Among its wide range of geoinformation, its high-resolution orthoimages are implemented in the web tool (Fig. 1c).

The reason for linking two seemingly equal base-maps (Fig. 1b and c) is because Google Maps offers more updated data than PNOA, while PNOA offers true orthoimages instead of the unrectified aerial images of Google Maps. This means that the PNOA base map can support geometric measurements directly, while Google Maps incorporates camera lens distortion in its images, invalidating the map as a reliable geometric information source [38]. In addition, the possibility of visualizing a map (Fig. 1a) or a photorealistic view (Fig. 1b or 1c) is included in order to adapt the tool to the preferences of the users, who can choose each visualization mode to facilitate their orientation in the study area.

2.1.2. Rooftop geometry layer

Rooftop geometry plays a key role in the SolarWeb platform, since it is the basis for computing the solar potential. Rooftop geometry is defined by several parameters that were previously obtained by the methodology presented in Ref. [14], which combines aerial images with LiDAR data and automatically obtains the orientation, tilt angle, height and dimensions of the rooftops. The advantages of using these two data sources together are reported in Ref. [14] and mainly result in increased precision. The output of this methodology is a vector layer (the rooftop geometry layer) with the parameters of orientation, tilt angle, height and useable surface associated to each roof. The vector layer is incorporated into the web tool as an overlay layer in GeoJSON format (Fig. 2a). This format is widely used in web mapping, due to its simplicity, computational lightness, quick visualization and easy data exchange.

2.1.3. Cadastral data

The cartographic database with associated building alphanumeric information which is linked to the web tool belongs to the Spanish Ministry of Finance. Among the building information provided is location, identification code, building use or destination (residential, commercial, etc.), built area, year and quality of the construction, common elements, and a graphic representation of the building (Fig. 2b). This layer is included because it is the only free dataset that provides information about the use of the building. In contrast with the base layers (raster layers), and the rooftop geometry (vector layer), cadastral data is included through a WMS connection. This type of inclusion is selected for two main reasons: (i) space optimization, and (ii) continuous updating of cadastral data in Ener3DMap Solar Web.

From all this information, Ener3DMap-SolarWeb only requires the built area and the use of the building to estimate the electrical consumption of the house or building under study. With this information, the web tool calculates the energy balance of the dwelling and the potential percentage of self-consumed electricity. This information is extracted after a proper query to the administrative database with the building location coordinates.

2.1.4. Solar radiation data

Solar radiation data is not included as a layer in the tool. In this case, due to the proven good performance of the PhotoVoltaic Geographic Information System (PVGIS) database [39], a connection is created between the SolarWeb platform and PVGIS. Specifically, a request is performed in PVGIS from the SolarWeb platform, and the results of the request are visualized in the SolarWeb platform.

PVGIS from the European Commission is a tool of the Energy Efficiency and Renewable Energy Unit of the Joint Research Center that provides solar radiation information. Ener3DMap-SolarWeb uses the PVGIS web service [40] to compute the amount of energy that can be obtained from different types of PV systems. The PVGIS solar radiation model computes the sun's irradiance by integrating the direct and diffuse radiation based on [41]. The direct radiation is modeled per hour as a function of the solar position relative to the position of interest for the computation (that is, the roof position), with computational inputs of extraterrestrial solar radiation, the distance between the Sun and the Earth, and the solar angle [42]. With this input data, the absorption of radiation from the atmosphere is included in the model. For the inclusion of the Sun position in the model, the hourly variation is considered using the position at the middle of each hour (Sun position at 10:30 for the hour between 10 and 11) [43]. The diffuse radiation is modeled based on the scattering caused by clouds as a fraction of the contribution of scattering to the direct radiation.

The radiation at each rooftop location is obtained in PVGIS after an inner query to meteorological data (solar irradiance and weather patterns) from the last 30 years, in such way that all the latest meteorological behavior variability is considered. The most important value in this case is the sky cloudiness, which determines the ratio between direct and diffuse radiation. In addition, the historical data computation allows setting different working ranges, to estimate variations in the real production of the installation in future years.

2.2. Methodology

Ener3DMap-SolarWeb Roofs is a SDI that correctly manages all the data and queries to the databases required to compute rooftop PV energy production, facilitating its use by both expert and novice users. For this reason, the tool development requires two main steps: (i) identification of the rooftop geometry and generation of the rooftop geometry layer; (ii) generation of the SDI, including its interface, the integration of all the data and the development of its functionalities (mainly the rooftop

Fig. 1. Set of base-maps offered by Ener3DMap-SolarWeb Roofs: (a) OpenStreetMap, (b) Google Maps, and (c) PNOA orthophotos.

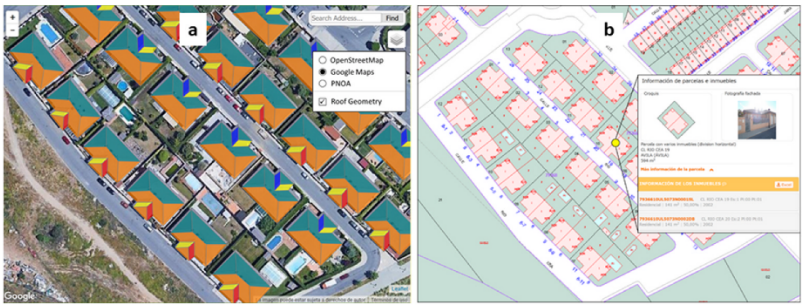


Fig. 2. (a) Layer with geometric information of existing roofs superimposed on the orthoimage from Google Maps and (b) graphic viewer of the cadastral cartographic database linked to Ener3DMap-SolarWeb Roofs.

selection and the connection with PVGIS).

2.2.1. Generation of the rooftop geometry layer from LiDAR and imagery data

The rooftop geometry (orientation, tilt angle and area) is automatically computed by using the processing methodology presented in Ref. [14]. In this methodology, detailed in Fig. 3, LiDAR data and aerial orthophotographies are combined to compute all the parameters required: LiDAR data provides orientation and tilt angle information, while the orthophotographies stand the computation of precise information about the rooftop area.

2.2.2. Generation of Ener3DMap-SolarWeb roofs SDI

Ener3DMap-SolarWeb Roofs is based on Leaflet, an Open Source JavaScript library launched in 2011. This library allows working with georeferenced data, similar to other libraries such as OpenLayers and Google Maps API. However, Leaflet is chosen because it is the only one that is open-access and lightweight, that has a fluid performance on mobile devices. In addition, Leaflet library is the most complete in terms of plugins available to extend its capabilities and functions. It supports several geographic data formats, such as WMS, GeoJSON, GeoCSV and KML, among others.

The main interface of the web tool is divided into two main parts: (a) a multilayer web viewer on the left side and (b) a complete form with all the parameters required to query the different databases on the right side (Fig. 4). In turn, the parametric part of the solar-web is subdivided into four sections: (i) a noneditable form with information regarding the location and geometry of the selected rooftop, (ii) an editable form with all the parameters required for the PV system design, (iii) a section to choose the output report format and (iv) a button that activates the query to the database for the PV energy production computation.

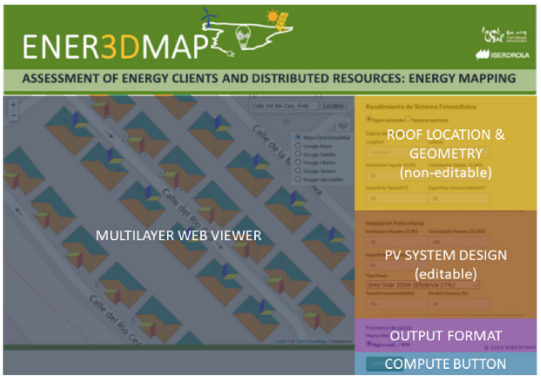


Fig. 4. The Ener3DMap-SolarWeb main interface, with its different sections.

The queries are made to the PVGIS application through its API. The call includes the location, the available area for solar panel installation, the orientation, the slope and, if desired, the climate data of the rooftop of interest. The web-based tool outputs the photovoltaic energy production of the area under study by applying the solar model behind PVGIS. The results, in JSON format, are collected from a PHP page to be

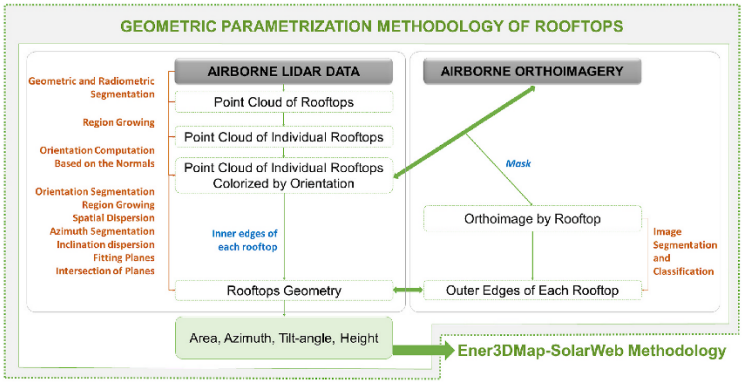


Fig. 3. Detailed workflow of the methodology for generating the Rooftop Geometry Layer from geospatial data.

displayed through the web or exported in PDF format. For the latter, the free FPDF class, written in PHP, is required.

The web-based tool is installed on a web server, using host memory; the user only needs an internet connection to access the data and results. This type of tool extends the usability of the tool to every user, without requiring a previous installation and optimizing the memory resources of the used device. The base layers and coding of the geo-database occupy 5.57 MB on the server, while the rooftop geometry layer size depends on the number of rooftops and their geometries. In particular, the memory size depends on the number of rooftop vertices; rectangular rooftops occupy more memory than triangular rooftops. In addition, data from a 1000 m^2 plot occupies the same memory as a 10 m^2 plot, provided the rooftop shape is the same. For example, the mean memory occupation is 5 KB per 1000 m^2 for the most typical rooftop shapes (hip roofs and M-shaped roofs in industry).

The integration of all the data layers and the connections to the Cadastral data and PVGIS APIs yields the following outputs:

- Production of the system under study per month (numerically and graphically)
- Mean solar radiation (in kWh per m^2 and day)
- Plane of Array, or total incident radiation on the panel, adding direct solar rays, diffuse irradiance and irradiation reflected by the ground onto the panel
- Estimated energy balance, which provides a guide for the energy savings produced by the installation

The energy balance consists of estimating the energy consumption based on the extension and use of the building, and computing the energy generated by the PV installation. The building extension and use information is extracted from the cadastral data integrated in the geo-database (Section 2.1.3). In this sense, ENER3DMAP-SolarWeb approximates the energy consumption as other solar tools provide (PVSOL), based on the building area. The approximation is based on a top-down approach [44] using consumer profiles from the Spanish Electricity Network. This approach is selected to avoid requesting more data from the user, since an accurate energy consumption estimation requires actual consumption data, or data about the time of active occupancy, the number of occupants [45] and the energy use [46]. In addition, accurately estimating the energy consumption is beyond the current scope and objective of the proposed tool, the focus and potential of which are in calculating the photovoltaic solar potential of built and unbuilt roofs.

3. Operation of the Ener3DMap SolarWeb-Roofs SDI

The main body of the SDI corresponds to the area which allows the visualization of the different layers integrated into the web tool: the base layers (Section 2.1.1) and the vector layer with the roof geometric information (Section 2.1.2). The latter represents the geometry of the existing rooftops in plan view with a color assigned to each slope according to its orientation: blue for North, turquoise for Northwest and Northeast, yellow for West and East, orange for Southwest and Southeast, and red for South. The user can easily locate the target rooftop (either existing or under construction) by using the search bar and investigate different possible PV panel configurations in the right side of the web tool.

The procedure for tool use and its inner operation are detailed in the following sections and described in Fig. 5.

There is no mention in the manuscript to the Forecast API included in Fig. 5 because it is a connection analogous to the connection made with PVGIS. The Forecast API has been linked to Ener3DMap SolarWeb-Roofs tool for management purposes once the PV installation is done. Since this part is out of the scope of the study, it has not been described in detail.

3.1. Case study selection

Once the target rooftop, or the desired solar installation position, are selected in the viewer by clicking on the base layer (Section 2.1.1), the query form is automatically populated or left for the user to define, depending on whether the analysis is targeted at pre-existing rooftops. For existing rooftops (Fig. 6a), the latitude, longitude, tilt angle, orientation and useable area of the target rooftop are automatically populated in the form from the Rooftop Geometry Layer. For a solar panel installation position without a building (Fig. 6b), once the area of interest has been selected, the longitude and latitude of the study location is automatically populated into the form.

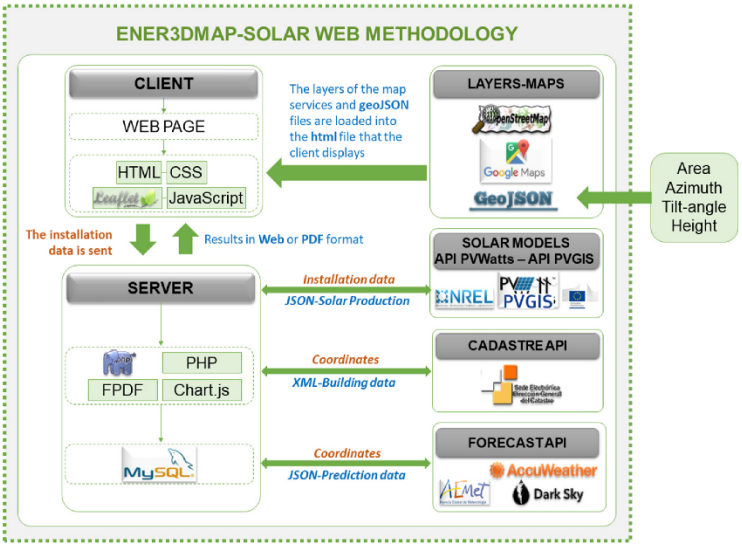


Fig. 5. Operation of the ENER3DMAP SolarWeb-Roofs SDI.

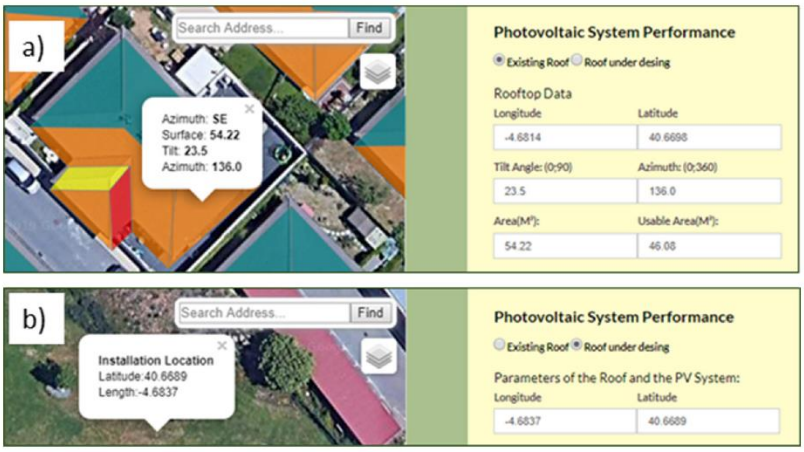


Fig. 6. Roof location parameters automatically populated into the form for a query for (a) an existing rooftop or (b) a parcel without a building.

3.2. PV system design

In addition to the rooftop data form, the interface has a form for PV system design. The contents of this form also depend on whether the target is an existing or projected rooftop. For both cases, the form is editable. For existing rooftops, the PV system design form is automatically populated with default parameters obtained from the rooftop geometry layer, assuming an integrated PV system: the panels follow the same orientation and inclination as the target rooftop (Fig. 7a). However, the form can be edited to set different configurations to check their energy performance. For locations with no building or no roof, there is a blank form with four tabs that must be filled by the user with the preferred design and configuration parameters and two tabs that are automatically populated based on the previous manually entered parameters (Fig. 7b).

In both cases, the three most widely used panel models and configurations are presented as user choices (Fig. 7): (i) Canadian Solar 275 W, (ii) Jinko Solar 330 W and (iii) Canadian Solar 330 W. In addition, for each model and configuration, the solar capacity in kW and the system losses in % are automatically computed.

This data entry form offers different alternatives for computing the PV potential for both existing rooftops and locations without a building.

3.2.1. Photovoltaic solar potential of existing rooftops

The rooftop geometry is important when calculating the PV solar potential of existing rooftops since it affects both the number of solar panels and the way in which the panels receive solar radiation. Due to the great PV technology development, currently there are several photovoltaic glass solutions and types of solar panels that suit any geometry. However, since the objective of this web tool is to offer information about the PV solar potential of rooftops when installing the most common panels in the market, the computation is performed for the panels selected in the tool, with approximately 2 m^2 of surface area.

In addition to rooftop geometry, the way in which the panels are installed is another factor that determines both the installation cost and the final energy production. PV panels can be (i) integrated on the rooftop or (ii) fixed on structural supports that modify their tilt angle and orientation. For both cases, the web tool performs the calculation based on the rooftop geometry, the site location, and the solar radiation model (Sections 2.1.2 and 2.1.4).

3.2.1.1. PV panels integrated on the rooftop. This is the most commonly

Fig. 7. Form for the introduction of data for the PV system design, for cases with (a) an existing rooftop or (b) a location without a building.

adopted solution; since the solar panels are completely aligned with the rooftop, no visual integration problems occur. It is necessary to ensure that the rooftop geometry layer is activated to show a layer in which the rooftop slopes are in color (blue, turquoise, orange, yellow or red, depending on the orientation).

The rooftop of interest must be located through the search bar or with the mouse controls and selected. The desired panel type among the three offered must be selected. The remaining rooftop parameters and those of the PV system are automatically populated.

By clicking on the compute button, the query to the solar radiation database is launched and after few seconds the solar potential of the chosen roof slope is shown in the user-selected output format. To obtain the total rooftop solar radiation, a query must be made for each roof slope [14].

3.2.1.2. PV panels on support structures modifying the rooftop geometry. This solution optimizes the production by modifying the inclination and orientation of the PV panels with respect to the rooftop. Fewer panels fit on the roof, since gaps between consecutive panel rows must be established to ensure that no shadows are projected among the panels. To compute the reduction of useable area for the panels, the date for which the largest shadow is projected must be considered, which corresponds to the winter solstice. On this date, a panel's shadow is approximately the same size as the panel, such that, for example, for 330 W panels, which are approximately 1.95 m^2, the spacing between consecutive panel rows must be approximately 2 m, reducing the rooftop's useable area by one-half.

For this reason, a percentage reduction is applied by default to compute the useable rooftop area. This percentage is computed as a function of the tilt angle selected for the solar PV installation.

3.2.2. Photovoltaic solar potential for locations without a building

When planning a PV solar installation during a building's design or construction phase, it is possible not only to determine the optimal rooftop orientation/tilt angle to maximize the production of the PV installation but also to optimize the rooftop geometry to maximize the useable area for the placement of solar panels, to choose the best roof material option to facilitate the installation of the panels and to consider PV installation loads together with wind loads on the rooftop to ensure the structure's stability.

In addition, the possibility of introducing the solar installation parameters manually allows the use of the web tool to design PV solar

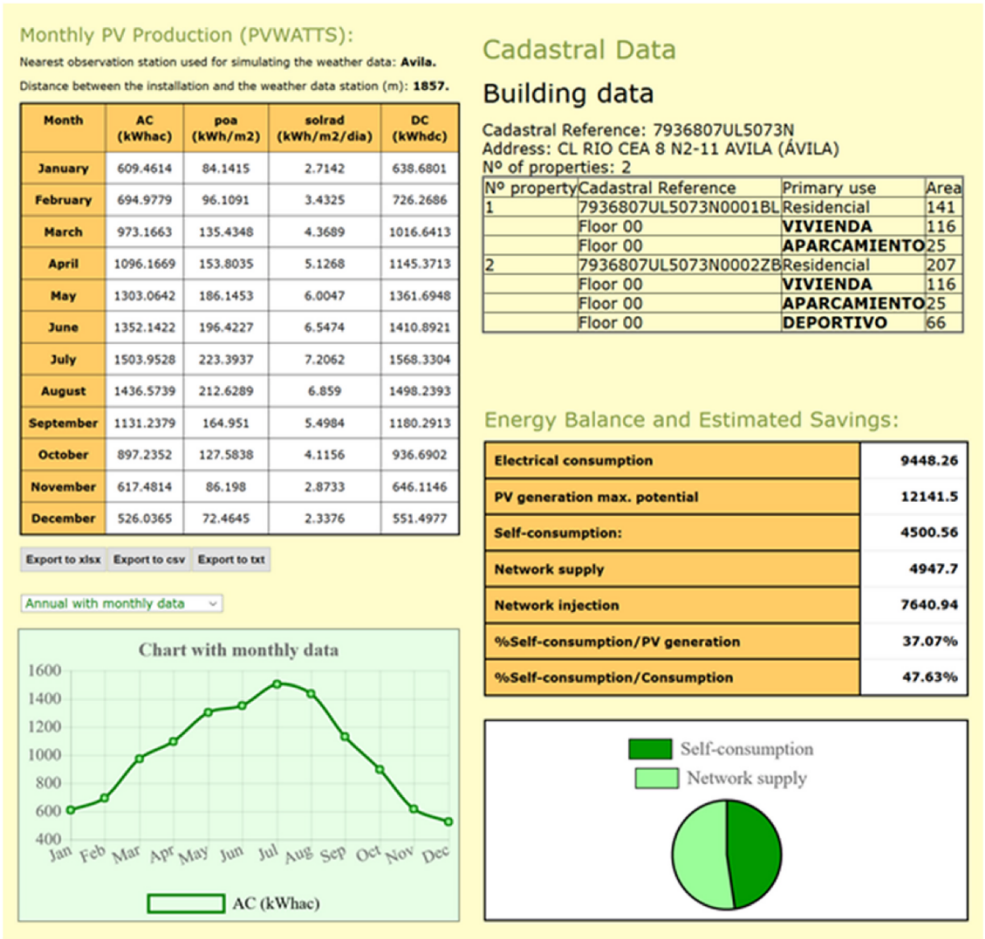


Fig. 8. Output of the web tool in web format: results of PV energy production and approximate energy balance of the case study.

farms directly on the ground, with no roof required.

3.3. Computation of solar energy parameters

Once the rooftop geometry parameters and the PV system configuration have been established, the result of solar production is obtained after establishing the output format, either via the web or through a pdf file (Figs. 8 and 9), and after pressing the execution button (Fig. 4). Output generation implies a query has been sent to the cadastral and the solar radiation (PVGIS) layers, which requires 4–5 s. However, the computation time depends on the user's web connection speed.

4. Evaluation of the Ener3DMap Solar-Web Roofs Tool

This section aims to show the versatility and robustness of the Solar-Web Roofs geospatial tool not only in terms of effective multisource data management but in terms of automation and query efficiency. Due to the data forms in the web tool, it is possible to make consecutive queries to find the optimal PV solution based on the end user's economic and/or energy priorities. Thus, this tool helps find ad hoc solutions-based not only on the 3D geometric characteristics of the rooftop under study and its PV solar potential but also considering different panel configurations to maximize energy production, minimize the number of PV panels or find intermediate solutions.

4.1. Experimental validation

To show these prospects, several studies have been performed to compute the energy production of adopting different rooftop and solar panel configurations. Specifically, three different-use buildings in the city of Avila were selected: a single-family residential building with a 7-slope hip-valley roof (Figs. 10–1a), a multifamily building with 8 residential units and a 4-slope hip roof (Figs. 10–2a), and an industrial building with a 4-slope M-shaped roof (Figs. 10–3a). For each building, three simulations with the following configurations were carried out:

- Fig. 10c: maximum number of PV panels integrated on the rooftop considering all roof slopes.
- Fig. 10d: maximum number of PV panels integrated on the rooftop considering only slopes facing West, Southwest, South, Southeast and East (W, SW, S, SE, E).
- Fig. 10e: maximum number of PV panels with the optimal tilt angle and orientation.

For the latter (Fig. 10e), the following assumptions were considered:

- Since the case studies were in the northern hemisphere, the simulation was established with the panels facing to the South (180° azimuth) and with the tilt angle that maximizes the annual energy production (tilt angle of 37° - 39°). For the case studies shown, located in the city of Ávila (at a 40.67° North latitude), a panel inclination of 34° guarantees the maximum annual energy production [47].
- Since slanted panels cast a shadow, enough space must be provided between consecutive rows of panels to ensure they are not shaded. To size the installation, the date of the winter solstice is considered, which corresponds to December 21st in the northern hemisphere and 13:12 p.m. local time for Ávila, when the sun reaches an inclination of 25°. On this date, the spacing between consecutive rows of panels of 330 W must be approximately 2 m, which reduces the rooftop useable area by one-half.

The main spatial and geometric characteristics of the studied rooftops as well as the PV system configuration for each case are listed in Table 2.

Table 3 shows the annual energy production simulation result with

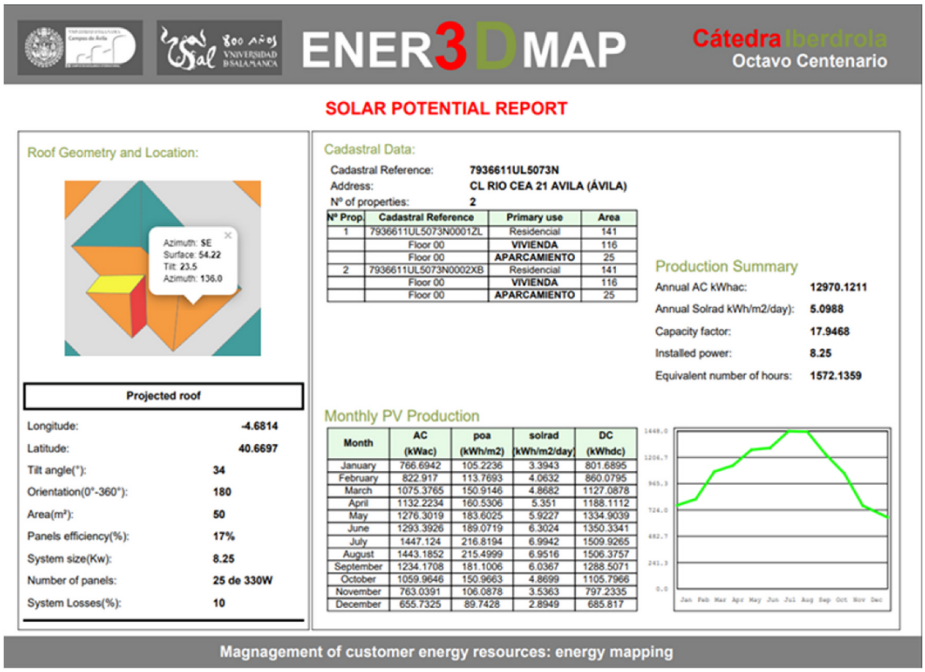


Fig. 9. Output of the web tool in pdf format: results of PV energy production and approximate energy balance of the case study.

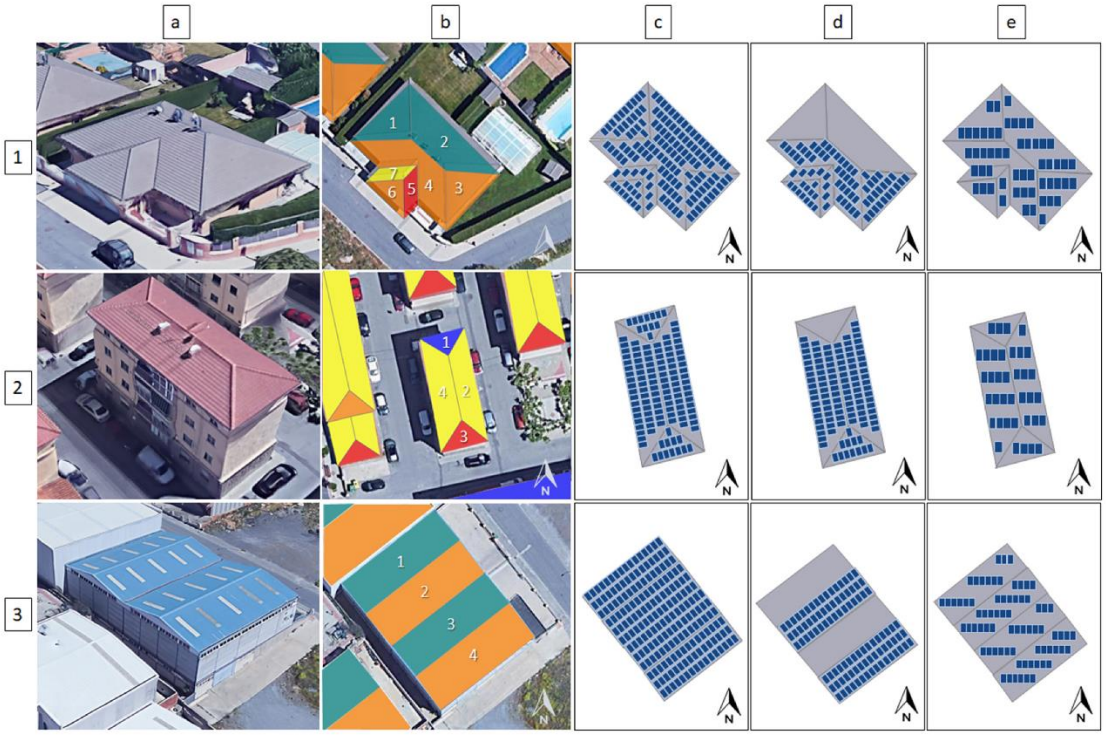


Fig. 10. Validation tests performed for the Ener3DMap-SolarWeb tool: 1, 2 and 3 are the different studied cases. Columns: (a) actual roofs; (b) rooftops automatically identified by the tool; (c) integration of the panels using all the useable area; (d) integration of the panels using only slopes facing West, Southwest, South, Southeast and East, considered optimal for the reception of solar radiation; (e) panel installation with optimal orientation and tilt-angle (nonintegrated panels).

Ener3DMap-SolarWeb Roofs and the three different PV panel installation configurations (c, d and e). To draw conclusions about the production-cost ratio of each PV configuration, in addition to the data offered by the web tool (total energy production per year and per month), the energy production per panel and per unit of rooftop surface was calculated (Table 2).

For all cases studied:

- The highest energy production per year is obtained with configuration (c), that is, covering the rooftop completely with PV panels integrated at the orientation and tilt angle of the rooftop (Fig. 9).
- The highest annual energy production per unit of rooftop surface (highest roof productivity) is obtained with configuration (d), that is, the installation of the maximum number of PV panels integrated on the rooftop considering only slopes facing W, SW, S, SE, and E.
- The highest annual energy production per panel (best panel performance) is obtained with configuration (e), that is, the installation of the maximum number of photovoltaic panels in their optimal inclination and orientation.

Fig. 11 shows the annual evolution of the energy production for each studied case and panel configuration. The small difference in production between configurations (c) and (d) in case study 2 is because only one roof slope presents a nonoptimal orientation and thus the reduction in panel area from one configuration to the other is minimal.

Table 4 compares the time required to perform each of the previous studies. The time required to analyze configurations c and d with Ener3DMap SolarWeb-Roofs tool is 40% less than that required with PVGIS. The reduction is due to proposed tool's automatic introduction of rooftop data. For the study of configuration e, PVGIS has similar computation times as Ener3DMap, because PVGIS includes an automatic introduction of parameters to "optimize tilt angle and orientation". As a complementary contribution, Ener3DMap SolarWeb-Roofs computes the number of panels for each configuration, in addition to the solar radiation per unit area, while PVGIS only computes the solar radiation per unit area.

The aim of these simulations is not to draw a generic conclusion about the ideal PV configuration since many configuration possibilities were not considered, and because the ideal configuration also depends on factors such as users' energy needs, their economic resources and the building use (main or secondary use, for example). The purpose of these simulations is to validate the usefulness of the different data types included in the tool, and the versatility and support offered by the web tool developed to calculate PV solutions. In addition, the tool does not perform a structural study of the roof including the PV panels, such that the preferred installation viability study is incomplete in this sense. In fact, Ener3DMap-SolarWeb Roofs gives a warning in this regard.

4.2. *Theoretical validation*

The precision and accuracy of the results obtained by the Ener3DMap Solar-Web Roofs tool involve a combination of the precision and accuracy of its two components: the accuracy of the geometry determination with the algorithms developed and presented in Refs. [14], and the model's precision in computing the solar energy production in PVGIS.

For the rooftop geometry [14], determines that the methodology

Table 2
Spatial and geometric characteristics of the analyzed rooftops and their PV panel configurations.

Case Study		Roof Location		Rooftop 3D Geometry				PV System			
Roof/config.	Roof slope	Lat.[a]	Long.[b]	Az.[c]	Or.[d]	Tilt[e]	A.[f]	Az.[c]	Or.[d]	Tilt[e]	A.[f]
1c	1	−4,68	40,67	318	NW	23,50	60,30	318	NW	23,50	51,26
	2			45	NE	24,00	112,21	45	NE	24,00	95,38
	3			136	SE	23,50	57,95	136	SE	23,50	49,25
	4			223	SW	21,00	107,49	223	SE	21,00	91,37
	5			182	S	16,00	18,51	182	S	16,00	15,73
	6			224	SW	19,00	25,47	224	SW	19,00	21,65
	7			275	W	17,00	16,54	275	W	17,00	14,06
1d	3	−4,68	40,67	136	SE	23,50	57,95	136	SE	23,50	49,25
	4			223	SW	21,00	107,49	223	SE	21,00	91,37
	5			182	S	16,00	18,51	182	S	16,00	15,73
	6			224	SW	19,00	25,47	224	SW	19,00	21,65
	7			275	W	17,00	16,54	275	W	17,00	14,06
1e	1	−4,68	40,67	318	NW	23,50	60,30	180	S	34,00	25,63
	2			45	NE	24,00	112,21	180	S	34,00	47,69
	3			136	SE	23,50	57,95	180	S	34,00	24,63
	4			223	SW	21,00	107,49	180	S	34,00	45,69
	5			182	S	16,00	18,51	180	S	34,00	7,87
	6			224	SW	19,00	25,47	180	S	34,00	10,83
	7			275	W	17,00	16,54	180	S	34,00	7,03
2c	1	−4,69	40,66	350	N	14,00	16,56	350	N	14,00	14,08
	2			82	E	16,00	61,85	82	E	16,00	52,57
	3			170	S	14,50	20,67	170	S	14,50	17,57
	4			260	W	17,00	73,55	260	W	17,00	62,52
2d	2	−4,69	40,66	82	E	16,00	61,85	82	E	16,00	52,57
	3			170	S	14,50	20,67	170	S	14,50	17,57
	4			260	W	17,00	73,55	260	W	17,00	62,52
2e	1	−4,69	40,66	350	N	14,00	16,56	180	S	34,00	7,04
	2			82	E	16,00	61,85	180	S	34,00	26,29
	3			170	S	14,50	20,67	180	S	34,00	8,79
	4			260	W	17,00	73,55	180	S	34,00	31,26
3c	1	−4,67	40,66	321	NW	10,50	285,18	321	NW	10,50	242,40
	2			143	SE	10,80	300,47	143	SE	10,80	255,40
	3			321	NW	10,50	289,55	321	NW	10,50	246,12
	4			143	SE	10,50	293,64	143	SE	10,50	249,59
3d	2	−4,67	40,66	143	SE	10,80	300,47	143	SE	10,80	255,40
	4			143	SE	10,50	293,64	143	SE	10,50	249,59
3e	1	4,67	40,66	321	NW	10,50	285,18	180	S	34,00	121,20
	2			143	SE	10,80	300,47	180	S	34,00	127,70
	3			321	NW	10,50	289,55	180	S	34,00	123,06
	4			143	SE	10,50	293,64	180	S	34,00	124,79

[a] Latitude (°).
[b] Longitude (°).
[c] Azimuth (0°–360°).
[d] Orientation.
[e] Tilt Angle (°).
[f] Area (m^2).

Table 3
Energy production for the cases studied.

Case	Num. Panels	Production[a]	Production/Panel[a]	Production/m^2 rooftop[a]	Installed Power	Equivalent hours PS[b]
1c	173	73,666	425,81	184,87	57,09	9346,60
1d	98	47,346	483,12	209,53	32,34	7217,58
1e	84	43,579	518,81	109,37	27,72	11,004,95
2c	75	32,731	436,41	189,60	24,75	5257,80
2d	68	30,103	442,70	192,88	22,44	4120,33
2e	36	18,677	518,81	108,19	11,88	6288,54
3c	509	223,100	438,31	190,87	167,97	5305,49
3d	259	122,219	471,90	205,72	85,47	2859,92
3e	254	131,776	518,81	112,74	83,82	6288,54

[a] kWh/year.
[b] Peak Sun.

error is 0.19% and 0.45% for the computation of the orientation and the tilt-angle of the roof, respectively, and 4.6% for the determination of the useable area. As stated in Ref. [14], this error can be considered negligible, as it implies the underestimation or overestimation of one or zero panels.

With respect to the PVGIS results, a thorough evaluation of the software used to compute the solar energy production is presented in Refs. [39]. The analysis shows that PVGIS is the most accurate free software package for estimating annual solar energy, with a mean deviation of 2.5%, compared to 8.8% and 8.1% for PVWatts and RETScreen, respectively. For all cases, the winter energy production has a higher deviation due to the complexity in predicting the weather for

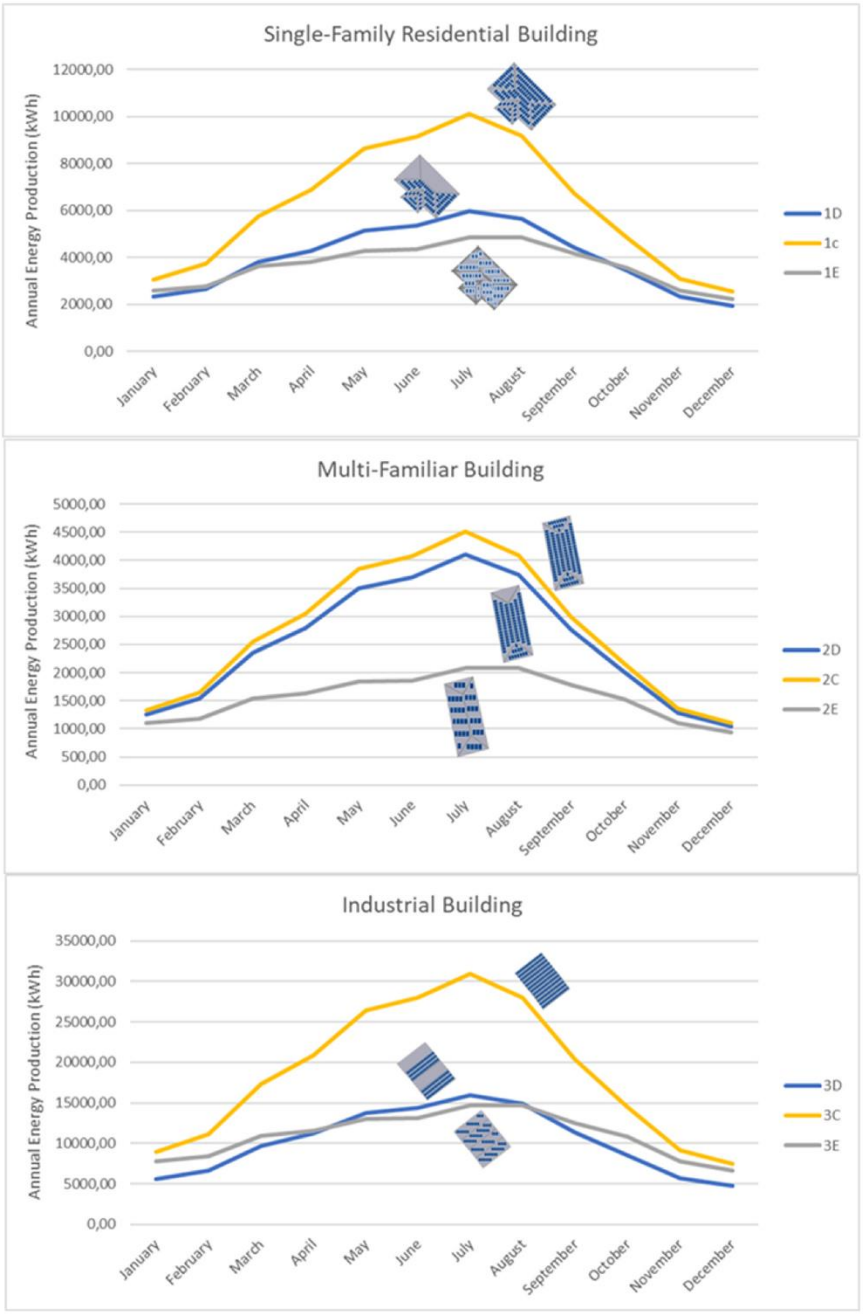


Fig. 11. Evolution of the annual energy production for different building uses as a function of the different solar panel installation configurations.

this season, despite that all approaches use historic weather data.

To assess the performance of PVGIS compared with commercial software [48], PVGIS and TRNSYS, Archelios, Polysun, PVSyst and PVSOL were analyzed with real data from a PV installation, all using weather data measured at the installation position, except for PVGIS. According to this study, the estimation accuracy of PVGIS is in the same range as the commercial software that perform PV planning and analysis, such as PVSyst and PVSOL, with similar RMSE, MAD and MAPE values (root mean square error, mean absolute deviation, mean absolute percentage error) of 244.88; 21.42 and 9.24%, respectively [49]. Also

Table 4
Processing time for the different case of studies and configurations: Ener3DMap vs. PVGIS.

	Case 1			Case 2			Case 3		
	Conf.c	Conf.d	Conf.e	Conf.c	Conf.d	Conf.e	Conf.c	Conf.d	Conf.e
Ener3DMap-SolarWeb	42 s.	24 s.	38 s.	21 s.	15 s.	20 s.	25 s.	11 s.	24 s.
PVGIS	76 s.	55 s.	45 s.	43 s.	30 s.	25 s.	46 s.	25 s.	32 s.

determines the error of PVGIS for the simulation of the global irradiance with respect to the ground measurements, with an R^2 between 0.82 and 0.92. Thus, PVGIS can be considered an accurate alternative for estimating solar PV potential, and the error regarding this issue in the Ener3DMap SolarWeb tool is considered acceptable, since the tool is based on PVGIS.

5. Conclusions

This paper presents the development of a web-based tool called Ener3dmap Solar-Web Roofs. The tool is based on Leaflet and supports the main formats for geospatial data: WMS, GeoJSON, GeoCSV and KML.

The tool integrates 3D rooftop geometry data with a widely validated solar radiation model, PVGIS, with a high level of automation. The level of automation is accomplished with algorithms that compute rooftop geometric parameters from LiDAR data and aerial orthophotography. The availability and format of these datasets is standardized in European countries under the INSPIRE Directive. With the use of these data, the replicability of the tool in other European countries is ensured.

In addition to automatic data entry, the tool provides user editing capability, both for the solar PV installation design and for the introduction of data for parcels without buildings. The developed tool matches the capability of existing tools, of allowing the introduction of data by the user.

Last, the integration of a validated solar radiation model such as PVGIS presents an advance over existing tools in the frequency with which the solar radiation results can be computed: hourly, monthly and annually. Most other tools only provide annual solar radiation data.

The tool presented in this work provides an accuracy below 0.50% for computing the rooftop angular parameters (orientation and tilt-angle), and below 5% for the determination of the useable area. The error in computing the solar energy production is 2.5%. With these values, the contribution of the tool is two-fold:

1 With respect to roof modeling tools, Ener3DMap SolarWeb Roofs provides the solar energy production per hour, month and year.
2 With respect to existing solar planning tools, Ener3DMap SolarWeb Roofs provides the rooftop geometric parameters automatically from geomatic data.

An example of the proposed tool's versatility is that it enables the study of different solar panel configurations with agility (with a mean computation time reduction of 40% with respect to the solar energy computation of PVGIS alone). In this way, for example, it was possible to determine that, for a certain rooftop, the highest energy production per year is obtained when the rooftop is completely covered with PV panels that are integrated with the rooftop's orientation and tilt angle.

Taking these facts into account, the main contribution of the study to society is the availability of a tool for the accurate but easy computation of solar PV installations, facilitating the incorporation of this renewable and clean energy to the electricity network and increasing the accessibility to solar energy for all types of users (residential, industrial), promoting self-consumption and reducing energy poverty.

Future studies will address the further improvement of the integrated solar radiation model, to reduce the 2.5% deviation of PVGIS, either by integrating additional historic weather data, or by incorporating satellite measurements (land surface temperature, albedo, solar radiation on the atmosphere) into the mathematical model.

Funding sources

This work was supported by Iberdrola S.L. and the University of Salamanca [Cátedra Iberdrola VIII Centenary]; and the Ministry of Science, Innovation and Universities [project RTC-2017-6291-3].

CRediT authorship contribution statement

M. Sánchez-Aparicio: Validation, Writing - original draft. **J. Martín-Jiménez:** Software, Visualization. **S. Del Pozo:** Methodology, Writing - original draft. **E. González-González:** Resources. **S. Lagüela:** Conceptualization, Writing - review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank AJE for the English editing and TIDOP Research Group for all the academic support and administrative guidance.

References

[1] European Commission (EC). Europe 2020: a strategy for smart, sustainable and inclusive growth. Working paper {COM (2010) 2020}. 2010.
[2] Mohajeri N, Perera ATD, Coccolo S, Mosca L, Le Guen M, Scartezzini JL. Integrating urban form and distributed energy systems: assessment of sustainable development scenarios for a Swiss village to 2050. Renew Energy 2019;143:810–26.
[3] Zhao Z, Zhang S, Hubbard B, Yao X. The emergence of the solar photovoltaic power industry in China. Renew Sustain Energy Rev 2013;21:229–36.
[4] Gautam B, Li F, Ru G. Assessment of urban roof top solar photovoltaic potential to solve power shortage problem in Nepal. Energy Build 2015;86:735–44.
[5] Hartner M, Ortner A, Hiesl A, Haas R. East to west – the optimal tilt angle and orientation of photovoltaic panels from an electricity system perspective. Appl Energy 2015;160:94–107.
[6] Luthander R, Widén J, Nilsson D, Palm J. Photovoltaic self-consumption in buildings: a review. Appl Energy 2015;142:80–94.
[7] Kabir E, Kumar P, Kumar S, Adelodun AA, Kim KH. Solar energy: potential and future prospects. Renew Sustain Energy Rev 2018;82:894–900.
[8] Zhang Y, Ren J, Pu Y, Wang P. Solar energy potential assessment: a framework to integrate geographic, technological, and economic indices for a potential analysis. Renew Energy 2020;149:577–86.
[9] Petrichenko K, Ürge-Vorsatz D, Cabeza LF. Modeling global and regional potentials for building-integrated solar energy generation. Energy Build 2019;198:329–39.
[10] Phap VM, Huong NTT, Hanh PT, Van Duy P, Van Binh D. Assessment of rooftop solar power technical potential in Hanoi city, Vietnam. J Building Eng 2020:101528.
[11] Zurqani HA, Post CJ, Mikhailova EA, Schlautman MA, Sharp JL. Geospatial analysis of land use change in the Savannah River basin using Google Earth Engine. Int J Appl Earth Obs Geoinf 2018;69:175–85.
[12] Eskandari S, Chuvieco E. Fire danger assessment in Iran based on geospatial information. Int J Appl Earth Obs Geoinf 2015;42:57–64.
[13] Choi Y, Suh J, Kim SM. GIS-based solar radiation mapping, site evaluation, and potential assessment: a review. Appl Sci 2019;9(9):1960.
[14] Martín-Jiménez J, Del Pozo S, Sánchez-Aparicio M, Lagüela S. Multi-scale roof characterization from lidar data and aerial orthoimagery: automatic computation of building photovoltaic capacity. Autom ConStruct 2020;109:102965.

[15] Directive, I. N. S. P. I. R. E.. Directive 2007/2/EC of the European parliament and of the council of 14 march 2007 establishing an infrastructure for spatial information in the European community (INSPIRE). Published in the official journal on the 25th april. 2007.
[16] Mahdy M, Bahaj AS. Multi criteria decision analysis for offshore wind energy potential in Egypt. Renew Energy 2018;118:278–89.
[17] Gigovic L, Pamucar D, Bozanic D, Ljubojevic S. Application of the GIS-DANP-MABAC multi-criteria model for selecting the location of wind farms: a case study of Vojvodina, Serbia. Renew Energy 2017;103:501–21.
[18] Santilano A, Donato A, Galgaro A, Montanari D, Menghini A, Viezzoli A, et al. An integrated 3D approach to assess the geothermal heat-exchange potential: the case study of western Sicily (southern Italy). Renew Energy 2016;97:611–24.
[19] Choudhury S, Parida A, Pant RM, Chatterjee S. GIS augmented computational intelligence technique for rural cluster electrification through prioritized site selection of micro-hydro power generation system. Renew Energy 2019;142: 487–96.
[20] Castro-Santos L, García GP, Simões T, Estanqueiro A. Planning of the installation of offshore renewable energies: a GIS approach of the Portuguese roadmap. Renew Energy 2019;132:1251–62.
[21] Chukwuma EC. Facility location allocation modelling for bio-energy system in Anambra State of Nigeria: integration of GIS and location model. Renew Energy 2019;141:460–7.
[22] Omitaomu OA, Blevins BR, Jochem WC, Mays GT, Belles R, Hadley SW, et al. Adapting a GIS-based multicriteria decision analysis approach for evaluating new power generating sites. Appl Energy 2012;96:292–301.
[23] Firozjaei MK, Nematollahi O, Mijani N, Shorabeh SN, Firozjaei HK, Toomanian A. An integrated GIS-based Ordered Weighted Averaging analysis for solar energy evaluation in Iran: current conditions and future planning. Renew Energy 2019; 136:1130–46.
[24] Hong T, Koo C, Park J, Park HD. A GIS-based optimization model for estimating the electricity generation of the rooftop PV (photovoltaic) system. Energy 2014;65: 190–9.
[25] Wong MS, Zhu R, Liu Z, Lu L, Peng J, Tang Z, et al. Estimation of Hong Kong's solar energy potential using GIS and remote sensing technologies. Renew Energy 2016; 99:325–35.
[26] Dehwah AHA, Asif M. Assessment of net energy contribution to buildings by rooftop photovoltaic systems in hot-humid climates. Renew Energy 2019;131: 1288–99.
[27] Budhathoki NR, Nedovic-Budic Z. Reconceptualizing the role of the user of spatial data infrastructure. Géo 2008;72(3–4):149–60.
[28] Freitas S, Catita C, Redweik P, Brito MC. Modelling solar potential in the urban environment: state-of-the-art review. Renew Sustain Energy Rev 2015;41:915–31.
[29] Huang Z, Mendis T, Xu S. Urban solar utilization potential mapping via deep learning technology: a case study of Wuhan, China. Appl Energy 2019;250:283–91.
[30] Chen S, Zhang Y, Nie K, Li X, Wang W. Extracting building areas from photogrammetric DSM and DOM by automatically selecting training samples from historical DLG data. ISPRS Int J Geo-Inf 2020;9(1):18.
[31] Zhou Z, Gong J. Automated residential building detection from airborne LiDAR data with deep neural networks. Adv Eng Inf 2018;36:229–41.
[32] https://www.google.com/get/sunroof. [Accessed 11 June 2019].
[33] Skopeliti A, Stamou L. Online map services: contemporary cartography or a new cartographic culture? ISPRS Int J Geo-Inf 2019;8(5):215.
[34] Wang Cong. Usability evaluation of public web mapping sites. ISPRS Ann Photog Rem Sens Spat Inform Sci 2014;2:4.
[35] https://www.openstreetmap.org/#map 6/40.007/-2.488. [Accessed 11 October 2019].
[36] https://www.google.com/maps. [Accessed 11 October 2019].
[37] https://pnoa.ign.es/. [Accessed 11 October 2019].
[38] Dall JA. Perfect base for a GIS: digital orthophoto, second national specialty conference on civil engineering applications of remote sensing and geographic information systems. 1991. Washington, USA.
[39] Psomopoulos CS, Ioannidis GC, Kaminaris SD, Mardikis KD, Katsikas NG. A comparative evaluation of photovoltaic electricity production assessment software (PVGIS, PVWatts and RETScreen). Environ Proc 2015;2:S175–89.
[40] PVGIS web service. https://re.jrc.ec.europa.eu/pvg_static/web_service.html. [Accessed 11 June 2019].
[41] Pérez R, Ineichen P, Seals R, Michalsky J, Stewart R. Modeling daylight availability and irradiance components from direct and global irradiance. Sol Energy 1990;44 (5):271–89.
[42] Ideriah FJK. A model for calculating direct and diffuse solar radiation. Sol Energy 1981;26:447–52.
[43] Dobos AP. PVWatts version 5 manual. NREL Publications; 2014. NREL/TP-6A20-62641.
[44] Swan LG, Ugursal VI. Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. Renew Sustain Energy Rev 2009;13(8): 1819–35.
[45] López-Rodríguez MA, Santiago I, Trillo-Montero D, Torriti J, Moreno-Muñoz A. Analysis and modeling of active occupancy of the residential sector in Spain: an indicator of residential electricity consumption. Energy Pol;62:742-751.
[46] McKerracher C, Torriti J. Energy consumption feedback in perspective: integrating Australian data to meta-analyses on in home displays. Energy Effi 2013;6(2): 387–405.
[47] Jain D, Lalwani M. A review on optimal inclination angles for solar arrays. Int J Renew Energy Resour 2017;7(3):1053–61.
[48] Axaopoulos PJ, Fylladitakis ED, Gkarakis K. Accuracy analysis of software for the estimation and planning of photovoltaic installations. Int J Energy Environ Eng 2014;5:1.
[49] Psiloglou BE, Kambezidis HD, Kaskaoutis DG, Karagiannis D, Polo JM. Comparison between MRM simulations, CAMS and PVGIS databases with measured solar radiation components at the Methoni station, Greece. Renew Energy 2020;146: 1372–91.

Capítulo VI

Conclusiones y líneas futuras

6. Conclusiones y líneas futuras

En este capítulo se incluyen las conclusiones de la investigación realizada, y se exponen las líneas futuras que se abren para continuar avanzando y complementando la investigación.

6.1. Conclusiones

El objetivo principal de la Tesis Doctoral, conseguir la automatización del procesado de datos geoespaciales de infraestructuras y edificios, ha sido conseguido. Del mismo modo, han sido conseguidos los objetivos específicos con las herramientas desarrolladas (apéndice B): ***inRoad inAlert*** para la categorización de los tramos de las vías según su riesgo en base a las características del trazado, ***Ener3DMap-SolarRoofs*** para la extracción de las características de cada agua del tejado, ***Ener3DMap-SolarWeb Roofs*** para el cálculo de la producción solar fotovoltaica enlazando con la API de PVGIS.

Esta investigación muestra que es posible definir estrategias, metodologías y algoritmos para automatizar el procesado de datos procedentes de digitalizaciones 3D de infraestructuras y edificios, para dotarlos de propiedades semánticas que facilitan el manejo de esa información. En esta Tesis Doctoral se aportan soluciones de automatización, cuando el volumen de información a procesar es muy grande. Este tamaño puede ser debido a que la resolución espacial es muy buena, si se trabaja con datos procedentes de sistemas de cartografiado móviles. Con resoluciones espaciales no tan grandes en el caso de datos procedentes de vuelos fotogramétricos, el gran volumen de datos a procesar puede venir por el número de infraestructuras o por el tamaño de las mismas.

En los estudios realizados se han obtenido soluciones que facilitan el trabajo con la información geoespacial para profesionales de distintos sectores. Además, se consigue con las herramientas web desarrolladas permitir el uso de la información procesada, de un modo simple e intuitivo, por parte de usuarios finales, que no tienen que ser expertos en la materia.

El procesado de nubes de puntos obtenidos con la tecnología MLS puede ofrecer la evaluación de riesgos de una infraestructura vial mediante: un mapeo automático de riesgos de la carretera basado en índices de consistencia geométrica y criterios de estabilidad precisos derivados de estos índices. Esto se consigue con el enfoque integral presentado, que combina geometría y riesgo, utilizando nubes de puntos MLS 3D. Este enfoque representa un método novedoso de evaluar la seguridad vial a partir

de los datos MLS y apoyado en un proceso de razonamiento inductivo, que proporciona una evaluación del riesgo potencial, basada exclusivamente en parámetros geométricos intrínsecos al trazado de la vía. Este índice de riesgo potencial, puede complementar adecuadamente otros enfoques de seguridad, como EuroRAP, que se basa en un enfoque estadístico-cualitativo, sensible al efecto distorsionador de los accidentes de tráfico. En definitiva, se ha obtenido un alto grado de fiabilidad en la extracción de los elementos geométricos del alineamiento horizontal de la vía. Esto es evidente considerando las bajas discrepancias obtenidas para los principales parámetros geométricos. También destacan las diferencias mínimas observadas en los índices de consistencia geométrica y en los criterios de estabilidad.

En el procesado de datos procedentes de la digitalización de infraestructuras de edificios, a partir de nubes de puntos LiDAR 3D e ortoimágenes, procedentes de vuelos fotogramétricos, aplicados al cálculo del potencial solar fotovoltaico, la metodología propuesta resolvió eficazmente la parametrización automática de tejados a nivel de ciudad, barrio y edificio. Esta parametrización resuelve para la gran mayoría de tipos de tejados existentes: tejados inclinados, tejados planos, a dos aguas y tejados piramidales. Además, la posibilidad del cálculo a nivel de ciudad permite la realización de estudios de prospectiva del análisis de posibilidades y establecimiento de normativas e incentivos por parte de las administraciones locales y autonómicas.

La metodología desarrollada, implica no solo la extracción automática de los bordes del tejado, sino también la parametrización de cada agua del tejado en términos de orientación azimutal, ángulo de inclinación y dimensiones efectivas, que son los parámetros necesarios para el cálculo de la capacidad fotovoltaica de los tejados. Hay que destacar que la metodología es sensible a cambios en la resolución de los datos de entrada, de manera que se obtienen importantes mejoras en la precisión con conjuntos de datos de mayor resolución. Específicamente, se han obtenido mejoras significativas en cuanto a la delimitación de las pendientes del tejado y el cálculo de sus dimensiones al comparar los resultados de ortoimágenes de mayor resolución espacial y datos LiDAR. Esto es debido a que una nube de puntos de mayor densidad proporciona una mejor definición de la geometría y, en el caso de las cubiertas de los edificios, permite una mejor diferenciación de cada agua y un mejor ajuste de planos. En cuanto al resto de parámetros, se obtienen resultados muy precisos en las medidas angulares tanto para datos aéreos de mayor y menor resolución espacial. Específicamente, se obtuvo un error promedio de 0,20% en la estimación de orientación azimutal, 1,29% en la estimación del ángulo de inclinación, 0,63% en la estimación de altitud y 2,93% en el área por estimación de pendiente del tejado.

La herramienta de mapas web Ener3dmap Solar-Web Roofs desarrollada aporta verticalidad a la investigación, completando todas las fases del tratamiento de los datos desde el procesado de los datos de origen, la incorporación de propiedades geométricas a los mismos, y el uso último de la herramienta por parte de los profesionales que realizan la instalación, o del usuario final. Esta aplicación, basada en Leaflet, es compatible con los principales formatos de datos geoespaciales. Además, integra datos geométricos de tejados en 3D con un modelo de radiación solar ampliamente validado, PVGIS. Proporciona un alto nivel de automatización, autocompletando los parámetros del tejado seleccionado en el formulario para el cálculo de la instalación. Permite la modificación de estos parámetros por parte del usuario, tanto para el diseño de la instalación solar fotovoltaica como para la introducción de datos para edificios proyectados, o para instalaciones en terrenos sin edificaciones existentes.

La integración de un modelo de radiación solar validado como PVGIS presenta un avance sobre las herramientas existentes en la frecuencia con la que se pueden calcular los resultados de la radiación solar: horaria, mensual, anual; y completando los parámetros geométricos de la cubierta automáticamente a partir de datos geomáticos.

Por todo lo anterior, se aporta una herramienta para el cálculo preciso pero sencillo de las instalaciones solares fotovoltaicas, facilitando la incorporación de esta energía renovable y limpia a la red eléctrica, promoviendo el autoconsumo y reduciendo la pobreza energética.

6.2. Líneas futuras

En el desarrollo de metodologías para la automatización del procesado de información, procedente de la digitalización de infraestructuras, realizado en esta Tesis Doctoral, se abren las siguientes líneas de trabajo para avanzar en la investigación. Con respecto a la seguridad vial y reducción de emisiones en las infraestructuras viarias:

- Incorporar el análisis de los parámetros verticales del trazado de las infraestructuras viales para completar el índice de riesgo potencial.
- Incorporar la visibilidad de parada, y la visibilidad de cruce al análisis del riesgo de cada tramo.
- Estudio de la electrificación de líneas de autobuses, a partir del estudio de la orografía de las rutas, para contribuir a la reducción de emisiones.
- Incorporación de mobiliario urbano, como marquesinas de autobuses, de soporte para paneles fotovoltaicos para producción de energía [44].

- Utilizar los datos MLS para inspeccionar las infraestructuras viales y detectar posibles deficiencias.
- Con las mejores resoluciones de datos procedentes de las nuevas pasadas de los vuelos fotogramétricos, del Instituto Geográfico Nacional, valorar si la aplicación de las metodologías descritas proporciona unos resultados aceptables para esos datos.

Con respecto a la aplicación en procesado de edificios, para fomentar el autoconsumo solar fotovoltaico:

- Mejorar la caracterización de las cubiertas incorporando los tejados a cuatro aguas.
- Diseñar procedimientos para detectar automáticamente las sombras proyectadas en los tejados, en base al valor de la altitud, haciendo un estudio de la línea de visión de cada edificio.
- Realizar el análisis del potencial fotovoltaico integrando el estudio del efecto de la geomorfología del suelo en la radiación solar incidente sobre las superficies fotovoltaicas.
- Mejorar el modelo de radiación solar utilizado en la herramienta web, para reducir la desviación del 2,5% de PVGIS, con la integración de datos meteorológicos históricos adicionales, o incorporando medidas de satélite (temperatura de la superficie terrestre, albedo, radiación solar en la atmósfera) en el modelo matemático [45].

7. Conclusions and future lines

This chapter includes the conclusions of the research carried out, and sets out the future lines that are open to continue advancing and complementing the research.

7.1. Conclusions

The main objective of this Doctoral Thesis, to achieve the automation of geospatial data processing of infrastructures and buildings, has been achieved. Similarly, the specific objectives have been achieved with the tools developed (Appendix B): ***inRoad inAlert*** for the categorization of road sections according to their risk based on the characteristics of the road alignment, ***Ener3DMap-SolarRoofs*** for the extraction of the characteristics of each roof slope, ***Ener3DMap-SolarWeb Roofs*** for the calculation of the photovoltaic solar production linking with the PVGIS API.

This research shows that it is possible to define strategies, methodologies and algorithms to automate the processing of data from 3D digitalization of infrastructures and buildings, to provide them with semantic properties that facilitate the handling of this information. In this Doctoral Thesis, automation solutions are provided for their use when the volume of information to be processed is very large. This size may be due to the fact that the spatial resolution is very good, if working with data from mobile mapping systems. With spatial resolutions not so high in the case of data from photogrammetric flights, the large volume of data to be processed may be due to the number of infrastructures or their size.

In the studies performed in this thesis, solutions have been obtained to facilitate the work of professionals from different sectors with geospatial information. In addition, the web tools developed allow the use of the processed information, in a simple and intuitive way, by end users, who do not have to be experts in the field.

The processing of point clouds obtained with MLS technology can provide risk assessment of a road infrastructure by means of: automatic risk mapping of the road based on geometric consistency indices and accurate stability criteria derived from these indices. This is achieved with the presented comprehensive approach, which combines geometry and risk, using 3D MLS point clouds. This approach represents a novel method of assessing road safety from MLS data and supported by an inductive reasoning process, which provides an assessment of potential risk based solely on geometric parameters intrinsic to the road alignment. This potential risk index can adequately complement other safety approaches, such as EuroRAP, which is based on a statistical-qualitative approach, sensitive to the distorting effect of traffic accidents.

All in all, a high degree of reliability has been reached in the extraction of the geometric elements of the horizontal alignment of the road. This is evident considering the low discrepancies obtained for the main geometric parameters. The minimal differences observed in the geometric consistency indices and in the stability criteria are also noteworthy.

In the data processing for the digitization of building infrastructures, from 3D LiDAR point clouds and orthoimages from photogrammetric flights, applied to the calculation of photovoltaic solar potential, the proposed methodology effectively solved the automatic parameterization of roofs at city, neighborhood and building levels. This parameterization solves for the vast majority of existing roof types: pitched roofs, flat roofs, gable roofs and pyramidal roofs. In addition, the possibility of application at city level allows the performance of prospective studies of the analysis of possibilities and the establishment of regulations and incentives by local and regional administrations.

The developed methodology involves not only the automatic extraction of roof edges, but also the parameterization of each roof slope in terms of azimuthal orientation, tilt angle and effective dimensions, which are the necessary parameters for the calculation of the photovoltaic capacity of the roofs. It should be noted that the methodology is sensitive to changes in the resolution of the input data, so that significant improvements in accuracy are obtained with higher resolution datasets. Specifically, significant improvements in terms of delineating roof slopes and calculating their dimensions have been obtained when comparing results from higher spatial resolution orthoimagery and LiDAR data. This is due to the fact that a higher density point cloud provides a better definition of the geometry and, in the case of building roofs, allows a better differentiation of each slope and a better adjustment of the planes. As for the rest of the parameters, very accurate results are obtained in the angular measurements for both higher and lower spatial resolution aerial data. Specifically, an average error of 0.20% was obtained in the estimation of azimuthal orientation, 1.29% in the estimation of tilt angle, 0.63% in the estimation of altitude and 2.93% in the area by estimation of roof slope.

The Ener3dmap Solar-Web Roofs web mapping tool developed brings verticality to the research, completing all the phases of the data treatment from the processing of the source data, the incorporation of their geometric properties, and the ultimate use of the tool by the professionals who perform the installation, or the end user. This application, based on Leaflet, is compatible with the main geospatial data formats. In addition, it integrates 3D roof geometry data with a widely validated solar radiation model, PVGIS. It provides a high level of automation, autocompleting the roof parameters selected in the form for installation calculation. It allows the modification of these

parameters by the user, both for the design of the solar photovoltaic installation and for the input of data for projected buildings, or for installations on land with no existing buildings.

The integration of a validated solar radiation model such as PVGIS presents an advance over existing tools in the frequency with which solar radiation results can be calculated: hourly, monthly, annual; and completing the geometric parameters of the roof automatically from geomatics data.

For all of the above reasons, a tool is provided for the precise but simple calculation of solar photovoltaic installations, facilitating the incorporation of this renewable and clean energy into the electricity grid, promoting self-consumption and reducing energy poverty.

7.2. Future lines

In the development of methodologies for the automation of information processing, coming from the digitalization of infrastructures, carried out in this Doctoral Thesis, the following lines of work are open to advance in the research. Regarding road safety and emission reduction in road infrastructures:

- Incorporation of the analysis of the vertical parameters of the road infrastructures to complete the potential risk index.
- Inclusion of stopping visibility and crossing visibility into the risk analysis of each section.
- Study of the electrification of bus lines, based on the study of the orography of the routes, to contribute to the reduction of emissions.
- Incorporation of street furniture, such as bus shelters, to support photovoltaic panels for energy production [44].
- Use of MLS data to inspect road infrastructure and detect possible deficiencies.
- With the best data resolutions coming from the new passes of the photogrammetric flights of the National Geographic Institute, assessment of whether the application of the described methodologies provides acceptable results for these data.

Regarding the developments on building processing, to promote photovoltaic solar self-consumption:

- Improvement of the characterization of roofs by incorporating hipped roofs.
- Design of procedures to automatically detect the shadows cast on roofs, based on the value of the altitude, making a study of the line of sight of each building.
- Analysis of photovoltaic potential by integrating the study of the effect of soil geomorphology on the incident solar radiation on photovoltaic surfaces.
- Improvement of the solar radiation model used in the web tool, to reduce the 2.5% deviation of PVGIS, with the integration of additional historical meteorological data, or by incorporating satellite measurements (land surface temperature, albedo, solar radiation in the atmosphere) in the mathematical model [45].

Referencias

8. Referencias

[1] Objetivos de Desarrollo Sostenible, Naciones Unidas. Accesible: https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/. Consultado: 02/11/2021

[2] Seguridad Vial, Visión Zero, Unión Europea. Accesible: https://ec.europa.eu/transport/road_safety/what-we-do_es. Consultado: 02/11/2021

[3] Bitenc, M., Lindenbergh, R., Khoshelham, K., van Waarden, A.P. Evaluation of a LIDAR Land-Based Mobile Mapping System for Monitoring Sandy Coasts. *Remote Sensing 3 (7)* (2011) 1472-1491.

[4] Gonzalez-Jorge, H., Puente, I., Riveiro, B., Martinez-Sanchez, J., Arias, P. Automatic segmentation of road overpasses and detection of mortar efflorescence using mobile LiDAR data. *Optics and Laser Technology 54* (2013) 353-361.

[5] Mc Elhinney, C., Kumar, P., Cahalane, C., McCarthy, T., 2010. Initial Results from European Road Safety Inspection (Eursi) Mobile Mapping Project. *Proceedings of the ISPRS Commission V Mid-Term Symposium Close Range Image Measurement Techniques 38* (2010) 440-445.

[6] Puente, I., Gonzalez-Jorge, H., Martinez-Sanchez, J., Arias, P. *Review of mobile mapping and surveying technologies. Measurement 46 (7)* (2013) 2127-2145.

[7] M. C. Brito, S. Freitas, S. Guimarães, C: Catita, P. Redweik. The importance of facades for the solar PV potential of a Mediterranean city using LiDAR data. *Renewable Energy 111* (2017) 85-94. DOI: 10.1016/j.renene.2017.03.085

[8] A. M. Martín, J. Domínguez, J. Amador. Applying LIDAR datasets and GIS based model to evaluate solar potential over roofs: a review. *AIMS Energy 3.3* (2015) 326-343. DOI: 10.3934/energy.2015.3.326

[9] Directive, I. N. S. P. I. R. E. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). *Published in the official Journal on the 25th April; 2007*. (2007)

[10] Centro de descargas – Instituto Geográfico Nacional de España: http://centrodedescargas.cnig.es/CentroDescargas/index.jsp. Consultado: 02/11/2021

[11] Geoportal de Inspire: https://inspire-geoportal.ec.europa.eu/results.html. Consultado: 02/11/2021

[12] Soille, P., Burger, A., De Marchi, D., Kempeneers, P., Rodriguez, D., Syrris, V., Vasilev, V. A versatile data-intensive computing platform for information retrieval from big geospatial data, *Future Generation Computer Systems 81*, (2018) 30-40.

[13] Camacho-Torregrosa, F.J., Pérez-Zuriaga, A.M., Manuel Campoy-Ungria, J., Garcia-Garcia, A. New geometric design consistency model based on operating speed profiles for road safety evaluation. *Accident Analysis and Prevention 61*, (2013) 33-42.

[14] da Costa, J.O., Prudencio Jacques, M.A., Cunha Soares, F.E., Freitas, E.F. Integration of geometric consistency contributory factors in three-leg junctions collision prediction models of Portuguese two-lane national highways. *Accident Analysis and Prevention 86*, (2016) 59-67.

[15] Siskind, V., Steinhardt, D., Sheehan, M., O'Connor, T., Hanks, H. Risk factors for fatal crashes in rural Australia. *Accident Analysis and Prevention 43 (3)*, (2011) 1082-1088.

[16] Garach, L., de Ona, J., Lopez, G., Baena, L. Development of safety performance functions for Spanish two-lane rural highways on flat terrain. *Accident Analysis and Prevention 95*, (2016) 250-265.

[17] López, G., de Oña, J., Garach, L., Baena, L. Influence of deficiencies in traffic control devices in crashes on two-lane rural roads. *Accident Analysis and Prevention 96*, (2016) 130-139.

[18] EuroRAP. Accesible: https://eurorap.org/, 2020. Consultado: 02/11/2021.

[19] Holgado-Barco, A., Gonzalez-Aguilera, D., Arias-Sanchez, P., Martinez-Sanchez, J. Semiautomatic Extraction of Road Horizontal Alignment from a Mobile LiDAR System. *Computer-Aided Civil and Infrastructure Engineering 30 (3)*, (2015) 217-228.

[20] Varela-Gonzalez, M., Gonzalez-Jorge, H., Riveiro, B., Arias, P. Automatic filtering of vehicles from mobile LiDAR datasets. *Measurement 53*, (2014) 215-223.

[21] Castro, M., Lopez-Cuervo, S., Parens-Gonzalez, M., de Santos-Berbel, C. LIDAR-based roadway and roadside modelling for sight distance studies. *Survey Review 48 (350)*, (2016) 309-315

[22] Riveiro, B., Gonzalez-Jorge, H., Martinez-Sanchez, J., Diaz-Vilarino, L., Arias, P. Automatic detection of zebra crossings from mobile LiDAR data. *Optics and Laser Technology 70*, (2015) 63-70.

[23] Ai, C., Tsai, Y.J. Critical Assessment of an Enhanced Traffic Sign Detection Method Using Mobile LiDAR and INS Technologies. *Journal of Transportation Engineering 141 (5)*, (2015) 04014096.

[24] Cabo, C., Kukko, A., Garcia-Cortes, S., Kaartinen, H., Hyyppa, J., Ordonez, C. An Algorithm for Automatic Road Asphalt Edge Delineation from Mobile Laser Scanner Data Using the Line Clouds Concept. *Remote Sensing 8 (9)*, (2016) 740.

[25] Yang, B., Fang, L., Li, J. Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing 79*, (2013) 80-93.

[26] Marinelli, G., Bassani, M., Piras, M., Lingua, A.M., 2017. Mobile mapping systems and spatial data collection strategies assessment in the identification of horizontal alignment of highways. *Transportation Research Part C-Emerging Technologies 79*, (2017) 257-273.

[27] Kumar, P., McElhinney, C.P., Lewis, P., McCarthy, T. Automated road markings extraction from mobile laser scanning data. *International Journal of Applied Earth Observation and Geoinformation 32*, (2014) 125-137.

[28] Kumar, P., McElhinney, C.P., Lewis, P., McCarthy, T. An automated algorithm for extracting road edges from terrestrial mobile LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing 85*, (2013) 44-55.

[29] Montella, A., Imbriani, L.L. Safety performance functions incorporating design consistency variables. *Accident Analysis and Prevention 74*, (2015) 133-144.

[30] Gargoum, S., El-Basyouny, K. Automated Extraction of Road Features using LiDAR Data: A Review of LiDAR applications in transportation (Contributed Paper). In: *4th International Conference on Transportation Information and Safety (ICTIS). 8–10 August 2017 Banff, Alberta, Canada*, (2017).

[31] Eurostat (2017). Community Innovation Survey of Renewable energy statistics of the European Communities. Office for official publications of the European communities. Disponible: https://ec.europa.eu/eurostat/statistics-explained. Consultado: 02/11/2021

[32] Spanish Official State Bulletin, BOE. BOE-A-2018-13593 about urgent measures for energy transition and consumer protection. Disponible: https://boe.es/eli/es/rdl/2018/10/05/15. Consultado: 02/11/2021

[33] S. Castellanos, D. A. Sunter, D. M. Kammen, Rooftop solar photovoltaic potential in cities: how scalable are assessment approaches? *Environmental Research Letters 12(12)* (2017) 125005. DOI: 10.1088/1748-9326/aa7857

[34] Mohajeri N, Perera ATD, Coccolo S, Mosca L, Le Guen M, Scartezzini JL. Integrating urban form and distributed energy systems: Assessment of sustainable development scenarios for a Swiss village to 2050. *Renew Energy, 143*, (2019) 810-826.

[35] M. Fionnuala, K. McDonnell, A feasibility assessment of photovoltaic power systems in Ireland; a case study for the Dublin region. *Sustainability 9(2)* (2017) 302. DOI: 10.3390/su9020302

[36] Google´s Project Sunroof. Disponible: https://google.com/get/sunroof. Consultado: 02/11/2021

[37] Sánchez-Aparicio M, Del Pozo S, Martín-Jiménez JA, González-González E, Andrés-Anaya P, Lagüela S. Influence of LiDAR Point Cloud Density in the Geometric Characterization of Rooftops for Solar Photovoltaic Studies in Cities. *Remote Sensing. 12(22)*, (2020) 3726. DOI:10.3390/rs12223726

[38] D. Palmer, E. Koubli, I. Cole, R. Gottschalg, T. Betts, A GIS-based method for identification of wide area rooftop suitability for minimum size PV systems using LiDAR data and photogrammetry. *Energies 11(12)* (2018) 3506. DOI: 10.3390/en11123506

[39] S. Schuffert, T. Voegtle, N. Tate, A. Ramírez. Quality assessment of roof planes extracted from height data for solar energy systems by the EAGLE platform, *Remote Sensing 7(12)* (2015) 17016-17034. DOI: 10.3390/rs71215866

[40] Y. Li, C. Liu. Estimating solar energy potentials on pitched roofs, *Energy and Buildings 139* (2017) 101-107. DOI: 10.1016/j.enbuild.2016.12.070

[41] J. Sarralde, D. J. Quinn, D. Wiesmann, K. Steemers. Solar energy and urban morphology: Scenarios for increasing the renewable energy potential of neighbourhoods in London. *Renewable energy 73* (2015) 10-17. DOI: 10.1016/j.renene.2014.06.028

[42] S. Kiti, V. Wang, B. Sharp. Rooftop solar potential based on LiDAR data: Bottom-up assessment at neighbourhood level. *Renewable Energy 111* (2017) 463-475. DOI: 10.1016/j.renene.2017.04.025

[43] T. Schenk, B. Csathó. Fusion of LIDAR data and aerial imagery for a more complete surface description. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences 34(3/A)* (2002) 310-317. DOI: 10.1109/DFUA.2003.1219962

[44] Sánchez-Aparicio M., Lagüela S., Martín-Jiménez J., Del Pozo S., González-González E., Andrés-Anaya P. Smart Mobility in Cities: GIS Analysis of Solar PV Potential for Lighting in Bus Shelters in the City of Ávila. In: *Nesmachnow S., Hernández Callejo L. (eds) Smart Cities. ICSC-CITIES 2020. Communications in Computer and Information Science, vol 1359. Springer, Cham.* (2021) DOI:10.1007/978-3-030-69136-3_11

[45] Sánchez-Aparicio, María, de Andres Anaya, Paula, Del Pozo, Susana, Lagüela, Susana. Retrieving Land Surface Temperature from Satellite Imagery with a Novel Combined Strategy. *Remote Sensing. 12. 277.* (2020) DOI:10.3390/rs12020277.

APÉNDICE A

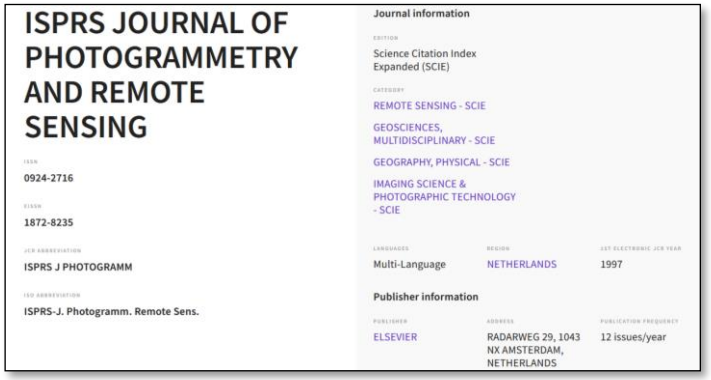
Indexación y factor de impacto de las revistas

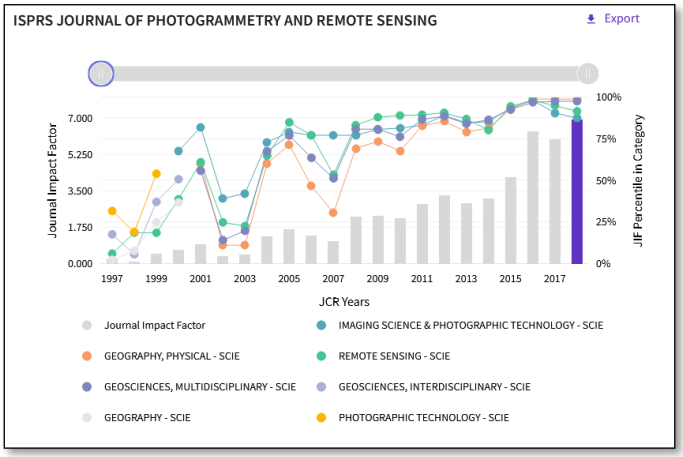
Apéndice A: Indexación y factor de impacto de las revistas

Publicación 1: *Road safety evaluation through automatic extraction of road horizontal alignments from Mobile LiDAR System and inductive reasoning based on decision tree.*

La publicación se encuentra en el cuartil Q1. Se muestra una tabla resumen y debajo unas imágenes con la captura de pantalla de los principales parámetros de la publicación.

Revista:	**ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING**
Editorial	**ELSEVIER**
ISSN:	**0924-2716**
eISSN:	**1872-8235**
Factor de Impacto (2018):	**6.942**
Ranking:	**1/50**
Cuartil:	**Q1**





Rank by Journal Impact Factor

Journals within a category are sorted in descending order by Journal Impact Factor (JIF) resulting in the Category Ranking below. A separate rank is shown for each category in which the journal is listed in JCR. Data for the most recent year is presented at the top of the list, with other years shown in reverse chronological order. Learn more

EDITION
Science Citation Index Expanded (SCIE)

CATEGORY
GEOGRAPHY, PHYSICAL
1/50

JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	1/50	Q1	99.00
2019	1/50	Q1	99.00
2018	1/50	Q1	99.00
2017	1/49	Q1	98.98
2018	1/50	Q1	99.00

EDITION
Science Citation Index Expanded (SCIE)

CATEGORY
GEOSCIENCES, MULTIDISCIPLINARY
5/196

JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	5/200	Q1	97.75
2019	5/200	Q1	97.75
2018	5/196	Q1	97.70
2017	5/190	Q1	97.63
2018	5/196	Q1	97.70

Rank by Journal Impact Factor

Journals within a category are sorted in descending order by Journal Impact Factor (JIF) resulting in the Category Ranking below. A separate rank is shown for each category in which the journal is listed in JCR. Data for the most recent year is presented at the top of the list, with other years shown in reverse chronological order. Learn more

EDITION
Science Citation Index Expanded (SCIE)

CATEGORY
IMAGING SCIENCE & PHOTOGRAPHIC TECHNOLOGY
4/28

JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	3/29	Q1	91.38
2019	3/27	Q1	90.74
2018	4/28	Q1	87.50
2017	3/27	Q1	90.74
2018	4/28	Q1	87.50

EDITION
Science Citation Index Expanded (SCIE)

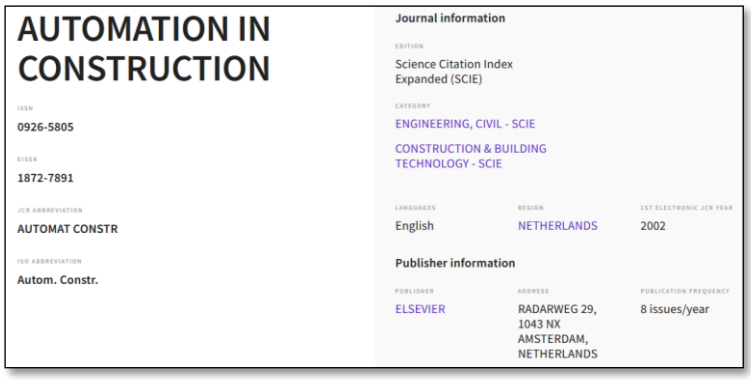
CATEGORY
REMOTE SENSING
3/30

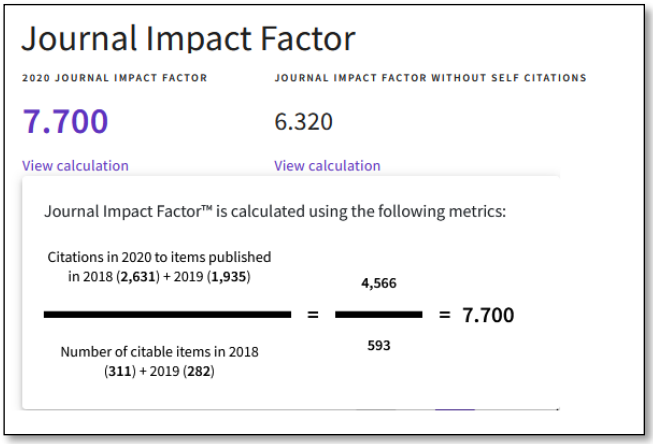
JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	2/32	Q1	95.31
2019	3/30	Q1	91.67
2018	3/30	Q1	91.67
2017	2/30	Q1	95.00
2018	3/30	Q1	91.67

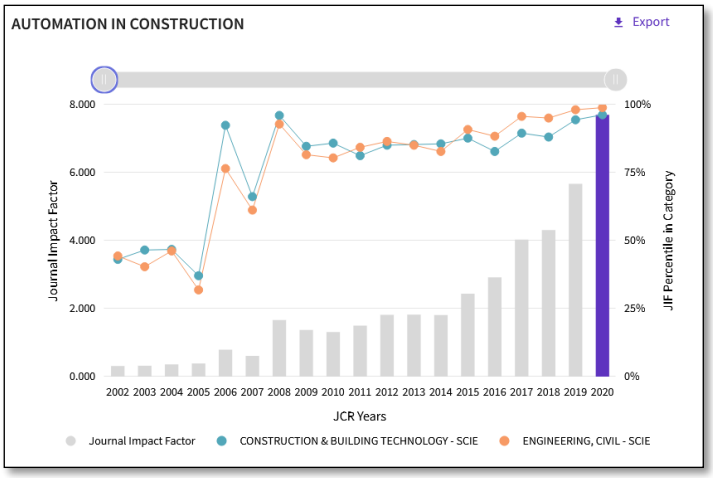
Publicación 2: *Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity.*

La publicación "Automation in Construction" se encuentra en el cuartil Q1. Se muestra una tabla resumen y debajo unas imágenes con la captura de pantalla de los principales parámetros de la publicación.

Revista:	**AUTOMATION IN CONSTRUCTION**
Editorial	**ELSEVIER**
ISSN:	**0926-5805**
eISSN:	**1872-7891**
Factor de Impacto (2020):	**7.700**
Ranking:	**2/137**
Cuartil:	**Q1**







Rank by Journal Impact Factor

Journals within a category are sorted in descending order by Journal Impact Factor (JIF) resulting in the Category Ranking below. A separate rank is shown for each category in which the journal is listed in JCR. Data for the most recent year is presented at the top of the list, with other years shown in reverse chronological order. Learn more

EDITION
Science Citation Index Expanded (SCIE)

CATEGORY
CONSTRUCTION & BUILDING TECHNOLOGY
3/67

JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	3/67	Q1	96.27
2019	4/63	Q1	94.44
2018	8/63	Q1	88.10
2017	7/62	Q1	89.52
2016	11/61	Q1	82.79

EDITION
Science Citation Index Expanded (SCIE)

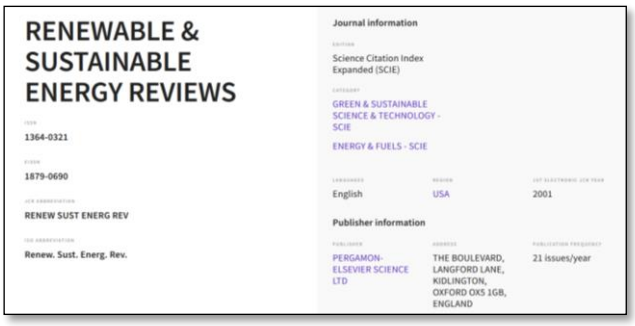
CATEGORY
ENGINEERING, CIVIL
2/137

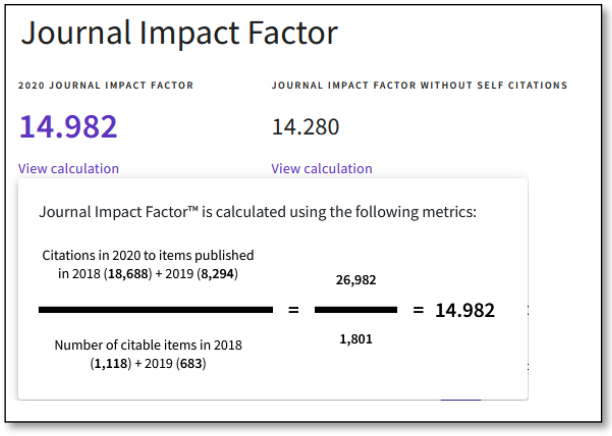
JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	2/137	Q1	98.91
2019	3/134	Q1	98.13
2018	7/132	Q1	95.08
2017	6/128	Q1	95.70
2016	15/125	Q1	88.40

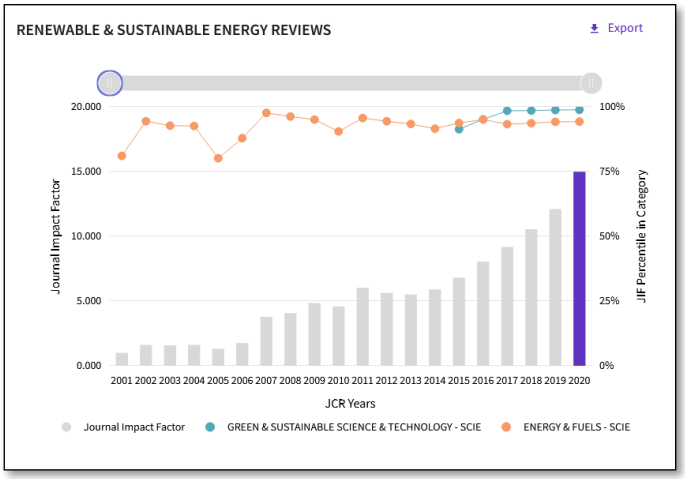
Publicación 3: *A geospatial web-based platform for the photovoltaic potential computation: Ener3DMap-SolarWeb Roofs.*

La publicación "Renewable & Sustainable Energy Reviews" se encuentra en el cuartil Q1. Se muestra una tabla resumen y debajo unas imágenes con la captura de pantalla de los principales parámetros de la publicación.

Revista:	**RENEWABLE & SUSTAINABLE ENERGY REVIEWS**
Editorial	**PERGAMON-ELSEVIER SCIENCE LTD**
ISSN:	**1364-0321**
Factor de Impacto (2020):	**14.982**
Ranking:	**1/44**
Cuartil:	**Q1**







Rank by Journal Impact Factor

Journals within a category are sorted in descending order by Journal Impact Factor (JIF) resulting in the Category Ranking below. A separate rank is shown for each category in which the journal is listed in JCR. Data for the most recent year is presented at the top of the list, with other years shown in reverse chronological order. Learn more

EDITION
Science Citation Index Expanded (SCIE)

CATEGORY
ENERGY & FUELS

7/114

JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	7/114	Q1	94.30
2019	7/112	Q1	94.20
2018	7/103	Q1	93.69
2017	7/97	Q1	93.30
2016	5/92	Q1	95.11

EDITION
Science Citation Index Expanded (SCIE)

CATEGORY
GREEN & SUSTAINABLE SCIENCE & TECHNOLOGY

1/44

JCR YEAR	JIF RANK	JIF QUARTILE	JIF PERCENTILE
2020	1/44	Q1	98.86
2019	1/41	Q1	98.78
2018	1/35	Q1	98.57
2017	1/33	Q1	98.48
2016	2/31	Q1	95.16

APÉNDICE B

Software desarrollado

Apéndice B. Software Desarrollado

Inroad inAlert

Tipo: Programa de ordenador
Universidad: Universidad de Salamanca-MovilData

- **Breve descripción:**

Inroad inAlert es una aplicación de escritorio que recibe como datos de entrada nubes de puntos de infraestructuras de viales: calles o carreteras, y los procesa extrayendo el eje. Realiza el análisis del eje para identificar tramos curvos, tramos rectos y clotoides. Calcula los parámetros de estabilidad de cada tramo y asigna un índice de peligrosidad según las características de los mismos.

- **Lenguaje de programación:**

Utiliza como entorno de desarrollo ***QT creator***, se usa el lenguaje de programación ***C++*** para optimizar el procesamiento de nubes de puntos se utiliza la librería ***PCL (Point Cloud Library)***.

- **Entorno operativo:**

Es una aplicación para ordenadores pudiéndose compilar para distintos sistemas operativos como Windows y Linux.

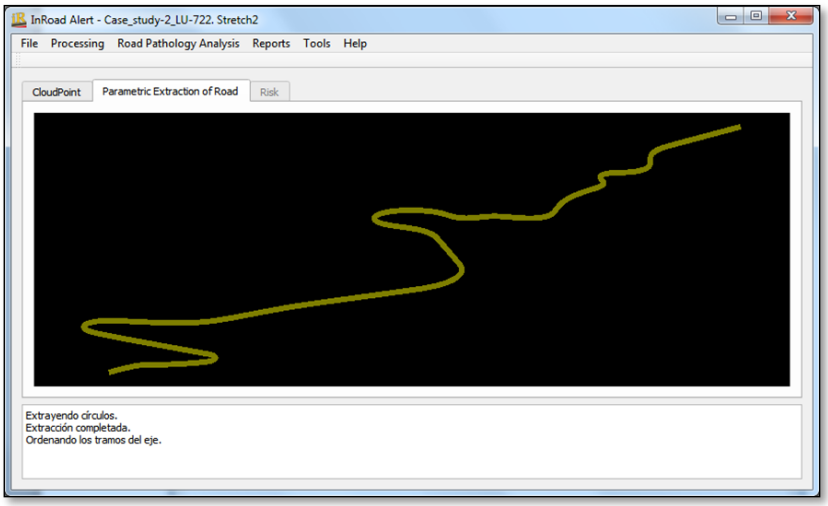


Figura B.1: Interfaz de la aplicación.

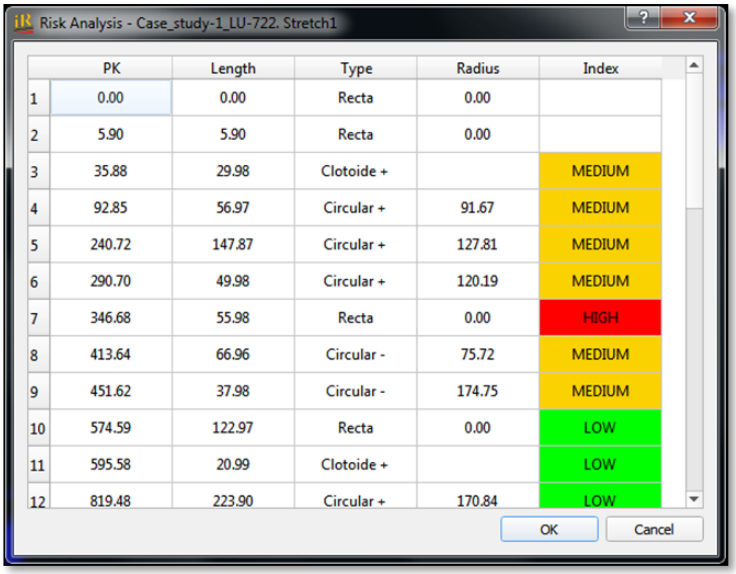
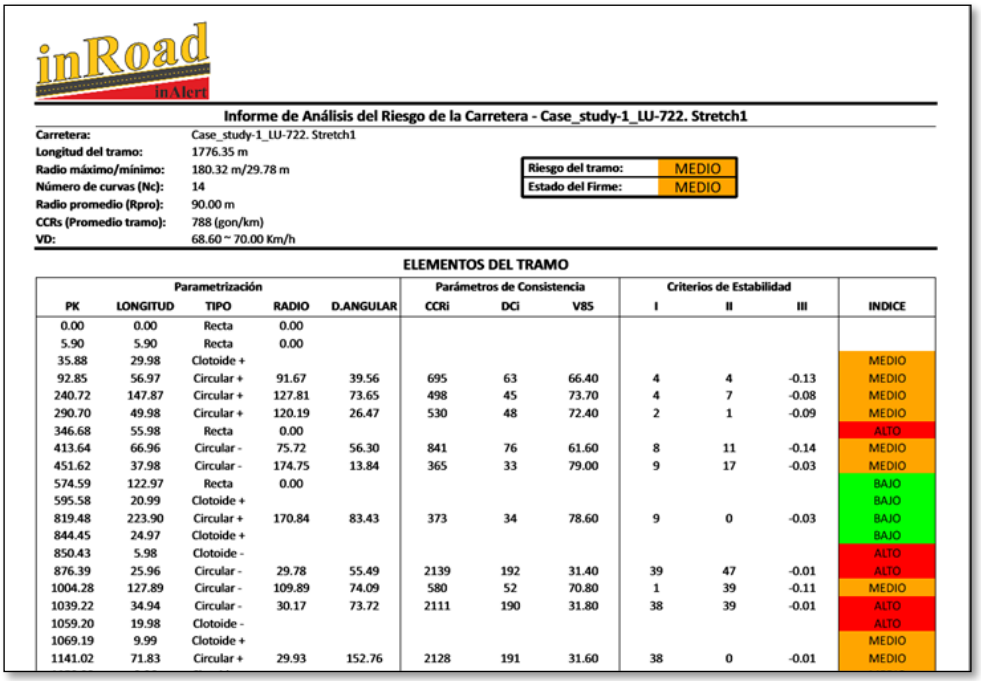


Figura B.2: Salida de datos en pantalla.



inRoad inAlert

Informe de Análisis del Riesgo de la Carretera - Case_study-1_LU-722. Stretch1

Carretera:	Case_study-1_LU-722. Stretch1
Longitud del tramo:	1776.35 m
Radio máximo/mínimo:	180.32 m/29.78 m
Número de curvas (Nc):	14
Radio promedio (Rpro):	90.00 m
CCRs (Promedio tramo):	788 (gon/km)
VD:	68.60 ~ 70.00 Km/h

Riesgo del tramo:	MEDIO
Estado del Firme:	MEDIO

ELEMENTOS DEL TRAMO

Parametrización					Parámetros de Consistencia			Criterios de Estabilidad			
PK	LONGITUD	TIPO	RADIO	D.ANGULAR	CCRi	DCi	V85	I	II	III	INDICE
0.00	0.00	Recta	0.00								
5.90	5.90	Recta	0.00								
35.88	29.98	Clotoide +									MEDIO
92.85	56.97	Circular +	91.67	39.56	695	63	66.40	4	4	-0.13	MEDIO
240.72	147.87	Circular +	127.81	73.65	498	45	73.70	4	7	-0.08	MEDIO
290.70	49.98	Circular +	120.19	26.47	530	48	72.40	2	1	-0.09	MEDIO
346.68	55.98	Recta	0.00								ALTO
413.64	66.96	Circular -	75.72	56.30	841	76	61.60	8	11	-0.14	MEDIO
451.62	37.98	Circular -	174.75	13.84	365	33	79.00	9	17	-0.03	MEDIO
574.59	122.97	Recta	0.00								BAJO
595.58	20.99	Clotoide +									BAJO
819.48	223.90	Circular +	170.84	83.43	373	34	78.60	9	0	-0.03	BAJO
844.45	24.97	Clotoide +									BAJO
850.43	5.98	Clotoide -									ALTO
876.39	25.96	Circular -	29.78	55.49	2139	192	31.40	39	47	-0.01	ALTO
1004.28	127.89	Circular -	109.89	74.09	580	52	70.80	1	39	-0.11	MEDIO
1039.22	34.94	Circular -	30.17	73.72	2111	190	31.80	38	39	-0.01	ALTO
1059.20	19.98	Clotoide -									ALTO
1069.19	9.99	Clotoide +									MEDIO
1141.02	71.83	Circular +	29.93	152.76	2128	191	31.60	38	0	-0.01	MEDIO

Figura B.3: Captura de parte del informe generado.

Autores:

- José Antonio Martín Jiménez
- Santiago Zazo del Dedo
- Pablo Rodriguez Gonzálvez
- Diego González Aguilera
- José Juan Arranz Justel

Ener3DMap-Solar Roofs

Tipo: Programa de ordenador

Universidad: Universidad de Salamanca

- **Breve descripción:**

SolarRoofs es una aplicación de escritorio desarrollada dentro de la Cátedra de Iberdrola VIII Centenario, proyecto de investigación bautizado como ENER3DMAP. Se trata de un programa que recibe como datos de entrada las nubes de puntos de infraestructuras: Edificios, los procesa y obtiene la geometría de los tejados para calcular el potencial solar de las instalaciones que se podrían realizar sobre ellos.

- **Lenguaje de programación:**

Utiliza como entorno de desarrollo ***QT creator***, se usa el lenguaje de programación ***C++*** para optimizar el procesamiento de nubes de puntos se utiliza la librería ***PCL (Point Cloud Library)***.

- **Entorno operativo:**

Es una aplicación para ordenadores pudiéndose compilar para distintos sistemas operativos como Windows y Linux.

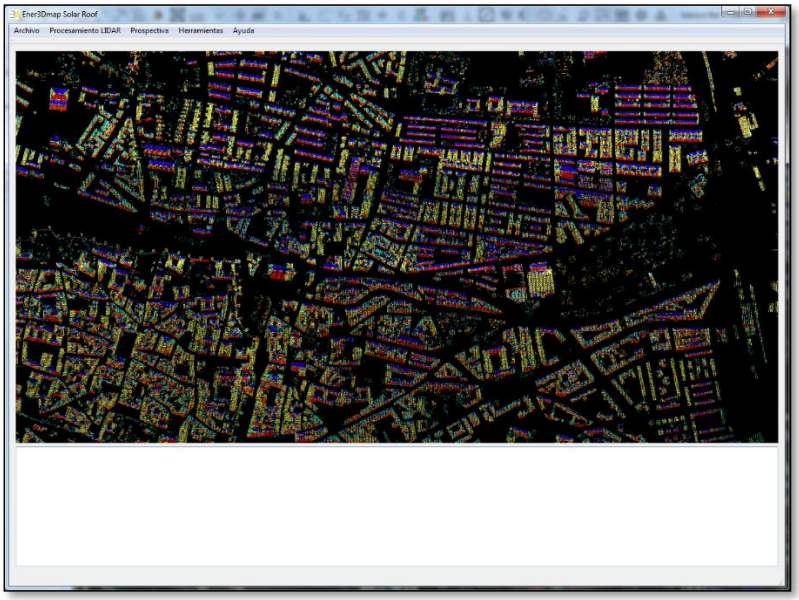


Figura B.4: Interfaz de la aplicación.

Autores:

- José Antonio Martín Jiménez
- Susana del Pozo Aguilera
- María Sánchez Aparicio
- Susana Lagüela López

Ener3DMap-SolarWeb Roofs

Tipo: Registro de propiedad intelectual

Objeto de la propiedad intelectual: Programa de ordenador

Referencia: SA-112-19

Asiento registral: 00/2019/3965

Universidad: Universidad de Salamanca

- **<u>Breve descripción:</u>**

SolarWeb es una Infraestructura de Datos Espaciales (IDE) desarrollada dentro de la Cátedra de Iberdrola VIII Centenario, proyecto de investigación bautizado como ENER3DMAP. Se trata de una herramienta web que permite a cualquier usuario conocer el potencial solar diario, mensual y anual de un edificio, así como el dimensionamiento y la configuración de la instalación fotovoltaica con tan solo hacer clic sobre el tejado del edificio de interés incluso sobre edificios proyectados (aun no existentes). Dicha herramienta está formada por una gran variedad de datos gratuitos de diferente naturaleza: desde datos geoespaciales como por ejemplo imágenes capturadas por aviones o por satélites hasta datos de radiación solar procedentes de una base de datos con un gran potencial como son PVGIS y PVWATTS. Además, utiliza datos de catastro y datos de predicción meteorológica.

Aunque ya existen herramientas similares, ***SolarWeb*** destaca por su elevado grado de automatización para la caracterización geométrica tridimensional de los edificios a partir de datos geomáticos gratuitos procedentes del Instituto Geográfico Nacional (IGN). En concreto, gracias a la combinación de imágenes aéreas con buena resolución y nube de puntos aéreas se obtiene de una manera precisa y automática los parámetros geométricos de interés para el cálculo del potencial solar en tejados: inclinación, orientación y superficie; sin necesidad de que el usuario los introduzca.

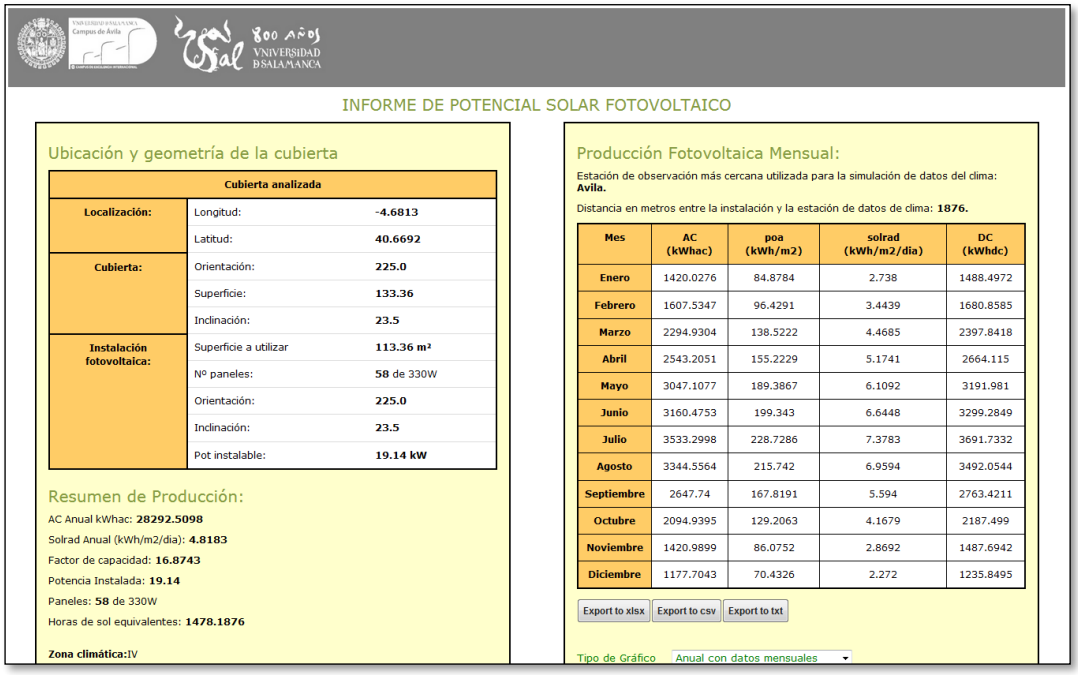
- **Lenguaje de programación:**

Utiliza un entorno web con ***HTML***, ***CSS*** y ***JavaScript*** para el lado del cliente con uso de ***Leaflet*** como API para interfaz de mapas. En la parte de programación del lado del servidor utiliza lenguaje ***PHP***.

- **Entorno operativo:**

Es una aplicación web para ordenadores y dispositivos móviles que tengan acceso a internet.

Figura B.5: Interfaz de la aplicación web.



INFORME DE POTENCIAL SOLAR FOTOVOLTAICO

Ubicación y geometría de la cubierta

Cubierta analizada		
Localización:	Longitud:	**-4.6813**
	Latitud:	**40.6692**
Cubierta:	Orientación:	**225.0**
	Superficie:	**133.36**
	Inclinación:	**23.5**
Instalación fotovoltaica:	Superficie a utilizar	**113.36 m²**
	Nº paneles:	**58** de 330W
	Orientación:	**225.0**
	Inclinación:	**23.5**
	Pot instalable:	**19.14 kW**

Resumen de Producción:

AC Anual kWhac: **28292.5098**
Solrad Anual (kWh/m2/dia): **4.8183**
Factor de capacidad: **16.8743**
Potencia Instalada: **19.14**
Paneles: **58** de 330W
Horas de sol equivalentes: **1478.1876**

Zona climática:IV

Producción Fotovoltaica Mensual:

Estación de observación más cercana utilizada para la simulación de datos del clima: **Avila.**

Distancia en metros entre la instalación y la estación de datos de clima: **1876.**

Mes	AC (kWhac)	poa (kWh/m2)	solrad (kWh/m2/dia)	DC (kWhdc)
Enero	1420.0276	84.8784	2.738	1488.4972
Febrero	1607.5347	96.4291	3.4439	1680.8585
Marzo	2294.9304	138.5222	4.4685	2397.8418
Abril	2543.2051	155.2229	5.1741	2664.115
Mayo	3047.1077	189.3867	6.1092	3191.981
Junio	3160.4753	199.343	6.6448	3299.2849
Julio	3533.2998	228.7286	7.3783	3691.7332
Agosto	3344.5564	215.742	6.9594	3492.0544
Septiembre	2647.74	167.8191	5.594	2763.4211
Octubre	2094.9395	129.2063	4.1679	2187.499
Noviembre	1420.9899	86.0752	2.8692	1487.6942
Diciembre	1177.7043	70.4326	2.272	1235.8495

Export to xlsx | Export to csv | Export to txt

Tipo de Gráfico: Anual con datos mensuales

Figura B.6: Parte de los resultados de producción generados.

Autores:

- José Antonio Martín Jiménez
- Susana del Pozo Aguilera
- María Sánchez Aparicio
- Susana Lagüela López
- Enrique González González

Ener3DMap-SolarWeb Cities

Tipo: Registro de propiedad intelectual

Objeto de la propiedad intelectual: Programa de ordenador

Referencia: SA-8-20

Asiento registral: 00/2020/645

Universidad: Universidad de Salamanca

- **Breve descripción:**

Solarweb Cities es una Infraestructura de Datos Espaciales (IDE), que se desarrolla para dar soluciones a las necesidades que plantea Iberdrola Clientes, dentro de la Cátedra de Iberdrola VIII Centenario. En la necesidad de promover las instalaciones de energía solar fotovoltaica, se diseña esta herramienta web, que permite evaluar la prospectiva del potencial solar de una zona de edificios, con el fin de priorizar las instalaciones, en los barrios con características más favorables. La herramienta utiliza los datos de orientaciones, inclinaciones y superficies, calculados para cada parcela catastral, con el fin de calcular el potencial solar mediante peticiones a los modelos de radiación solar PVGIS y PVWATTS.

SolarWeb Cities permite seleccionar las orientaciones objetivo de estudio y calcular el potencial solar de ciudades o barrios. Además, proporciona una tabla de resultados con superficies, potencial solar y producción anual, para cada combinación de orientaciones e inclinaciones, considerando un intervalo de 10 grados en las orientaciones, y de 5 grados en las inclinaciones.

- **Lenguaje de programación:**

Utiliza un entorno web con ***HTML***, ***CSS*** y ***JavaScript*** para el lado del cliente con uso de ***Leaflet*** como API para interfaz de mapas. En la parte de programación del lado del servidor utiliza lenguaje ***PHP***.

- **Entorno operativo:**

Es una aplicación web para ordenadores y dispositivos móviles que tengan acceso a internet.

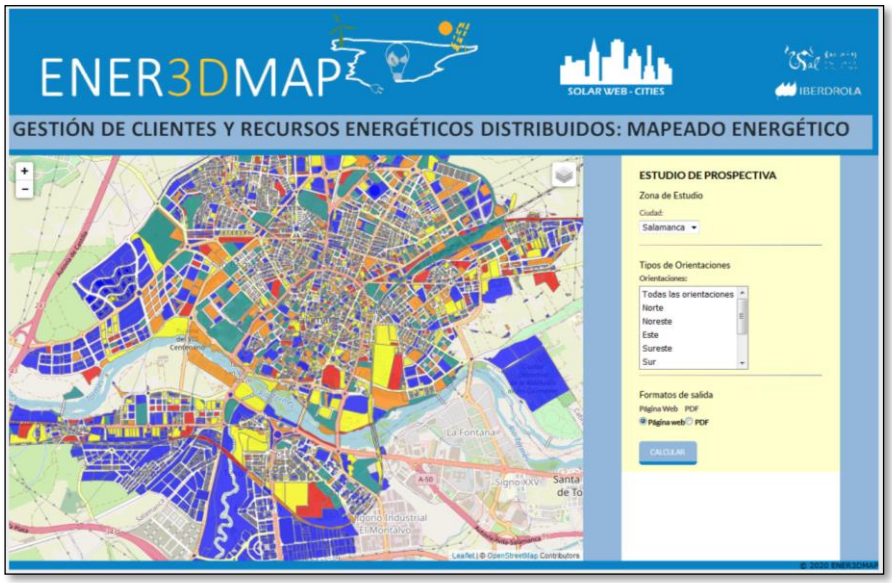


Figura B.7: Interfaz de la aplicación web.

Orientación	Área (m²)	Potencia (kWp)	Producción Anual(kWh)
Norte	211.992,00	35.963,40	39.903.132,00
Noreste	140.540,00	23.842,17	28.496.940,00
Este	188.060,00	31.905,39	42.080.372,40
Sureste	151.168,00	25.644,96	36.090.756,00
Sur	211.288,00	35.848,89	51.483.888,00
Suroeste	152.008,00	25.788,84	35.846.472,00
Oeste	198.212,00	33.627,66	42.957.508,80
Noroeste	157.856,00	26.781,48	30.996.360,00
Totales	1.411.124,00	239.402,79	307.855.429,20

Figura B.8: Parte de los resultados de prospectiva generados.

Autores:

- José Antonio Martín Jiménez
- Susana del Pozo Aguilera
- María Sánchez Aparicio
- Enrique González González
- Susana Lagüela López