

# Construcción automatizada de modelos 3D de árboles, a partir de datos LiDAR

## Automated construction of 3D tree models based on LiDAR data

Gizela Andrea Guzmán <sup>1</sup>,  
01 de diciembre de 2025

### Resumen

El presente trabajo aborda la construcción automatizada de modelos 3D de árboles a partir de datos LiDAR, con el propósito de analizar su geometría y estructura mediante diferentes Niveles de Detalle (LOD). La metodología implementada integra procesos de clasificación, segmentación, depuración y modelado geométrico dentro de un flujo de trabajo reproducible y orientado a datos reales. El procedimiento inicia con la clasificación preliminar de la nube LiDAR, a partir de la cual se aíslan los puntos correspondientes a vegetación. Sobre este subconjunto filtrado se aplica una segmentación espacial diseñada para separar árboles individuales dentro de la masa vegetada. Dado que esta etapa puede introducir errores por sobresegmentación, copas fragmentadas o ruido inherente al LiDAR, se implementa una fase de depuración y limpieza, que elimina puntos atípicos, planos residuales y segmentos no arbóreos mediante criterios morfométricos y verificaciones geométricas. Los segmentos validados sirven como base para la reconstrucción de modelos 3D, generados en distintos niveles de detalle. Los LOD inferiores (LOD0–LOD1) permiten representar la distribución espacial y el volumen general de cada árbol mediante formas simples; mientras que los LOD intermedios (LOD2) incorporan parámetros como altura, base de la copa y radios, produciendo geometrías más cercanas a la estructura real. Finalmente, el LOD3 emplea directamente la nube de puntos para construir volúmenes mediante envolventes convexas. En conjunto, los resultados obtenidos muestran que la metodología es capaz de producir modelos 3D coherentes en múltiples niveles de detalle, aunque su fidelidad geométrica depende de manera crítica de la densidad, continuidad y calidad de la nube de puntos.

**Palabras clave:** LiDAR, modelos 3D, arboles, LOD

Código disponible en: [https://github.com/gizeandre13/LISAR\\_Informe\\_2\\_Gizela\\_Guzman](https://github.com/gizeandre13/LISAR_Informe_2_Gizela_Guzman)

### Abstract

This paper addresses the automated construction of 3D tree models from LiDAR data, with the aim of analyzing their geometry and structure using different Levels of Detail (LOD). The methodology implemented integrates classification, segmentation, cleaning, and geometric modeling processes within a reproducible workflow oriented towards real data. The procedure begins with the preliminary classification of the LiDAR cloud, from which the points corresponding to vegetation are isolated. A spatial segmentation designed to separate individual trees within the vegetated mass is applied to this filtered subset. Since this stage can introduce errors due to oversegmentation, fragmented crowns, or noise inherent to LiDAR, a refinement and cleaning phase is implemented, which eliminates outliers, residual planes, and non-tree segments using morphometric criteria and geometric checks. The validated segments serve as the basis for the reconstruction of 3D models, generated at different levels of detail. Lower LODs (LOD0–LOD1)

allow the spatial distribution and overall volume of each tree to be represented using simple shapes, while intermediate LODs (LOD2) incorporate parameters such as height, crown base, and radii, producing geometries closer to the actual structure. Finally, LOD3 directly uses the point cloud to construct volumes using convex hulls.

**Keywords:** LiDAR, 3D models, trees, LOD

## 1 Introducción

El desarrollo urbano contemporáneo exige soluciones innovadoras capaces de abordar y gestionar la compleja dinámica de las ciudades actuales. En este contexto, el Gemelo Digital Urbano (Urban Digital Twin, UDT) se ha consolidado como una herramienta tecnológica clave, al permitir la representación, simulación y análisis de escenarios urbanos en entornos digitales altamente detallados. Los UDT facilitan la comprensión integral del funcionamiento de las ciudades, apoyan la evaluación del impacto de diferentes estrategias de gestión y fortalecen la toma de decisiones orientadas hacia una planificación más eficiente, resiliente y sostenible (Luo et al., 2025)

Dentro de este panorama, los modelos urbanos tridimensionales han dejado de percibirse como un objetivo futurista, complejo o difícil de implementar, para convertirse en un componente esencial en múltiples aplicaciones urbanas y territoriales. Diversos estudios señalan que la demanda por modelos 3D precisos y actualizables es cada vez mayor, impulsada por necesidades relacionadas con planificación urbana, movilidad, energía, medio ambiente y gestión del riesgo (Biljecki et al., 2015; Jan, 2020)

Aunque los modelos 3D de edificios han avanzado considerablemente, la representación precisa de la vegetación, especialmente de los árboles, sigue siendo un desafío y una necesidad reconocida en la literatura científica. Los árboles urbanos influyen significativamente en el microclima, la mitigación del efecto isla de calor y la captura de carbono, por lo que su modelado detallado es esencial para simulaciones ambientales y planificación urbana sostenible (Li et al., 2025; H. Xu et al., 2021)

Los estudios actuales sobre reconstrucción de vegetación urbana se han concentrado

principalmente en aspectos geométricos y morfológicos de la vegetación, incorporando en algunos casos características fisiológicas como la especie del árbol, la distribución del follaje o los rasgos ópticos de las hojas. Sin embargo, estas aproximaciones no abordan de manera integral las características geométricas básicas necesarias para describir de forma completa la estructura tridimensional del arbolado (H. Xu et al., 2021).

La necesidad de incorporar modelos 3D de árboles dentro de los modelos urbanos existentes es evidente; sin embargo, para que estos modelos puedan integrarse adecuadamente en flujos de análisis urbano, deben generarse bajo estándares formales de representación, particularmente los Niveles de Detalle (LOD). El concepto de LOD es define el grado de complejidad geométrica y fidelidad que debe tener cada modelo según su finalidad. Esto permite disponer de versiones del árbol que van desde representaciones simplificadas hasta modelos altamente detallados, adaptados a diferentes requerimientos computacionales y de análisis.

En este marco, la construcción automatizada de árboles en múltiples LOD se convierte en un componente esencial para enriquecer los modelos urbanos 3D y garantizar su interoperabilidad con plataformas como CityJSON o CityGML, así como con herramientas de simulación ambiental. Por tanto, el presente trabajo implementa un flujo de trabajo reproducible y automatizado para la construcción de modelos tridimensionales de árboles a partir de datos LiDAR, generando representaciones en distintos Niveles de Detalle (LOD) y abordando la brecha existente en la modelación del arbolado urbano. Con ello, se contribuye a mejorar la calidad de los modelos urbanos y a fortalecer su integración en gemelos digitales urbanos y aplicaciones avanzadas de planificación y análisis ambiental.

## 2 Materiales y métodos

### 2.1 Área de estudio

La presente investigación se enfoca en la generación automatizada de modelos 3D de árboles a partir de datos LiDAR, aplicada a un área piloto de pequeña escala en los Países Bajos. La implementación se evalúa en Noordereiland, un sector urbano localizado en la ciudad de Róterdam (Provincia de Zuid-Holland).

Noordereiland es una isla fluvial situada sobre el río Mosa, caracterizada por límites geográficos claramente definidos, una estructura urbana consolidada y presencia significativa de vegetación urbana distribuida en corredores viales, parques y zonas residenciales.



### 2.2 Características de los datos

Para el desarrollo de este trabajo se emplean datos de nubes de puntos provenientes del Actueel Hoogtebestand Nederland (AHN3), disponibles públicamente a través de la plataforma holandesa de geodatos Publieke Dienstverlening Op de Kaart (PDOK). AHN corresponde al modelo nacional de elevaciones de los Países Bajos, generado mediante tecnología LiDAR aerotransportada (Airborne Laser Scanning), con el objetivo de levantar información topográfica de alta precisión para todo el territorio (Kadaster, 2018).

La captura de los datos se realizó entre 2014 y 2019, mediante vuelos con aeronaves equipadas con sensores láser, alcanzando una precisión vertical de hasta  $\pm 5$  cm (Google Earth

Engine, 2025). El producto AHN3 se caracteriza por tener una densidad promedio entre 6 y 10 puntos/m<sup>2</sup> (AHN, 2015), lo que permite identificar objetos individuales como edificaciones, vegetación y estructuras lineales. Esta densidad es suficiente para la reconstrucción de copas arbóreas y el modelado geométrico de estructuras tridimensionales.

La nube de puntos AHN3 se clasifica en cinco clases. Cuatro de estas clases se basan en las Clases de Puntos LiDAR Estándar de la Sociedad Americana de Fotogrametría y Teledetección (ASPRS) y una está definida por la Actueel Hoogtebestand Nederland (AHN, 2015). En la siguiente tabla se relacionan las clasificaciones.

**Tabla 1.** Clasificaciones AHN3

Valor de Clasificación	Clase (en inglés)	Clase en español	Estándar
1	Unclassified	Sin clasificar	Estándar ASPRS
2	Ground	Terreno (suelo)	Estándar ASPRS
6	Buildings	Edificaciones	Estándar ASPRS
9	Water	Agua	Estándar ASPRS
26	Civil Structure	Estructura civil	Definida por el usuario

**Fuente:** AHN, 2015

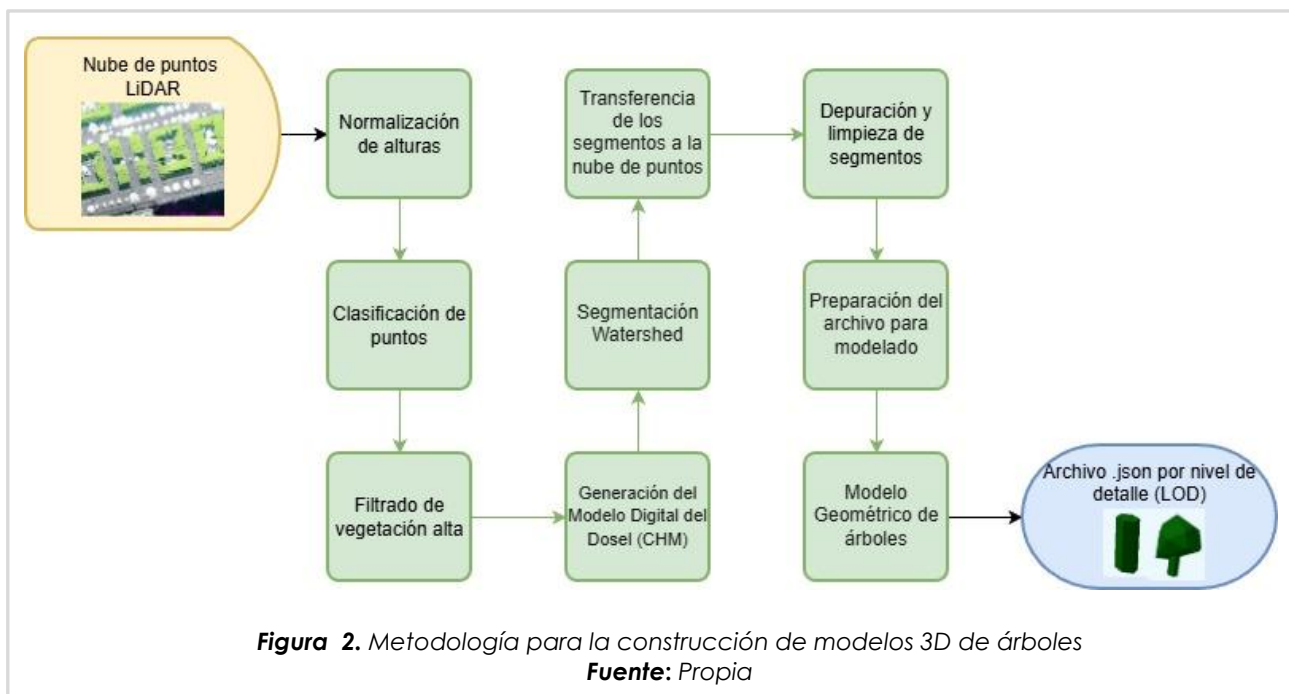
### 2.3 Construcción de modelos 3D de árboles

La metodología implementada para la construcción automatizada de modelos 3D de árboles se basa en una secuencia estructurada de etapas que combinan procesamiento de datos LiDAR, segmentación espacial, extracción de atributos geométricos y modelado paramétrico, con el fin de generar representaciones tridimensionales coherentes en diferentes Niveles de Detalle (LOD). En la **Figura 2** se presenta el flujo metodológico general, mediante el cual se ilustra el proceso completo desde la adquisición de los datos hasta la generación de los modelos 3D finales.

Este trabajo se fundamenta parcialmente en la implementación de la línea base metodológica propuesta por Jan, (2020) en la tesis "Automatic construction of 3D tree models in multiple levels of detail from airborne LiDAR

data", desarrollada en la Universidad Técnica de Delft (TU Delft). Dicha propuesta sirve como marco conceptual y técnico para la generación de modelos tridimensionales paramétricos a partir de datos LiDAR, adaptándola en este estudio a un contexto

experimental urbano específico, con ajustes en los procesos de segmentación, parametrización y estructura final de los modelos generados.



### 2.3.1 Clasificación de vegetación alta

La nube de puntos AHN3 empleada en esta investigación cuenta con una clasificación inicial de cinco clases. Sin embargo, para el modelado específico de árboles se trabaja con la clase sin clasificar, con el propósito de generar un conjunto de datos compuesto exclusivamente por puntos correspondientes a vegetación. Esto implica eliminar objetos como vehículos, embarcaciones, edificaciones, mobiliario urbano y superficies planas no vegetadas.

La reclasificación se realiza utilizando la información geométrica de los puntos (coordenadas X, Y, y altura normalizada) para estimar la rugosidad de la vecindad. Según Vukomanovic & Orr, (2014), las agrupaciones con alta variabilidad en altura presentan un comportamiento rugoso típico del follaje y estructura arbórea, mientras que los puntos con valores de altura similares

corresponden a superficies planas como techos, pavimento o suelo.

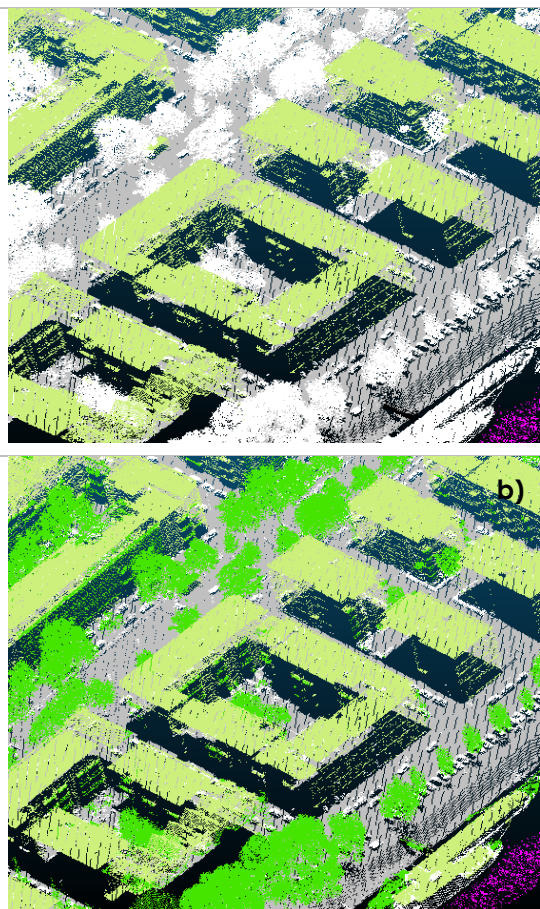
Para la reclasificación de la nube de puntos se emplea LAStools, una suite de software especializada para el manejo, análisis y filtrado de datos LiDAR; la cual ofrece algoritmos precompilados y altamente optimizados.

La reclasificación inicia con la normalización las alturas utilizando la herramienta lasheight de LAStools. Esta normalización consiste en recalcular la altura de cada punto con respecto al terreno, generando un valor de altura relativa que posteriormente se almacena como un nuevo atributo dentro del archivo procesado. El procedimiento se apoya en los puntos clasificados como Ground (clase 2) en los datos AHN3, lo cual permite estimar de manera precisa la diferencia vertical entre cada punto y la superficie del terreno más cercana.



Posteriormente, se utiliza la herramienta lasclassify (LAStools) para reclasificar automáticamente los puntos LiDAR con el fin de identificar vegetación alta (árboles) y estructuras artificiales, como edificaciones. Para ello se utiliza los atributos geométricos de los puntos (X, Y y altura normalizada previamente calculado), para evaluar la rugosidad local, es decir, la variación de altura entre un punto y sus vecinos inmediatos. De esta manera, se puede distinguir entre suelo, vegetación y superficies planas.

El resultado de este proceso se observa en la **Figura 3** donde los puntos originales (unclassified y ground) son transformados en puntos clasificados como terreno, árboles y objetos remanentes no clasificados.



**Figura 3.** Resultado clasificación. a) Nube de puntos inicial b) nube con Vegetación clasificada  
**Fuente:** Propia

Tras la clasificación automática de los puntos LiDAR, se realiza un proceso de filtrado para

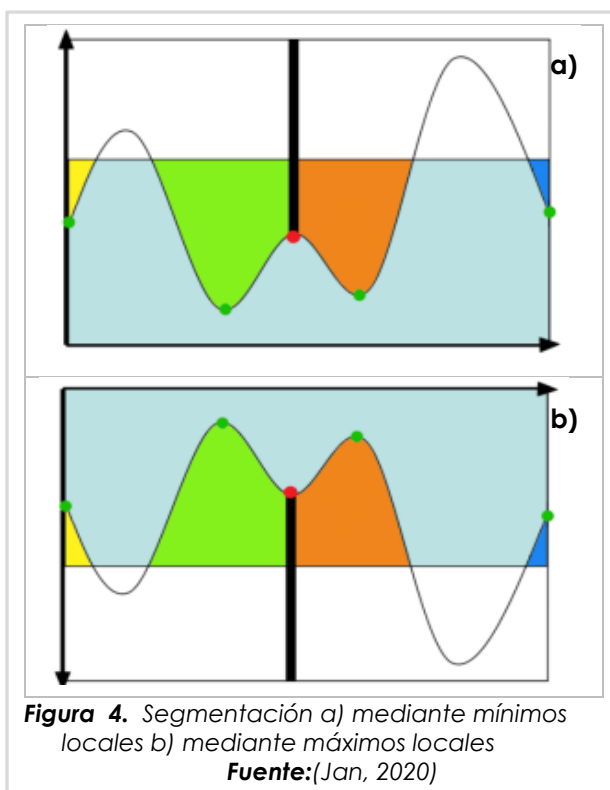
conservar únicamente aquellos puntos correspondientes a vegetación alta; el cual se realiza mediante el módulo las2las (LAStools), que permite extraer selectivamente una o varias clases dentro de un archivo LAS/LAZ. En este caso, se utiliza el parámetro -keep\_class 5 para conservar exclusivamente la clase 5 (High Vegetation).

### 2.3.2 Segmentación individual de árboles

La segmentación de la nube de puntos filtrada (vegetación alta) se realiza mediante un enfoque basado en la segmentación de cuencas hidrográficas (Watershed), módulo implementado en el software System for Automated Geoscientific Analyses (SAGA).

La segmentación por cuencas hidrográficas se basa en la interpretación de un Modelo Digital de Elevación (DEM) como una superficie continua en la que es posible delimitar regiones mediante la identificación de depresiones locales. En un DEM tradicional, los mínimos locales funcionan como semillas que marcan el inicio de cada cuenca. A medida que la superficie asciende desde cada semilla, se determinan los límites entre cuencas adyacentes, los cuales suelen ubicarse en puntos más elevados denominados sillas de montar. Estas sillas permiten establecer si dos cuencas corresponden a una misma unidad hidrológica o si representan entidades independientes. En la **Figura 4** se presenta una ilustración simplificada en 2D del funcionamiento de este proceso mediante semillas y sus respectivas sillas de montar.

Para adaptar este enfoque para los árboles, es necesario invertir el criterio, en lugar de trabajar con mínimos locales, se utilizan los máximos locales, ya que la copa del árbol corresponde al punto más alto de su geometría. Las zonas donde dos árboles se aproximan generan las equivalentes sillas de montar, que permiten definir el límite entre copas contiguas. El principio operativo se mantiene, la diferencia entre un máximo local (semilla) y la silla de montar determina si dos copas pertenecen al mismo segmento o si deben clasificarse como árboles individuales.



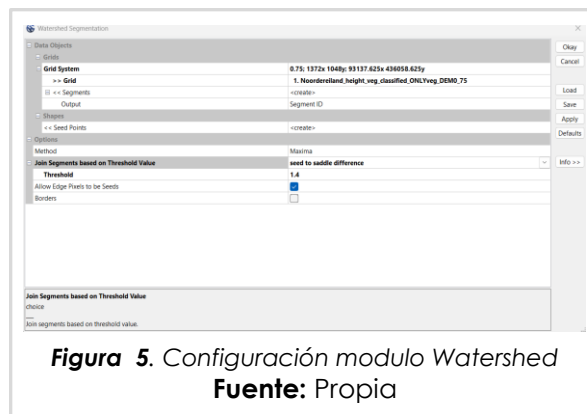
Como ya se explicó, este método funciona con un DEM, por lo que primero se debe convertir la nube de puntos a un DEM. Para ello, se emplea el módulo lasgrid (LAStools), el cual convierte los puntos LiDAR en raster, generando un Modelo Digital de Copas (Canopy Height Model, CHM) a partir del valor máximo de altura normalizada contenido en cada celda; se define un tamaño de celda de 0.75 metros, que ofrece un equilibrio adecuado entre resolución espacial y costo computacional (Jan, 2020)

Sobre el CHM, se aplica el módulo "Watershed Segmentation Module" en SAGA, basado en la lógica de segmentación de cuencas hidrográficas utilizando los siguientes parámetros:

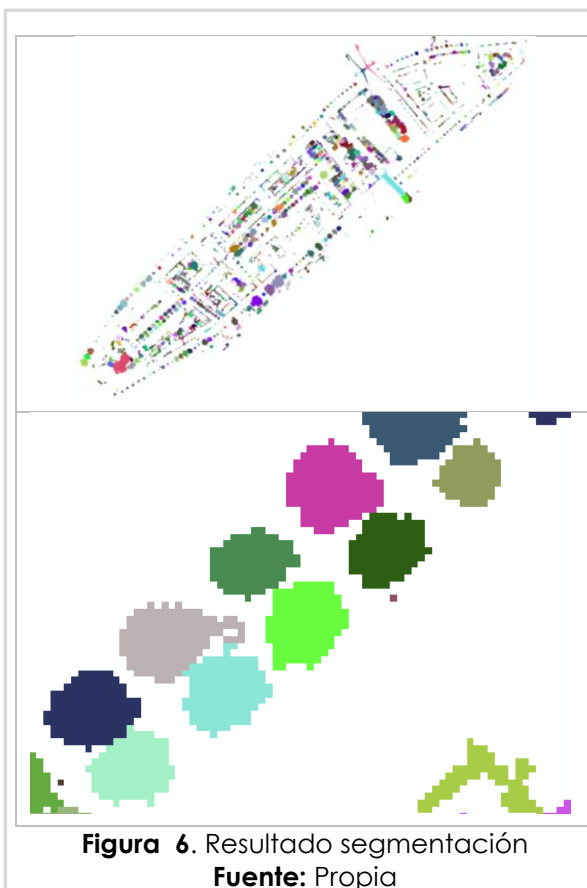
- Salida: "Segment ID" campo donde se almacena el ID de cada árbol individual.
- Método: "Máxima", ya que es el opuesto de una cuenca, un árbol.
- Unión segmentos según umbral: "Seed to saddle difference, ya que los árboles suelen estar cerca unos de otros, la silla se define como el lugar donde se encuentran las copas de

los árboles y las semillas son los máximos locales

- Umbral: 1.4 m, si la diferencia entre las semillas y las sillas de montar está por debajo del valor umbral, se considera un solo árbol.

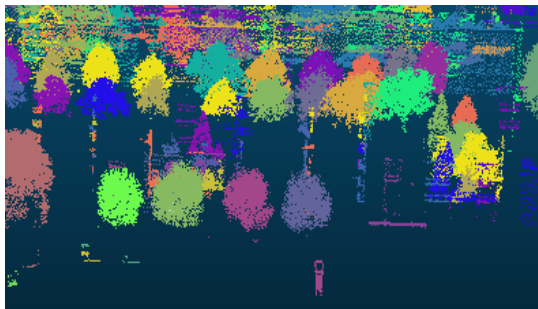


El resultado es un raster segmentado, donde cada píxel recibe un identificador único correspondiente a un árbol.



Posteriormente, el raster segmentado se superpone a la nube de puntos original,

asignándole a cada punto el identificador del segmento sobre el cual se encuentra. Este paso convierte el modelo raster en información tridimensional discreta por árbol. Así, se obtiene una nube de puntos segmentada individualmente, donde cada árbol no solo conserva su forma y altura real, sino que cuenta con una identidad única que facilita su análisis, visualización y modelado.



**Figura 7.** Nube de puntos segmentada  
**Fuente:** Propia

### 2.3.3 Depuración y limpieza de segmentos

Una vez que los puntos LiDAR han sido clasificados como vegetación y agrupados por segmentos (árboles potenciales), es necesario aplicar una fase de limpieza para garantizar que cada segmento represente efectivamente un árbol y que esté libre de ruido o geometrías no deseadas. Esta limpieza se realiza en varios pasos secuenciales.

En primer lugar, se aplican una serie de filtros iniciales a nivel de segmento. Se exige un mínimo de 50 puntos por segmento, ya que conjuntos más pequeños suelen corresponder a ruido, fragmentos de objetos o partes de copas mal delimitadas. Adicionalmente, se analiza la intensidad media y el número medio de retornos, los árboles tienden a presentar intensidades relativamente bajas (típicamente en rangos de 30 a 60), mientras que superficies altamente reflectantes como tejados metálicos o infraestructuras muestran intensidades más elevadas. Por este motivo, los segmentos cuya intensidad media supera un umbral de 100 se descartan como no arbóreos. De forma análoga, el número

medio de retornos de un segmento debe ser superior a 1.5, ya que la vegetación genera múltiples retornos debido a la penetración del haz láser en la copa, mientras que muchas superficies sólidas producen un único retorno. Finalmente, se impone una altura máxima de 50 m para descartar segmentos con errores altimétricos.

En segundo lugar, se evalúa la planaridad de cada segmento mediante un ajuste robusto de planos utilizando el algoritmo RANSAC (implementado en Scikit-learn). Este procedimiento permite verificar si los puntos del segmento pueden aproximarse a una superficie plana, adaptando el modelo para trabajar directamente en tres dimensiones. A diferencia de las regresiones lineales tradicionales, el ajuste se realiza sobre la nube de puntos completa en el espacio 3D, permitiendo descartar valores atípicos y obtener una estimación fiable del plano asociado al segmento.

El algoritmo selecciona iterativamente subconjuntos mínimos de puntos (3 puntos) y estima el plano que mejor se ajusta a la mayoría de puntos internos (inliers). Luego, se calcula la distancia media al plano entre las coordenadas reales ( $X$ ,  $Y$ ,  $Z$ ) y las coordenadas proyectadas sobre el plano. Cuando esta distancia media es inferior a un umbral de 0.1 m, se considera que el segmento es prácticamente plano y, por tanto, no corresponde a un árbol, sino a elementos como tejados, cubiertas o plataformas, por lo que se descarta del análisis.

En tercer lugar, se aplica la detección de planos parciales y valores atípicos dentro del propio segmento. Se identifican y eliminan subsecciones que se comportan como planos locales, así como puntos atípicos alejados de la estructura principal de la copa. Este procedimiento se repite de forma iterativa hasta que no se detectan nuevos planos ni outliers significativos. Solo los segmentos que conservan una estructura tridimensional coherente, con suficiente número de puntos y sin geometrías planas dominantes, se consideran árboles válidos y se utilizan en la etapa posterior de extracción de parámetros y modelado 3D.

### 2.3.4 Preparación del archivo para modelado

Una vez completadas las etapas de depuración y parametrización de los segmentos arbóreos, se ejecuta un proceso de clasificación automática del tipo de árbol. Este procedimiento se basa en el entrenamiento supervisado de una red neuronal, la cual permite predecir la especie o tipología del árbol (hoja caduca, conífera, ornamental, entre otros), igualmente el modelo asigna una probabilidad asociada a dicha predicción (classification certainty), que posteriormente se almacena como metadato en el archivo CityJSON.

En el flujo de trabajo original, estos atributos (TreeType y Classification\_Certainty) cumplen un papel relevante, dado que no solo enriquecen la descripción semántica de cada objeto, sino que también influyen en la representación geométrica final del árbol, ajustando la forma de la copa o la proporción de sus componentes según la especie estimada.

No obstante, en el presente estudio no se dispone del modelo previamente entrenado, ni se cuenta con un conjunto de datos rotulado que permita realizar el entrenamiento supervisado. Por esta razón, la etapa de clasificación automática no pudo ser reproducida.

Con el fin de preservar la compatibilidad con la estructura de datos requerida por el módulo de reconstrucción geométrica, se decidió simular la salida del clasificador, generando de manera controlada las columnas TreeType y Classification\_Certainty para cada segmento. A TreeType se le asignó el valor fijo de "GenericTree" que *representa la categoría taxonómica*; y para Classification\_Certainty, se definió un valor constante de 1.0 con el propósito de simular el nivel de confianza que generaría un clasificador real. Este procedimiento garantiza que el script de modelado 3D pueda ejecutarse sin modificaciones, respetando la lógica del flujo original, aun cuando la clasificación real no se haya realizado.

### 2.3.5 Modelo geométrico 3D de árboles

La última etapa del flujo consiste en la generación de modelos 3D explícitos de los árboles a partir de los parámetros geométricos previamente calculados.

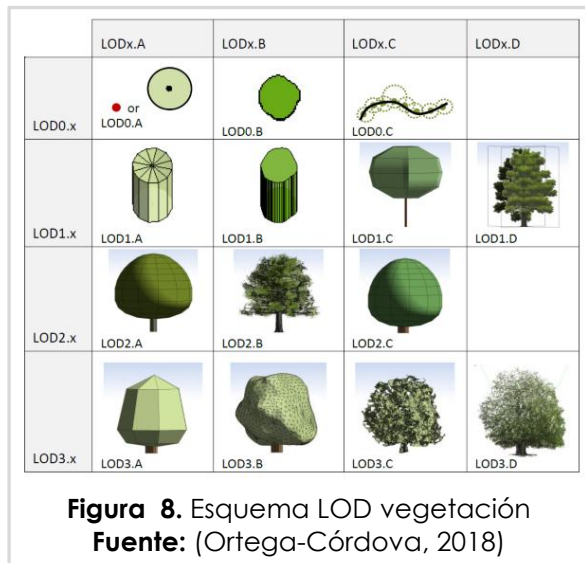
En primera instancia, se establece la estructura base del archivo CityJSON que contendrá los modelos generados. Esto incluye la creación del contenedor general donde se almacenarán los objetos urbanos (CityObjects), la definición de la lista de vértices que conformarán las geometrías tridimensionales y la configuración de los elementos visuales (tronco y copa).

Una vez preparado este esquema inicial, se procede a recorrer el conjunto de árboles identificados en la nube de puntos. Para cada individuo, el sistema genera una representación tridimensional específica, ajustada al Nivel de Detalle (LOD) seleccionado. De este modo, la geometría del árbol se construye de manera personalizada, variando desde formas simples y generalizadas hasta modelos más complejos que incorporan la estructura del tronco y la morfología de la copa con mayor precisión.

El Nivel de Detalle (LOD) es un atributo clave para valorar la calidad de un modelo 3D y su grado de correspondencia con la realidad, pues determina cuán fielmente se representa la geometría y la complejidad del objeto, influyendo directamente en su utilidad para diversas aplicaciones urbanas y ambientales (Biljecki et al., 2014).

En el caso de los árboles, uno de los esquemas más completos es el propuesto por Ortega-Córdova, (2018), presentada en la **Figura 8**, quien define distintos LOD que van desde representaciones simplificadas hasta modelos que describen de manera detallada la copa y el tronco, según los datos disponibles y el propósito del modelo.





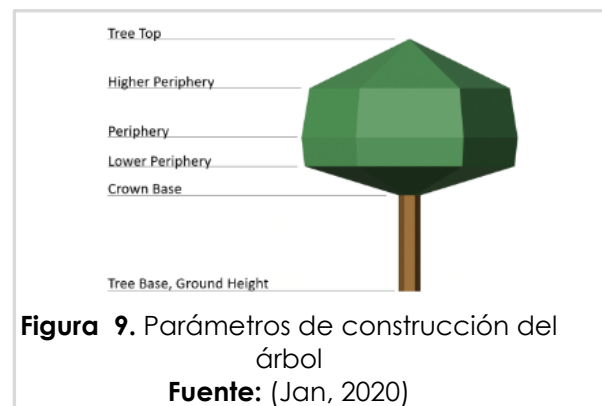
El propósito final de esta implementación es generar, como salida del proceso automático, al menos un modelo representativo de cada Nivel de Detalle, siguiendo el esquema conceptual de Ortega-Córdova, (2018). Los modelos implícitos utilizados en esta investigación se basan en geometrías hexagonales, ya que esta forma permite mantener un número reducido de vértices sin perder la apariencia circular necesaria para representar copas o secciones del árbol. A diferencia de formas con mayor complejidad como los círculos, que requieren muchos más vértices los hexágonos ofrecen un buen equilibrio entre simplicidad geométrica y fidelidad visual.

Los vértices de cada hexágono se construyen a partir del centro del árbol (coordenadas X e Y) y del radio correspondiente a la sección que se desea representar. Se construyen seis puntos que definen el hexágono, y posteriormente se conectan para formar las caras del modelo.

En cuanto a la parametrización, la mayoría de los árboles se generan utilizando valores extraídos directamente de su segmento en la nube de puntos. Antes de calcular estos parámetros, se aplica una corrección a las alturas, ajustando el valor Z mediante la resta del promedio de las diferencias entre la altura real y la altura previamente estimada. Este ajuste asegura que los modelos resultantes mantengan la coherencia vertical necesaria para su construcción tridimensional.

Cada segmento se divide verticalmente en secciones, definidas entre la altura mínima y máxima del árbol. En cada sección se registran atributos básicos: número de puntos, centroide, altura representativa y radio estimado. A partir de esta discretización se derivan los parámetros estructurales que permiten modelar la forma del árbol. Los principales parámetros obtenidos son los siguientes:

- ✓ Base del árbol: Considerada como cero después de aplicar la corrección vertical necesaria para asegurar coherencia en el modelo 3D.
- ✓ Radio del tronco: Estimado como el 10% del radio de periferia.
- ✓ Base de la copa: Definida como el percentil 5 de la distribución de alturas del segmento, lo cual representa el inicio de la estructura de copa en la mayoría de los árboles.
- ✓ Altura y radio de la periferia principal: Determinados a partir de la división que contiene la mayor densidad de puntos, generalmente coincidente con la parte más volumétrica de la copa.
- ✓ Periferia inferior (altura y radio): La altura se calcula como el punto medio entre la base de la copa y la altura periférica principal. El radio se extrae de la sección en la que se encuentra esta altura.
- ✓ Periferia superior (altura y radio): Se define como el punto medio entre la altura periférica principal y la parte superior del árbol. El radio se toma de la sección correspondiente.
- ✓ Cima del árbol: Asignada al percentil 99 de la distribución de alturas, evitando valores atípicos.



Cada LOD utiliza una combinación diferente de parámetros extraídos para construir el modelo 3D de árbol.

En LOD0, solo utiliza el radio de la periferia y la base del árbol, se representa mediante un hexágono plano en planta, centrado en la posición del árbol y con un radio equivalente al radio estimado. Esta representación funciona como una huella geométrica simplificada que aproxima la proyección de la copa sobre el terreno.

En LOD1, se utiliza el radio de la periferia, la base del árbol y la copa del árbol, la representación se eleva a un volumen tridimensional básico mediante un prisma hexagonal que se extiende desde la base del árbol hasta la altura máxima registrada. Este modelo volumétrico permite visualizar de manera ligera y eficiente la ocupación espacial del arbolado en el entorno urbano.

En LOD2, se emplean todos los parámetros derivados directamente de la nube de puntos. El tronco se representa mediante un prisma hexagonal, mientras que la copa se reconstruye a partir de varias coronas hexagonales apiladas, lo que permite aproximar con mayor fidelidad la morfología general del árbol a partir de su distribución vertical y radial.

En LOD3, se utiliza la nube de puntos clasificada por árbol para generar un modelo que refleja con mayor precisión la geometría real. Según la configuración adoptada, se calcula una envolvente tridimensional mediante un casco convexo, obteniendo una malla que reproduce explícitamente el volumen de la copa. Sobre esta superficie se integra un tronco paramétrico similar al empleado en LOD2, resultando en un modelo altamente detallado que captura la forma específica del árbol a partir de los datos LiDAR.

Finalmente, cada árbol se almacena en los CityObjects junto con sus atributos principales, entre ellos el identificador de segmento, el número total de puntos, la intensidad media y el retorno promedio. Las

coordenadas utilizadas en la geometría se normalizan a metros (mediante división por 1000) y se redondean con el fin de optimizar el tamaño del archivo sin afectar la representación. El modelo completo se exporta en formato CityJSON, quedando listo para su integración en visores 3D, plataformas de gemelos digitales urbanos o bases de datos de modelos de ciudad en tres dimensiones.

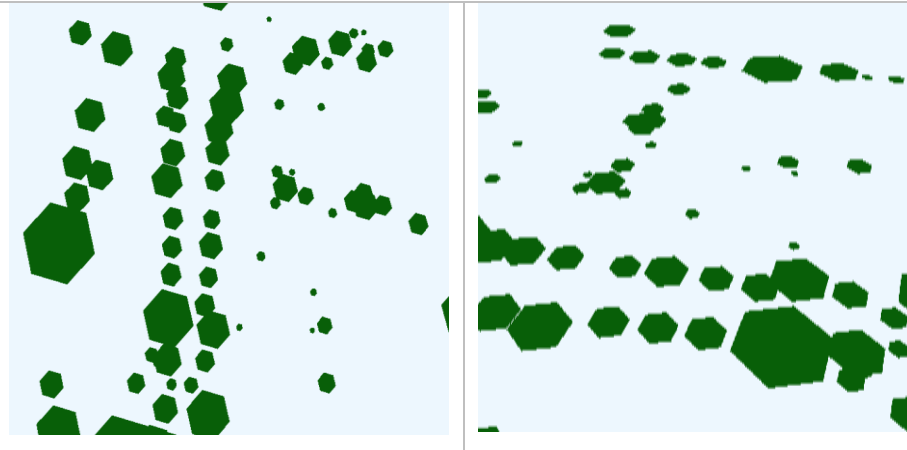
Para la exploración y validación de los resultados se empleó CityJSON Ninja, un visor oficial desarrollado por la iniciativa CityJSON/CityGML. Esta herramienta permite cargar, visualizar e inspeccionar modelos 3D directamente desde el navegador, sin necesidad de instalar software adicional (ver video en el repositorio Github).

### 3 Resultados

En esta sección se presentan los resultados obtenidos a partir del proceso automatizado de reconstrucción tridimensional de árboles utilizando datos LiDAR. Las visualizaciones generadas permiten evaluar el comportamiento del método propuesto en cada uno de los Niveles de Detalle (LOD) implementados, desde representaciones geométricas simplificadas hasta modelos que incorporan la estructura completa del árbol.

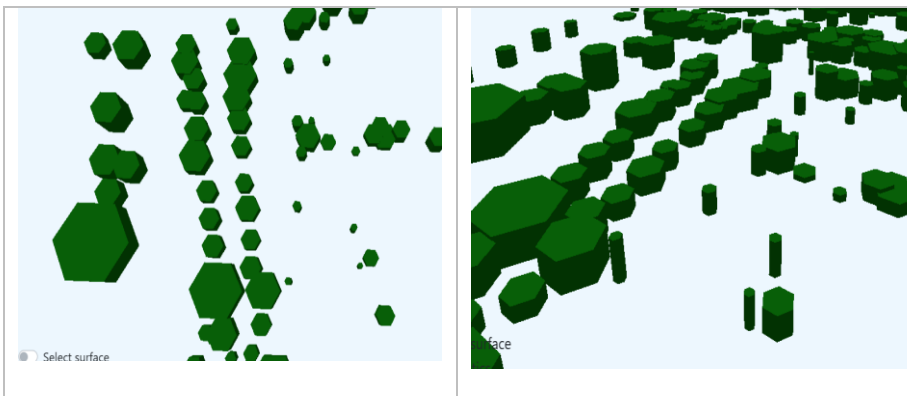
En total, se construyeron 5.773 segmentos correspondientes a árboles, los cuales pasaron por las etapas de limpieza, extracción de parámetros y modelado geométrico según el nivel de detalle definido.

A continuación, se muestran las visualizaciones correspondientes a los modelos generados en LOD0, LOD1, LOD2 y LOD3, destacando las diferencias en estructura, nivel de detalle y fidelidad geométrica respecto a la nube de puntos original, desde una forma básica hexagonal hasta volúmenes detallados basados en envolventes convexas.



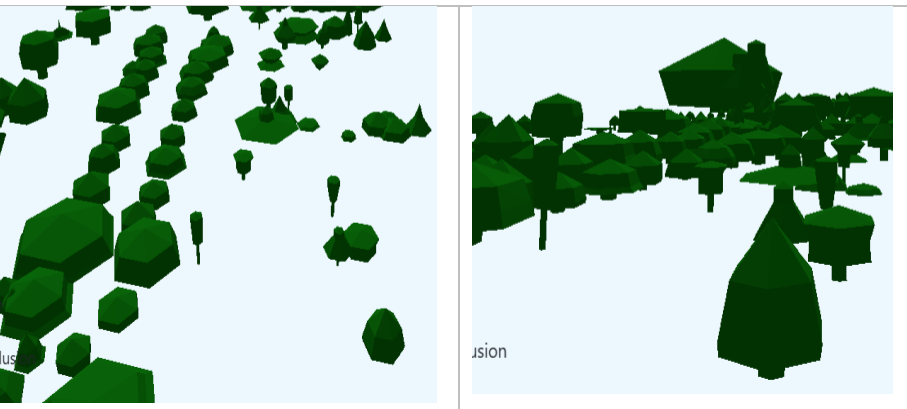
**Figura 10.** Visualización LOD0

**Fuente:** Propia



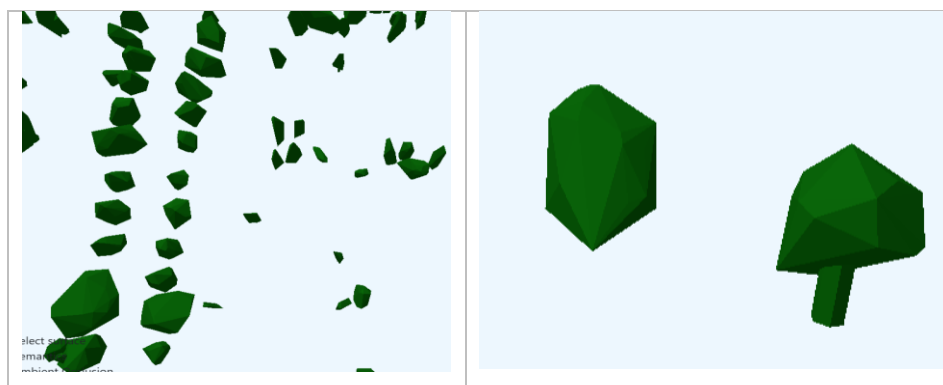
**Figura 11.** Visualización LOD1

**Fuente:** Propia



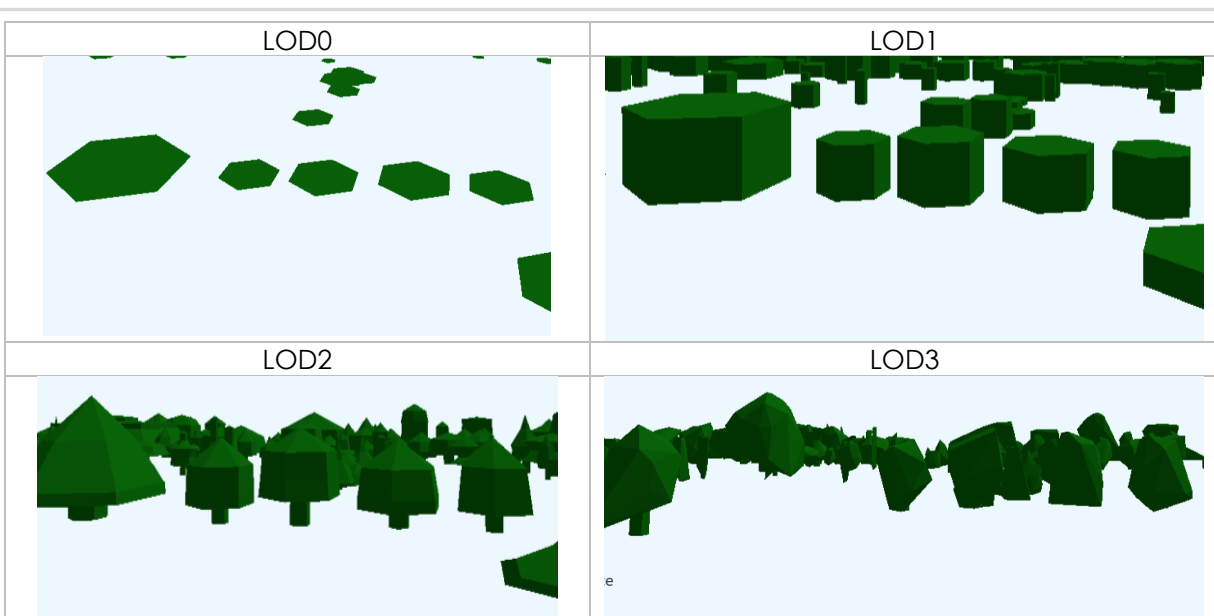
**Figura 12.** Visualización LOD2

**Fuente:** Propia



**Figura 13.** Visualización LOD3

**Fuente:** Propia



**Figura 14.** Comparación LOD

**Fuente:** Propia

Los resultados muestran la progresión geométrica de los modelos conforme aumenta el LOD, evidenciando cómo los parámetros extraídos de la nube de puntos influyen en la forma final de los árboles. De este modo, las figuras incluidas ilustran no solo la calidad de las geometrías obtenidas, sino también la capacidad del enfoque paramétrico para adaptar la complejidad del modelo según la disponibilidad de datos y el propósito de la representación.

Sin embargo, no todos los árboles reconstruidos presentan la forma esperada. Esto se debe a diversos factores inherentes a la calidad de los datos y a las limitaciones del proceso de limpieza previa. En particular, se

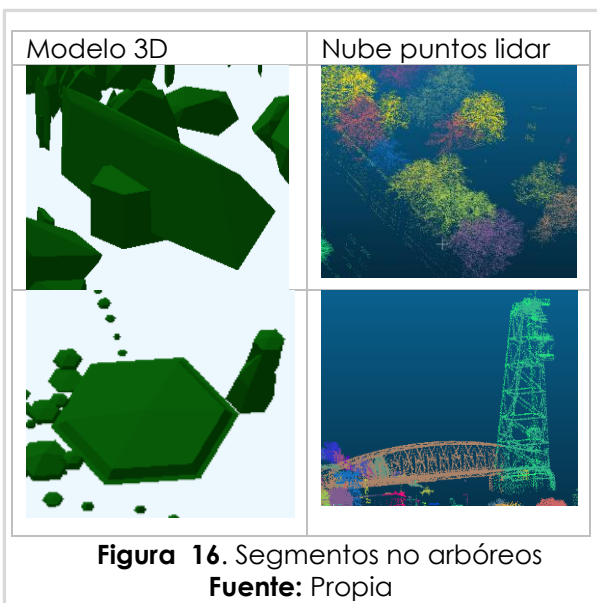
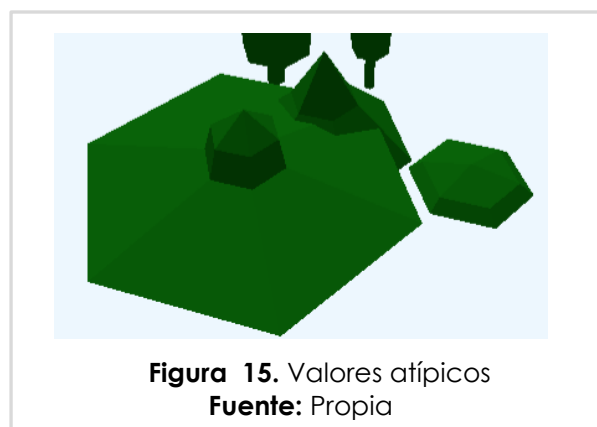
identifican dos situaciones principales que afectan la geometría final:

Valores atípicos no eliminados: cuando ciertos outliers permanecen dentro del segmento, estos generan irregularidades visibles en la copa del árbol, que se manifiestan como picos o “espinas” que sobresalen del volumen principal. Estas deformaciones se originan en puntos aislados con posiciones anómalas que alteran la envolvente 3D utilizada para modelar el árbol (Ver **Figura 15**).

Segmentos que no representan árboles reales: en algunos casos, ciertos grupos de puntos que deberían haber sido filtrados



durante la fase de depuración permanecen en el conjunto de datos. Al pasar por el modelador geométrico, estos segmentos no arbóreos se traducen en formas incoherentes o artefactos que no corresponden a la vegetación, afectando la limpieza visual y la interpretación final del modelo. En la **Figura 16** se compara el modelo 3d con los puntos lidar, donde se evidencia que el modelo 3D está reconstruyendo geometrías que no corresponden a árboles reales.



#### 4 Discusión

La metodología implementada demostró ser capaz de generar modelos tridimensionales de árboles a partir de nubes LiDAR, siguiendo un flujo de trabajo que integra segmentación, limpieza, extracción de parámetros morfométricos y modelado geométrico bajo el estándar CityJSON.

En las figuras 10 a 13 se puede observar la evolución progresiva de los modelos arbóreos a medida que aumenta el Nivel de Detalle (LOD). Cada representación responde a una lógica distinta de construcción, basada en parámetros morfométricos derivados de la nube de puntos o, en el caso de LOD3, en una reconstrucción directa basada en la geometría real del árbol. Estas visualizaciones permiten apreciar cómo cada nivel captura con mayor o menor éxito la estructura de la vegetación presente en el área de estudio.

En LOD0 (ver **Figura 10**) el nivel más básico, los árboles se representan como figuras planas, semejantes a polígonos regulares apoyados sobre el plano del terreno. A simple vista, esta representación proporciona una visión general de la distribución espacial de la vegetación, las formas aparecen alineadas siguiendo las hileras reales de árboles, lo que confirma que el proceso de segmentación logró recuperar su ubicación en el espacio. Sin embargo, también se nota la presencia de numerosos polígonos pequeños o muy próximos entre sí, reflejo de la fragmentación de copas que suele ocurrir en nubes LiDAR con variaciones internas de densidad (Sun et al., 2022). Aunque estas formas son simples, sirven como punto de partida para comprender la disposición general y la densidad del arbolado.

Por otro lado, el nivel LOD1 (ver **Figura 11**) los modelos dejan de ser figuras planas y se convierten en prismas que representan el volumen global del árbol, lo que permite apreciar diferencias en altura y tamaño entre los distintos segmentos, logrando una percepción más realista. El paisaje tridimensional resultante muestra hileras de volúmenes que ya comienzan a asemejarse a árboles. Sin embargo, también se observan cuerpos delgados o muy pequeños que corresponden a segmentos residuales. Estas limitaciones se deben tanto a la complejidad geométrica de la vegetación urbana como a la naturaleza irregular de los datos LiDAR.

En este sentido, diversos estudios han señalado la necesidad de aplicar un refinamiento post-segmentación, tal como recomienda (Wan Mohd Jaafar et al., 2018),

con el fin de mejorar la delimitación de árboles individuales. Dicho refinamiento incluye operaciones como la fusión de segmentos adyacentes, la eliminación o resegmentación de copas con geometrías anómalas y de segmentos demasiado pequeños para considerarse válidos. La incorporación de este ajuste secundario permite reducir errores de sobresgmentación y subsegmentación.

Por otro parte, el nivel LOD2 (ver **Figura 12**) presenta una apariencia más natural, donde se separa la copa del tronco y se introducen volúmenes más estilizados que evocan la geometría real de un árbol. Las copas aparecen redondeadas o ligeramente cónicas, y los troncos adquieren proporciones más visibles; permitiendo reconocer árboles individuales y distinguir variaciones de tamaño y forma. No obstante, también se identifican casos donde la geometría es incompleta o exagerada, producto de segmentos que contienen pocos puntos o muestran estructuras parciales.

Asimismo, se evidencian deformaciones visibles en la superficie estimada de la copa, manifestándose como protuberancias o picos que sobresalen del volumen principal (ver **Figura 15**). Este fenómeno ha sido documentado en diferentes enfoques de reconstrucción arbórea, donde los outliers generan artefactos geométricos tanto en modelos paramétricos como en envolventes, generando resultados no representativos de la estructura real (Jan, 2020; Wei et al., 2022)

A diferencia de los niveles anteriores, LOD3 (ver **Figura 13** **Figura 10**) no se construye mediante parámetros estimados, sino directamente a partir de la forma real de la nube de puntos. Este método intenta capturar la silueta precisa de cada árbol mediante envolventes convexas, lo que en teoría debería producir el modelo más fiel. Sin embargo, los resultados muestran que esta aproximación depende profundamente de la densidad y continuidad de los datos. En las visualizaciones se observan copas distorsionadas, superficies irregulares y, especialmente, troncos que no se

reconstruyen adecuadamente o desaparecen por completo. Esto ocurre porque la nube LiDAR empleada presenta pocos puntos en la zona del tronco y variaciones importantes dentro de las copas, lo que dificulta formar volúmenes cerrados y estables.

Por ejemplo, Tarsha Kurdi et al., (2024) señala que los parámetros utilizados en el modelado dependen directamente de la densidad de puntos, independientemente de las características geométricas reales del árbol; esto implica que copas con baja densidad pueden generar estimaciones sesgadas. Asimismo, Jan, (2020) menciona que un número reducido de puntos en la copa o el tronco puede generar modelos con triángulos excesivamente grandes, lo que produce geometrías que pueden parecer visualmente extrañas, aunque no necesariamente incorrectas desde el punto de vista topológico.

Esta situación se debe a que el algoritmo debe reconstruir la envolvente con la mínima información disponible, lo que obliga a interpolar superficies amplias ante la ausencia de detalles finos. Este comportamiento reafirma la dependencia directa del proceso de modelado respecto a la densidad y distribución de los datos LiDAR, lo que condiciona la fidelidad geométrica alcanzable.

Por otra parte, se evidenciaron muchos segmentos no arbóreos (ver **Figura 16**) en todos los niveles de detalle, a pesar de que el conjunto de datos fue sometido a una depuración previa para eliminar puntos no correspondientes a vegetación, la segmentación inicial de la nube LiDAR produjo un total de 5.773 segmentos, muchos de los cuales no representaban árboles completos ni estructuras coherentes. Posteriormente, al aplicar los filtros morfométricos como el mínimo de puntos, intensidad promedio, número de retornos, altura máxima, verificación de planaridad mediante RANSAC y limpieza de valores atípicos solo 685 segmentos cumplieron las condiciones necesarias para ser considerados árboles reales.

La clasificación inicial de la nube de puntos fue realizada mediante *lasclassify*, la cual

permitió separar de forma preliminar la vegetación alta de superficies planas como suelo y edificaciones; sin embargo, tiene una limitación, ya que se basa únicamente en las coordenadas X, Y y en la altura normalizada de cada punto, sin incorporar otras propiedades espectrales o geométricas que podrían mejorar la separación entre clases. Esta simplicidad metodológica restringe su capacidad de discriminación, y diversos estudios han demostrado que su precisión oscila entre **60–70%**, principalmente debido a errores de clasificación y a un número considerable de puntos que permanecen etiquetados como *unclassified* (McIver et al., 2017). En este sentido, la experiencia obtenida confirma que la clasificación automática basada únicamente en rugosidad y altura, aunque útil, requiere complementarse con filtros más específicos o enfoques supervisados para minimizar la propagación de errores.

Igualmente se evidencio que el proceso de reconstrucción 3D no incorporó estos criterios de depuración y limpieza. En consecuencia, el modelador interpretó cada segmento, válido o no, como una entidad arbórea independiente, generando automáticamente un modelo 3D para cada uno de ellos. Esto explica por qué, aun contando únicamente 685 árboles válidos, se terminaron creando más de cinco mil modelos. Muchos de estos corresponden a fragmentos de copas, arbustos, cúmulos aislados de puntos e incluso ruido residual presente en la nube LiDAR, que la etapa de segmentación no logró fusionar ni descartar adecuadamente.

Para garantizar que el modelado 3D reflejara únicamente árboles reales, se debió limitar explícitamente la entrada del constructor geométrico a los segmentos validados por los filtros morfométricos. Es decir que era necesario depurar no solo los puntos, sino los segmentos como unidades estructurales, y utilizar únicamente esos segmentos aprobados como base para la reconstrucción (Itakura et al., 2022; X. Xu et al., 2023). Esto habría evitado la generación de miles de modelos espurios sin correspondencia real a la vegetación del área de estudio.

## 5 Conclusiones

- i. Los niveles de detalle LOD0, LOD1 y LOD2 lograron representar la vegetación con grados crecientes de complejidad, mostrando un progreso coherente entre la ubicación espacial (LOD0), el volumen básico (LOD1) y la forma arbórea diferenciada (LOD2). En particular, el LOD2 alcanzó la representación más equilibrada entre simplicidad y naturalidad, logrando distinguir tronco y copa.
- ii. El nivel LOD3, construido directamente a partir de la nube de puntos mediante envolventes convexas, evidenció fuertes limitaciones derivadas de la baja densidad en troncos y copas. La falta de puntos suficientes impide generar superficies cerradas y obliga al algoritmo a interpolar grandes triángulos, lo que ocasiona geometrías distorsionadas, copas incompletas y la ausencia de troncos.
- iii. La clasificación preliminar basada únicamente en altura normalizada y rugosidad, aunque útil para separar vegetación del terreno, tiene una capacidad discriminante limitada y tiende a producir errores de clasificación que se propagan hacia etapas posteriores del proceso, afectando la fidelidad del modelado.
- iv. Los resultados mostraron que la aproximación basada en un Modelo Digital del Dosel (CHM) y segmentación basada en cuencas hidrográficas Watershed constituye un punto de partida sólido para la identificación de copas individuales, tal como ha sido reportado previamente en estudios de análisis forestal y teledetección urbana.
- v. Se evidenció una importante discrepancia entre el número total de segmentos generados (5.773) y los árboles considerados válidos tras aplicar los filtros morfométricos (685), demostrando que la segmentación inicial produjo una cantidad elevada

de fragmentos no arbóreos, y además se constató que el modelador 3D no incorporó estos criterios y, en consecuencia, generó un modelo por cada segmento, independientemente de su validez estructural.

## 6 Recomendaciones

- i. Se recomienda contar con un conjunto de datos de referencia que permita evaluar la calidad geométrica y temática de los modelos obtenidos. Esta validación es esencial para determinar si el flujo de trabajo propuesto logra reproducir con fidelidad las dimensiones, alturas, alineamientos y relaciones espaciales presentes en el entorno real.
- ii. Para evitar la generación masiva de modelos 3D no válidos, se recomienda incorporar un filtro previo al modelado, en el que únicamente los segmentos que superaron las pruebas morfométricas sean enviados al constructor de geometrías.
- iii. La clasificación inicial debe complementarse con variables como intensidad o índices derivados del número de retornos; para mejorar la separación entre vegetación, edificaciones y ruido, reduciendo errores sistemáticos
- iv. Se sugiere integrar un procedimiento de clasificación automática de tipos de árbol antes del modelado geométrico, ya que permite ajustar la representación geométrica de manera más realista, modificando parámetros como la forma de la copa, la proporción del tronco o la distribución volumétrica en función de la especie estimada.
- v. Se recomienda evaluar otros métodos de iconización que permitan mejorar la estabilidad y el realismo de los modelos, tales como SILVigenesis / SILVI-ESTRELLA (Ortega-Córdova, 2018) o incorporar formas

alfa (Kreveld et al., 2011) como alternativa a la envolvente convexa; estos enfoques proporcionan un mayor control sobre el nivel de detalle y permiten representar de manera más precisa las irregularidades naturales de la copa.

## 7 Agradecimientos

Agradezco al autor (Jan, 2020) su investigación y la publicación de su repositorio en GitHub. Su propuesta constituye un referente conceptual y técnico fundamental para la construcción automática de modelos tridimensionales de árboles a partir de datos LiDAR.

## 8 Referencias

- AHN, A. H. N. (2015). *Besteksvoorwaarden "inwinning lands dekkende dataset ahn2014 -2019"*.
- Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., & Çöltekin, A. (2015). Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information*, 4(4), 2842–2889. <https://doi.org/10.3390/ijgi4042842>
- Itakura, K., Miyatani, S., & Hosoi, F. (2022). Estimating Tree Structural Parameters via Automatic Tree Segmentation from LiDAR Point Cloud Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 555–564. <https://doi.org/10.1109/JSTARS.2021.3135491>
- Jan, G. (2020). *Automatic construction of 3D tree models in multiple levels of detail from airborne LiDAR data*.
- Li, R., Wang, L., Zhai, Y., Huang, Z., Jia, J., Wang, H., Ding, M., Fang, J., Yao, Y., Ye, Z., Hao, S., & Fan, Y. (2025). Modeling LiDAR-Derived 3D Structural Metric Estimates of Individual Tree Aboveground Biomass in Urban Forests: A Systematic Review of Empirical Studies. *Forests*, 16(3), 390. <https://doi.org/10.3390/f16030390>
- Luo, J., Liu, P., Xu, W., Zhao, T., & Biljecki, F. (2025). A perception-powered urban digital twin to support human-centered



- urban planning and sustainable city development. *Cities*, 156, 105473. <https://doi.org/10.1016/J.CITIES.2024.105473>
- McIver, C. A., Metcalf, J. P., & Olsen, R. C. (2017). *Spectral LiDAR analysis for terrain classification* (M. D. Turner & G. W. Kamerman, Eds.; p. 101910J). <https://doi.org/10.1117/12.2276658>
- Ortega-Córdova, L. M. (2018). *MSc thesis in Geomatics for the Built Environment Urban Vegetation Modeling 3D Levels of Detail*.
- Sun, C., Huang, C., Zhang, H., Chen, B., An, F., Wang, L., & Yun, T. (2022). Individual Tree Crown Segmentation and Crown Width Extraction From a Heightmap Derived From Aerial Laser Scanning Data Using a Deep Learning Framework. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.914974>
- Tarsha Kurdi, F., Gharineiat, Z., Lewandowicz, E., & Shan, J. (2024). Modeling the Geometry of Tree Trunks Using LiDAR Data. *Forests*, 15(2), 368. <https://doi.org/10.3390/f15020368>
- Vukomanovic, J., & Orr, B. (2014). Landscape Aesthetics and the Scenic Drivers of Amenity Migration in the New West: Naturalness, Visual Scale, and Complexity. *Land*, 3(2), 390–413. <https://doi.org/10.3390/land3020390>
- Wan Mohd Jaafar, W. S., Woodhouse, I. H., Silva, C. A., Omar, H., Abdul Maulud, K. N., Hudak, A. T., Klauberg, C., Cardil, A., & Mohan, M. (2018). Improving Individual Tree Crown Delineation and Attributes Estimation of Tropical Forests Using Airborne LiDAR Data. *Forests*, 9(12), 759. <https://doi.org/10.3390/f9120759>
- Wei, X., Liu, M., Ling, Z., & Su, H. (2022). Approximate convex decomposition for 3D meshes with collision-aware concavity and tree search. *ACM Transactions on Graphics*, 41(4), 1–18. <https://doi.org/10.1145/3528223.3530103>
- Xu, H., Wang, C. C., Shen, X., & Zlatanova, S. (2021). 3D Tree Reconstruction in Support of Urban Microclimate Simulation: A Comprehensive Literature Review. *Buildings*, 11(9), 417. <https://doi.org/10.3390/buildings11090417>
- Xu, X., Iuricich, F., Calders, K., Armston, J., & De Floriani, L. (2023). Topology-based individual tree segmentation for automated processing of terrestrial laser scanning point clouds. *International Journal of Applied Earth Observation and Geoinformation*, 116. <https://doi.org/10.1016/j.jag.2022.103145>