

PROPUESTA METODOLÓGICA PARA EL CÁLCULO A GRAN ESCALA DE
ÁRBOLES INDIVIDUALES MEDIANTE EL USO DE DATOS LIDAR (Airborne Light
Scanner, ALS)

Autores

Christian Ricardo Gómez González

William Alberto Sáenz Blanco

UNIVERSIDAD ANTONIO NARIÑO

FACULTAD DE INGENIERÍA AMBIENTAL Y CIVIL

ESPECIALIZACIÓN EN SISTEMAS DE INFORMACIÓN GEOGRÁFICA

DIRECTOR: ANDRÉS FELIPE CARVAJAL VANEGAS, DR.

12 de julio de 2021

Tabla de contenido

Resumen.....	4
Abstract.....	5
Introducción.....	6
Estado del Conocimiento.....	8
Marco Teórico.....	10
Objetivos.....	13
Objetivo general.....	13
Objetivos específicos.....	13
Metodología.....	14
Área de Estudio.....	14
Sensor.....	15
Procesado de Datos LIDAR.....	15
Extracción de datos crudos.....	16
Trayectoria.....	16
Clasificación y edición.....	16
Detección de Árboles.....	17
Agrupación.....	17
Segmentación.....	18
Extracción.....	18
Control de calidad.....	18
Geográfico.....	18
Estadística.....	19
Diagrama de Flujo Metodología.....	20
Resultados.....	21
Definición de la Metodología.....	21
Paso 1. Adquisición de la información.....	21
Paso 2. Cálculo de la trayectoria.....	22
Ajuste y clasificación de la nube de puntos.....	23
Agrupación.....	24
Extracción de los resultados de la segmentación.....	27
Evaluación de la Precisión y Calidad.....	28
Áreas de muestreo.....	28
Sub segmentación y Sobre segmentación.....	32

Caracterización	33
La exhaustividad (Recall).....	34
Conclusiones	37
Recomendaciones	39
Bibliografía	40
Anexos	42

Lista de Tablas

Tabla 1. Matriz de confusión para clasificación binaria	19
Tabla 2. Parámetros	26
Tabla 3. Resultados Iniciales.....	29
Tabla 4. Resultados Área 1	30
Tabla 5. Resultados Área 2	31
Tabla 6. Resultados Área 3	32
Tabla 7. Modelo de Matriz de confusión.....	33
Tabla 8. Matriz para Área 1	34
Tabla 9. Matriz para Área 2	34
Tabla 10. Matriz para Área 3.....	34
Tabla 11. Resultados Área 1 P (Precisión), E (Exhaustividad), V-F1 (Valor-F1).....	35
Tabla 12. Resultados Área 2 P (Precisión), E (Exhaustividad), V-F1 (Valor-F1).....	35
Tabla 13. Resultados Área 3 P (Precisión), E (Exhaustividad), V-F1 (Valor-F1).....	35

Lista de Figuras

Figura 1. Área de Estudio	14
Figura 9. Agrupación de los Individuos.....	25
Figura 10. Estimación segmentación en grupos	26
Figura 11. Estimación segmentación en grupos	27
Figura 12. Extracción datos csv	27
Figura 14. Área 1	30
Figura 15. Área 2	31
Figura 16. Área 3	32
Figura 17. Ejemplo sub segmentación	32
Figura 18. Ejemplo sobre segmentación.....	33

Resumen

En este estudio se presentó una alternativa metodológica para la detección de árboles individuales mediante el uso de nubes de puntos LIDAR, donde se describió el paso a paso para el cálculo de los árboles y extracción de la información en formato shapefile con los softwares MicroStation Connect, terrasolid version 19, Inertial Explorer 8.7, Spatial Explorer. Se realizó un control de calidad a la metodología presentada usando tres áreas al azar en las cuales se ejecutó la identificación manual de los árboles identificando la sub segmentación y la sobre segmentación, los cuales se compararon bajo una matriz de confusión para medir la precisión en el rendimiento de la metodología, la información utilizada para este estudio fue tomada en China, en la provincia insular Hainan un área de 7,8ha y suministrada por la compañía Ewatt aerospace inc, la cual utilizo para la captura de información un sensor Riegl VUX LR de 820 KHz con una IMU LN200 en una plataforma M600-UAV de la casa DJI.

Palabras clave: agrupación, árboles, calidad, exhaustividad, LIDAR, sensor remoto, segmentación.

Abstract

In this study, a methodological alternative was presented for the detection of individual trees through the use of LIDAR point clouds, where the step by step for the calculation of the trees and extraction of the information in shapefile format through the use MicroStation Connect, Terrasolid version 19, Inertial Explorer 8.7, Spatial Explorer was described. Quality control was carried out on the methodology presented using three random areas in which the manual identification of the trees was performed, identifying the sub-segmentation and the over-segmentation, which were compared under a confusion matrix to measure the precision in the performance of the methodology, the information used for this study was taken in China, in the island province Hainan an area of 7.8ha and supplied by the company Ewart aerospace inc, which used a Riegl VUX LR sensor of 820 KHz with an LN200 IMU on a DJI M600-UAV platform.

Keywords: grouping, trees, quality, completeness, LIDAR, remote sensing, segmentation.

Introducción

El cálculo de árboles a gran escala ha sido una de las necesidades de las comunidades en el mundo a través del tiempo ya que es importante mantener el inventario forestal para poder tomar decisiones sobre los mismos como la tala de árboles o el mantenimiento de las especies, por lo cual se han desarrollado diferentes tipos de métodos para el mismo trabajo como el conteo manual, estadístico, fotografías aéreas, LIDAR, etc.

Teniendo en cuenta los avances en la tecnología, así como en los procesos topográficos, los sensores remotos LIDAR (Light Detection and Ranging) tienen un gran potencial para el monitoreo forestal debido a su calidad altamente relevante Kim et al (2009), estos han logrado satisfacer muchas necesidades en diferentes campos ya que al obtener gran cantidad de información y con gran precisión en un corto tiempo

El uso de datos de sensores LIDAR, también conocido como altimetría láser, está creciendo rápidamente y LIDAR está a punto de convertirse en una importante tecnología de teledetección para aplicaciones forestales Ronald E. McRoberts, Warren B. Cohen, Erik Næsset, Stephen V. Stehman y Erkki O. Tomppo (2010) ya que se obtiene información de gran calidad, dependiendo del tipo de metodología utilizado para el levantamiento y procesamiento de la información, las metodologías utilizadas se han generado a partir del conocimiento práctico de cada uno de los usuarios.

Por otra parte, uno de los grandes retos del levantamiento de información por medio de tecnología LIDAR en el campo de los inventarios de árboles es la precisión de la metodología ya que se tendría que realizar un conteo manual de todas las unidades arbóreas para verificar la información resultante lo cual conllevaría una gran cantidad de recursos y tiempo.

Así las cosas, en el presente documento se pretende proponer una metodología y evaluar la misma para el cálculo a gran escala de árboles individuales mediante el uso de

datos obtenidos a partir de tecnología LIDAR en la isla Hainan ubicada en la provincia insular de China. Donde se pueden encontrar grandes zonas montañosas cubiertas de bosques, esta evaluación será realizada a partir del conteo uno a uno de diferentes muestras seleccionadas al azar.

Los programas a utilizar para el desarrollo de la metodología y evaluación de la precisión de la información obtenida son de MicroStation Connect con terrasolid version 19, Inertial Explorer 8.7, Spatial Explorer, ArcGis 10.8. Los cuáles serán utilizados para el procesamiento y cálculo.

Estado del Conocimiento

De acuerdo a Ronald E. McRoberts, Warren B. Cohen, Erik Næsset, Stephen V. Stehman y Erkki O. Tomppo (2010) el principio de los LIDAR comenzó a principios de los años de 1960 y la detección moderna comenzó en los años de 1970 con la Administración Nacional de Aeronáutica y del Espacio (NASA) para medir propiedades atmosféricas, agua del océano, las copas de bosques y las capas de hielo.

Aunque esta tecnología LIDAR está en completo desarrollo y diferentes pruebas para ver la satisfacción de la misma en las áreas forestales se puede evidenciar de acuerdo Ronald E. McRoberts, Warren B. Cohen, Erik Næsset, Stephen V. Stehman y Erkki O. Tomppo (2010), que el primer estudio de muestreo forestal con lidar de perfiles aerotransportados se llevó a cabo en Delaware, EE. UU. (Nelson et al., 2003b , 2004), donde se tomaron 56 líneas de vuelo espaciadas a 1 km., el inventario forestal utilizando el LIDAR PALS y un diseño similar en la prefectura de Ehime, Japón (Tsuzuki et al., 2008). Luego en el 2006, el condado de Hedmark, Noruega, se tomaron 105 líneas espaciadas a 3 Km en el cual se construyeron modelos de regresión basada en la biomasa versus altura y densidad de dosel derivadas del LIDAR con dos estimadores basados en muestras, sin embargo, el uso de perfiles LIDAR se ha limitado a encuestas por muestreo y no ha encontrado ninguna aplicación en gestión de inventarios.

Por otra parte se evidencio que el Glas el cual es un LIDAR espacial a bordo del satélite ICESat, con el que se tomó los datos para un inventario forestal para el año 2008 Nelson et al. (2009), donde los resultados describieron dificultades donde los pulsos láser tomaban la misma información en copa de árboles bajo como en las pendientes de terreno haciéndose difícil interpretar y diferenciar los mismos, finalmente describiendo que las huellas deben ser más pequeñas para tener una mejor calidad de la información Harding & Carajabal (2005).

Por otra parte, se ha demostrado que los estimadores que utilizan modelos de regresión sintéticos para dominios pequeños, como rodales de bosques individuales para los que no se dispone de datos de campo, son muy precisos, aunque posiblemente con grandes sesgos (Särndal et al., 1992, págs. 410–411). La estimación no proporciona ningún medio para controlar el sesgo potencial y depende completamente de modelos correctamente especificados. Naesset (2007) informó que para las condiciones nórdicas, la aplicación de modelos sintéticos a las áreas de interés locales sin el uso de datos locales para ajustar las estimaciones, el sesgo a nivel del rodal es moderado, generalmente menos del 10% y en la mayoría de los casos menos del 5%. En un estudio de simulación, Andersen y Breidenbach (2007) compararon estimadores que utilizan modelos sintéticos sin ajuste local con estimadores modificados con ajustes locales e informaron que los datos de parcelas locales específicos del rodal pueden usarse para ajustar las estimaciones y reducir los efectos de la especificación incorrecta del modelo local. Sin embargo, la recopilación de datos de campo a nivel de rodal no es factible para el control de sesgos locales en inventarios operativos de manejo forestal de grandes áreas.

Marco Teórico

Los inventarios forestales de acuerdo a la FAO consisten en la recolección de información sobre los recursos forestales en una zona determinada, los cuales permiten la evaluación del estado actual y presenta un punto de partida para la planificación y el análisis de una gestión forestal sostenible, una de las técnicas más innovadoras actualmente en la toma de información son los sistemas LIDAR que en sus siglas en ingles *Light Detection and Ranging* y de acuerdo a National Oceanic and Atmospheric Administration U.S. Department of Commerce. Recuperado 5 de febrero de 2021, es un sistema remoto que utiliza la luz de forma de pulso laser para medir distancias hasta un objeto o una superficie, estos pulsos combinados de otras tecnologías recolectan una gran cantidad de información en tres dimensiones, el resultado de la toma de esta información es de gran calidad y precisión donde se pueden generar modelos digitales de superficie y ayuda a diversos grupos de profesionales a determinar con diferentes técnicas los inventarios de árboles.

Para poder determinar la ubicación espacial con gran calidad en los levantamientos LIDAR se ha determinado la utilización del sistema de posicionamiento (GNSS), el cual determina la posición y la altura de la aeronave, donde se captura información en el orden de 0,005 segundos conformando el trazado o trayectoria realizada por la aeronave durante todo el tiempo de captura, complementado la información con el sistema de navegación inercial (INS) que en sus siglas en inglés (Inertial Navigation System) y de acuerdo a *Coordinate Systems and Conventions. (2017). Performance of the Jet Transport Airplane* es el dispositivo que registra y detecta las variaciones en la aceleración, orientación y los giros de la aeronave los cuales en conjunto con los datos de posición registrados por el receptor o antena GNSS aportan la información posición con un alto grado de precisión. Estas

correcciones se integran a la trayectoria GNSS procesada adicionando los valores roll, pitch, yaw (*attitude*) los cuales son los ángulos de corrección en cada punto de la trayectoria con respecto a su sistema de coordenadas de navegación.

Es de tener en cuenta que la toma de información con LIDAR para el conteo de árboles se realiza desde aviones por lo cual es muy importante la planificación de los vuelos donde se debe tener en cuenta el área de la cobertura donde el objetivo principal es lograr obtener de la mayor cantidad de puntos por metro cuadrado, teniendo en cuenta áreas de traslapo que permitan la verificación de la calibración entre las líneas de vuelo y de ser necesario realizar procesos de calibración de los ángulos y desplazamientos, para esto se debe considerar tener trayectorias perpendiculares y horizontales en sentido adelante y atrás. Adicionalmente se deben adecuar a la frecuencia y el rango de altitud de operación del sensor para buscar la máxima penetración de la cobertura vegetal.

La Configuración y preparación de los datos consiste en la recolección de los datos crudos del sensor (LIDAR), la conversión en formatos estándar que permitan el procesamiento en el software a emplear esto puede variar dependiendo del sensor, los receptores, y los programas disponibles, se debe tener en cuenta el refinamiento del cálculo de trayectoria el cual se realiza mediante programas especializado INERTIAL EXPLORER 8.7 se realiza el refinamiento de la posición del ALS a través de correcciones diferenciales y la unión de los datos GNSS y la IMU (*Inertial Measurements Unit*) se incrementa la precisión de los datos LIDAR y cualquier sensor adicional que componga el sistema, a su vez los datos GNSS de una estación en tierra proporcionan un datum de referencia para dar un apropiado amarre al proyecto.

La información de datos LIDAR se trabaja en formatos LAS, de acuerdo al estándar definido por The American Society for Photogrammetry and Remote Sensing. el formato LAS es

un archivo binario que almacena información específica a la naturaleza del LIDAR sin que esta sea demasiado compleja para su gestión.

Una de las técnicas utilizadas para el conteo de árboles fue desarrollada por Hudak y Col. (2008) donde comparó un enfoque de bosques aleatorios y varios enfoques de vecinos más cercanos para estimar el área basal y la densidad de árboles utilizando datos LIDAR, el estudio de 14. Gobakken, Terje & Næsset, Erik. (2009) con la Evaluación de los efectos de los errores de posicionamiento y el tamaño de la parcela de muestra en las propiedades biofísicas del soporte derivadas de los datos del escáner láser en el aire, también se encuentra el estudio de Duncanson et al (2009) donde estima la altura del dosel a partir datos LIDAR, el estudio de Ståhl et al. (2010) donde utilizo un modelo de regresión con datos LIDAR para concluir con los valores de la biomasa, el estudio de McRoberts y col. (2010) donde el tema de las solicitudes de inventarios forestales nacionales.

En Colombia Cristian D. Martínez, Jhon E. Aunta, Jorge A. Valero (2013) realizaron un estudio sobre la aplicación de datos LIDAR en la estimación del volumen forestal en el Parque Metropolitano Bosque San Carlos en el cual a partir de la nube de puntos LIDAR interpolados se producen modelos digitales en un formato ráster para la creación de un modelo digital de vegetación y mediante modelos de regresión determinar la estimación del volumen forestal.

Objetivos

Objetivo general

Proponer una metodología para el cálculo individual de árboles a partir de datos capturados mediante un escáner láser aerotransportado (*Airborne light scanner, ALS*).

Objetivos específicos

- Definir una metodología para el conteo de árboles a partir de datos LIDAR
- Evaluar la precisión y la calidad de la metodología para la detección de árboles (segmentación correcta, sobreestimación o subestimación de individuos), mediante el control en áreas aleatorias donde se realiza conteo manual.

Metodología

En la estructuración del conteo de árboles a gran escala con sensores remotos LIDAR es de tener en claro que el procedimiento, de toma de datos y trabajo en oficina, se determina de acuerdo a las experticias sobre la información propia. A continuación, se describirá el procedimiento para la toma de información, procesamiento de la misma y su respectiva evaluación frente a muestras de conteo manual.

Área de Estudio

Área de 7,8ha ubicada en la provincia insular Hainan en China en el punto más meridional del país como se visualiza en la *figura 1*. Con áreas montañosas cubiertas de bosques, el área de estudio ha sido seleccionada debido a la disponibilidad de los datos y en especial las características de la captura.

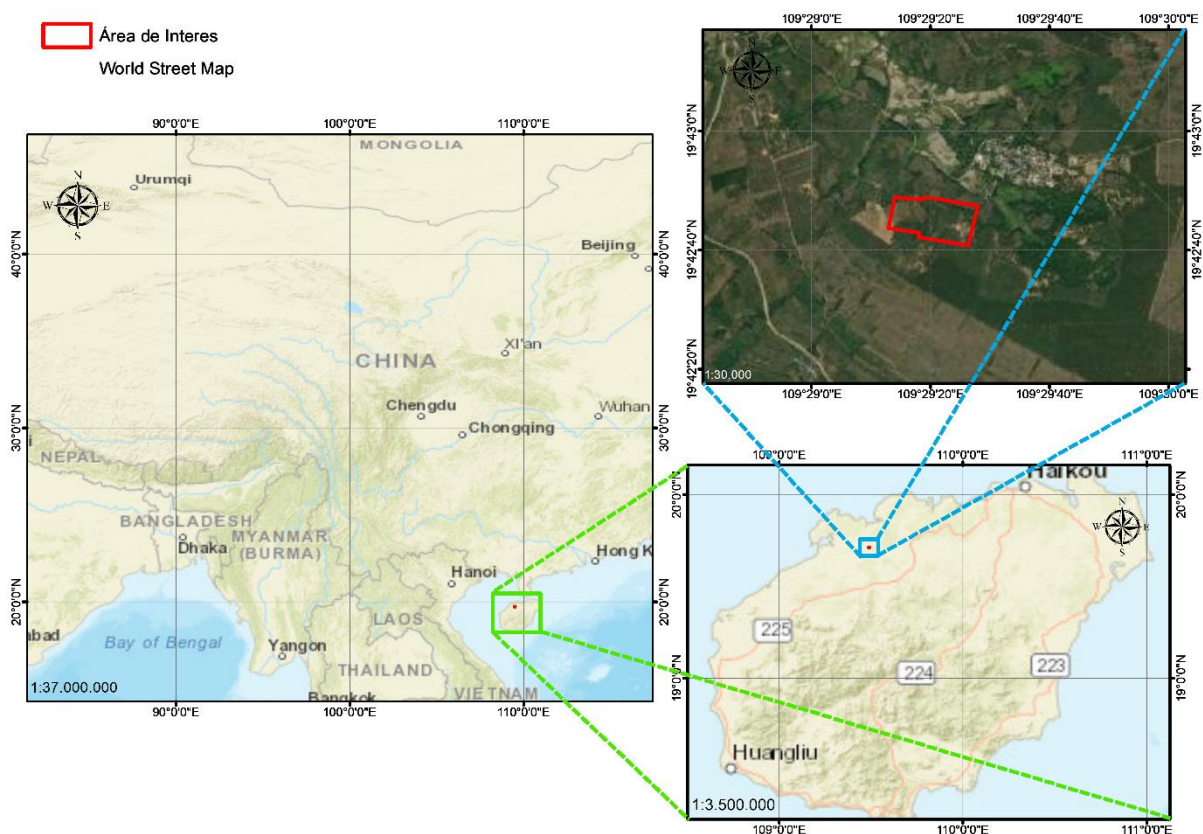


Figura 1. Área de Estudio

Sensor

Para la adquisición de esta información se utilizó el sensor rieggl VUX-LR, una plataforma DJI M600 Pro, la unidad de medida inercial IMU LN200C con una densidad de puntos 202 puntos/ m² por línea como se muestra en la *figura 2*.



Figura 2. Sensor Riegl VUX-LR Tomada en el área de estudio, Equipo Ewatt AeroSpace vuelo de calibración

Procesado de Datos LIDAR

El procesamiento de la información es de suma importancia para validar la información tomada en campo y ajustar la precisión de la misma, en los siguientes puntos se describirán los pasos del procesamiento de la información.

Extracción de datos crudos.

De debe extraer la información de los diferentes equipos utilizados, en primer lugar, la extracción de los datos crudos del sensor, la IMU, el sistema GNSS y de la información de posición de la referencia en tierra.

Es necesario obtener las coordenadas de la base GNSS de referencia en tierra que, dependiendo de las características del proyecto, la localización, extensión del área de interés, se determinará el número de bases, el tipo a utilizar (bases posicionadas o estaciones permanentes) y el tipo de procesado.

Trayectoria

Los dispositivos necesarios para determinar la posición de la aeronave provienen de dos dispositivos, el primero consta de una antena GNSS aerotransportado, el segundo una unidad de medición inercial (IMU *Inertial Measurement Unit*). La combinación de los datos de posición geográfica (XYZ) y los datos de las correcciones angulares referentes a la aeronave conocidos como (roll, pitch y yaw), derivan en el archivo correspondiente a la trayectoria el cual contiene los valores de marca de tiempo; localización Este, Norte, altura elipsoidal y posición Roll, pitch, yaw.

Al combinar la trayectoria procesada junto con los datos crudos provenientes de sensor se obtiene el archivo de la nube de puntos en formato. LAS (Laser Airborne Scanner) el cual contiene todas las mediciones realizadas por el sensor junto con los correspondientes ajustes en la posición dada por el sistema GNSS y las correcciones angulares (attitude) obtenidas por medio la IMU.

Clasificación y edición

Las nubes de puntos LIDAR dependiendo de la extensión del proyecto pueden trabajarse directamente o puede dividirse en bloques de áreas homogéneas más pequeñas, de

esta manera permitir al usuario un mejor desempeño en el proceso de clasificación, también facilitando la manipulación de los puntos al momento de la edición y a su vez una mejor gestión de los recursos de estación de trabajo.

La clasificación y edición de los puntos de terreno se realizará usando los módulos tarrascan, y terraModel, inicialmente realizando una clasificación del terreno teniendo en cuenta principalmente las características de las del terreno que se va a clasificar.

Una vez realizada la clasificación inicial se revisarán los resultados y la validación de los resultados de la clasificación del terreno, modificar o corregir posibles anomalías para posteriormente reclasificar la nube de puntos excluyendo puntos que no correspondan al terreno y continuar con la categoría de vegetación, así como cualquier otra clasificación que sea de interés en la nube de puntos.

Detección de Árboles

Posteriormente al procesado de la nube de puntos y su clasificación el procedimiento que permite la identificación, la localización, cómputo y posterior extracción de información correspondiente a los árboles del área de interés se deberá tener en cuenta los siguientes

Agrupación

El siguiente aspecto en el proceso para la detección de árboles individuales consiste en la asignación de un grupo identificador el cual se asigna a los puntos y es almacenado como de acuerdo a los criterios especificados como densidad de puntos, altura mínima, mínimo de puntos requeridos y la estructura lógica la cual se asigna haciendo uso de un modelo el cual se obtiene de árboles identificados en una categoría de clasificación especial. En estos modelos se debe tener en cuenta las características del árbol existente y serán tomados de diferentes ubicaciones dentro del área del proyecto, si hay diferentes

características o especies adicionalmente si hay variaciones en altura, volumen de la corona y otras características que puedan contribuir con una mejor detección.

Para optimizar los resultados proceso de agrupación de los puntos, se realizará el previo filtrado de manera que solo la categoría de vegetación sea tenida en cuenta durante el proceso.

Segmentación

Obtenidos y validados los grupos resultantes en la nube de puntos se realizará el procesamiento y computo los estadísticos como identificador de posición (x,y) de cada árbol, la altura del árbol, el diámetro de la corona.

Extracción

La extracción de los datos estadísticos se realizará en formato CSV y Shapefile, para posteriormente ser analizados en ArcGIS y continuar el proceso de control de calidad de la información extraída de nube de puntos que en este proceso se analizan.

Control de calidad

Para la determinación del nivel de precisión se analizaron los datos obtenidos mediante dos técnicas, el primer análisis espacial que permita comparar los elementos vectoriales obtenidos y la segunda mediante del valor esperado y el valor obtenido

Geográfico

Para determinar la precisión y la eficiencia en el proceso de la segmentación se tuvieron en cuenta una serie de áreas aleatorias, dentro de estas áreas se realizará el conteo y mapeo manual de los árboles durante el cual se obtendrá la información de posición, altura, y

diámetro de la corona, esta información también será exportada en formato CSV y finalmente a Shapefile para la comparación con la información obtenida del proceso de segmentación.

La metodología del control será la validación de en las áreas de control de los árboles identificados correctamente (Segmentado), la identificación de los casos en que un árbol no fue detectado, es decir no identificado como un individuo o más de un árbol fue detectado como uno solo (sub-segmentado), adicionalmente los casos donde un solo individuo fue segmentado como dos o más (sobre-segmentado).

Estadística

También se usaron los indicadores de la clasificación de la segmentación; (Segmentado), (sub-segmentado), (sobre-segmentado) para evaluar el desempeño del procesamiento basado en la matriz de confusión con la cual se podrá estimar la precisión del elemento (Goutte and Gaussier, 2005; Sokolova et al., 2006).

Tabla 1. Matriz de confusión para clasificación binaria

		Assignment z	
		+	-
Label ℓ	+	TP	FN
	-	FP	TN

De acuerdo a la tabla 1. TP son los árboles correctamente segmentados, FN los árboles sub segmentados y FP el sobre segmentado.

Bajo esta clasificación ser realizó un cálculo de la precisión (p) la exhaustividad (r) y la medida-F (f), con las siguientes formulas:

$$p = \frac{TP}{TP + FP}$$

$$r = \frac{TP}{TP + FN}$$

$$F = 2 \times \frac{r \times p}{r + p}$$

En donde (P) indica la precisión de la detección, (R) la tasa de detección en el procesamiento y (F) la precisión general del modelo considerando la precisión, la exhaustividad, [10].

Diagrama de Flujo Metodología

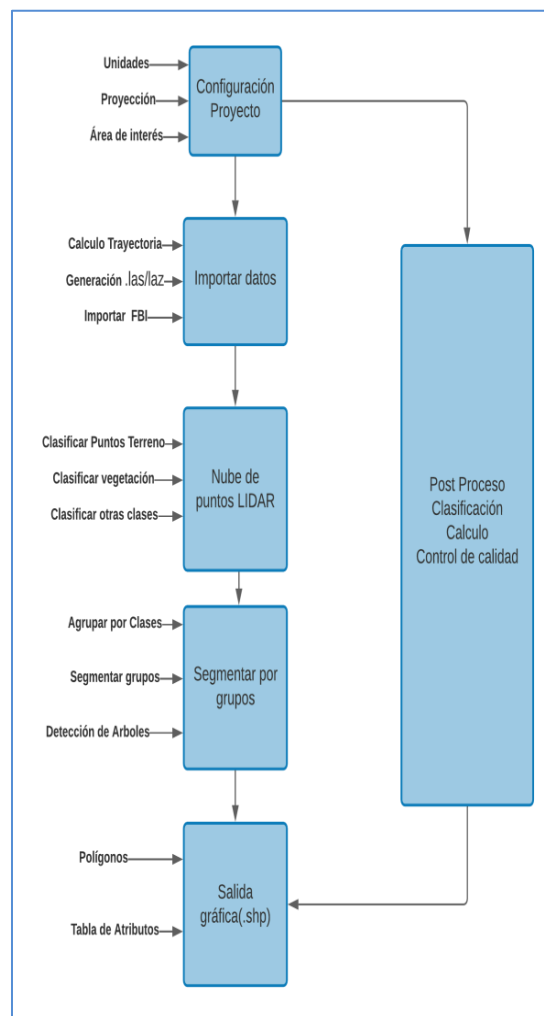


Figura 3. Metodología para el cálculo a gran escala de árboles individuales mediante el uso de datos LIDAR

Resultados

Los resultados obtenidos para el conteo de árboles a partir de los datos LIDAR se muestran en dos partes, la primera corresponde a la metodología establecida para el procesamiento preliminar los resultados obtenidos para la nube de datos LIDAR, el cálculo y ajuste de la trayectoria, la clasificación y filtrado, los modelos manualmente generados para la posterior identificación de individuos arbóreos, los resultados de la localización individual y las comparativa realizada con las áreas de control de calidad.

Definición de la Metodología

Con esta serie de pasos se describe un procedimiento adecuado para el conteo de árboles a gran escala a partir de datos LIDAR.

Paso 1. Adquisición de la información

La captura de los datos fue realizada el 07 de junio del 2019 en la región China de Hainan por parte de la compañía Ewatt aerospace inc. como parte del programa de la compañía para el desarrollo y la implementación de soluciones en la gestión forestal y producción de maderables. Los datos usados para el desarrollo de la presente metodología fueron suministrados por Ewatt aerospace inc. propietaria de los datos, para su uso en el presente estudio.

La plataforma que se utilizó fue un sensor Riegl VUX LR de 820 KHz con una IMU LN200 en una plataforma M600-UAV de la casa DJI. La captura se realizó a una altura de vuelo de 115 metros y se obtuvo una densidad superior a los 200 puntos por metro cuadrado, en las áreas de traslape se alcanzaron densidades superiores a 400 puntos por metro cuadrado.

Paso 2. Cálculo de la trayectoria

El procesado de la trayectoria es el proceso mediante el cual se obtiene la información de la posición y la orientación, en conjunto con la estación de referencia en tierra.

El programa utilizado para el cálculo es Inertial Explorer, el cual mediante un proceso diferencial robusto combina las observaciones GNSS y las mediciones de la unidad Inercial INS.

En el cálculo de la trayectoria se debe tener en cuenta que factores de calidad como la cobertura en la lectura del receptor GNSS IMU y la estación de referencia en tierra como se muestra en la *figura 4*, Numero de satélites de 7 para el cálculo de posición, ángulos de giro de la aeronave de entre -0.25 y + 0.25 arco minuto como se muestra en la *figura 5*, y una estimación de la precisión de la posición por debajo de los 0.015m en XY y menor a 0.020 m en altura. (Zhang et al., 2019) como se muestra en la *figura 6*.

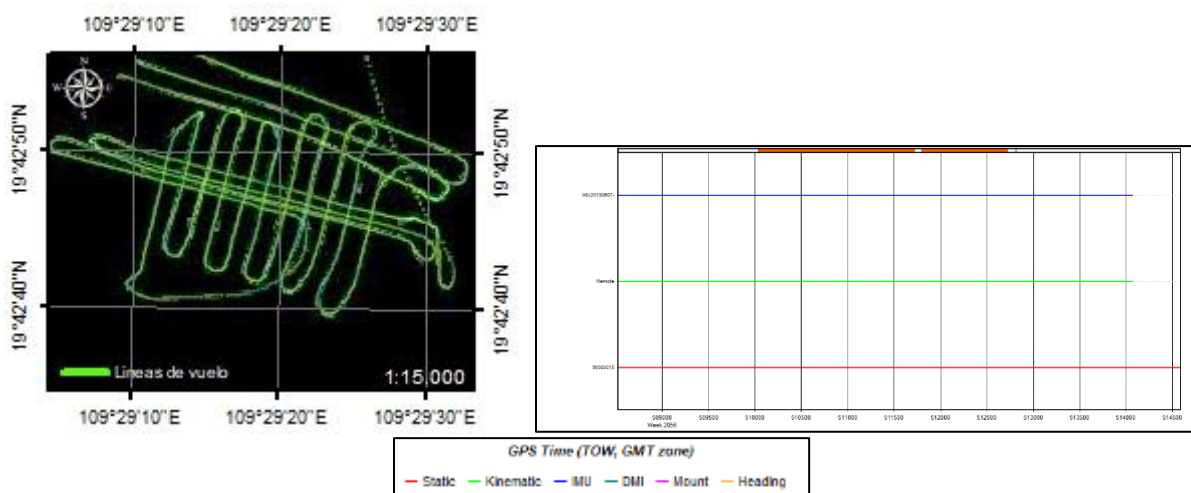


Figura 4. Trayectoria y cobertura del receptor GNSS, IMU y Base de referencia

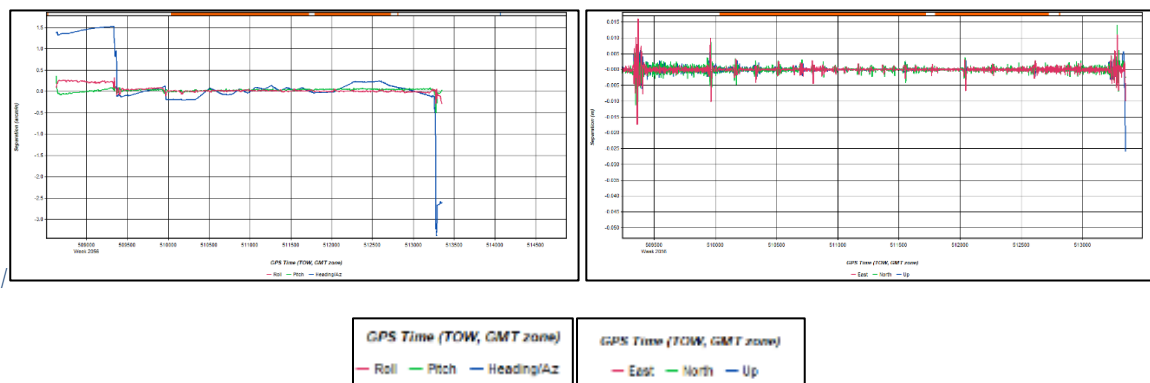


Figura 5. Control calidad Procesamiento trayectoria

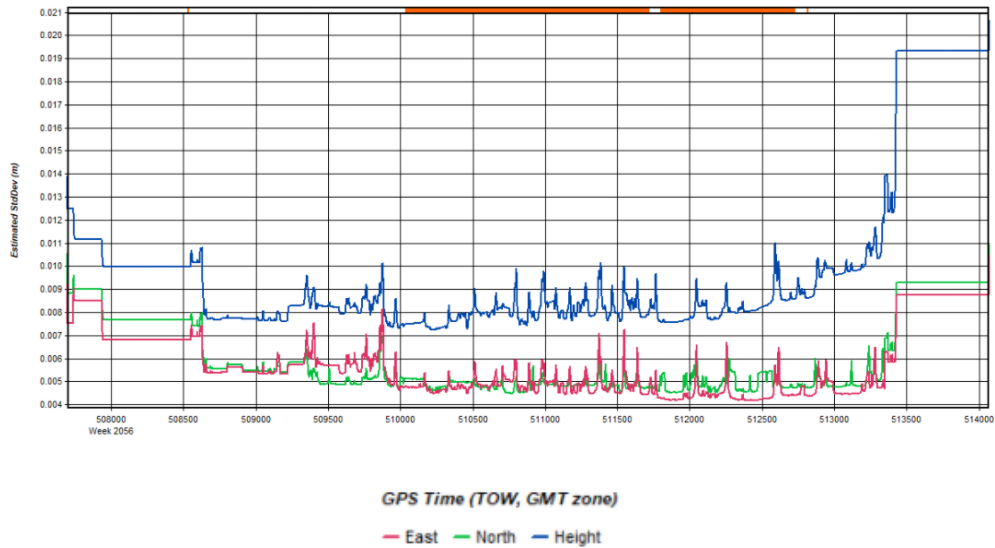


Figura 6. Control calidad estimación de la posición de la trayectoria

Ajuste y clasificación de la nube de puntos

Finalizado el proceso del cálculo de la trayectoria se integra con la información del sensor para obtener una nube de puntos georeferenciada de 15.625.774 puntos.

Posteriormente se realizó un proceso de filtrado con el fin de remover puntos comúnmente denominados como ruido. Finalizado el filtrado la nube de puntos fue clasificada según los códigos de clasificación definidos por la ASPRS (American Society for Photogrammetry and Remote Sensing) en el caso de los puntos de terreno, aplicando el modelo de densificación basado en triángulos (distancias y ángulos máximos) como se muestra en la *figura 7*.

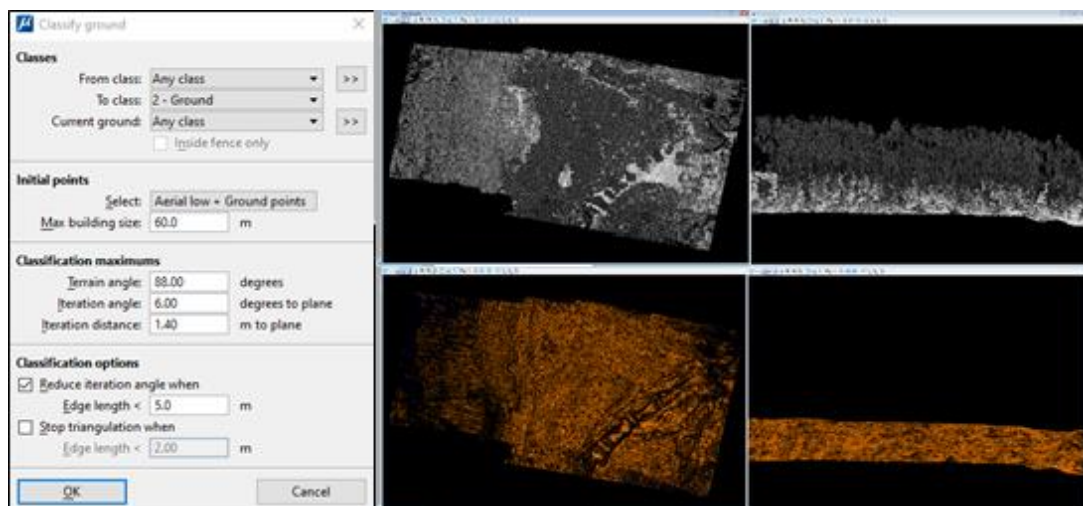


Figura 7. Clasificación Nube de Puntos

Posteriormente se realizó una clasificación preliminar de la vegetación basada en alturas sobre los puntos sobre el terreno como se muestra en la *figura 8*, identificando vegetación baja, media y alta.

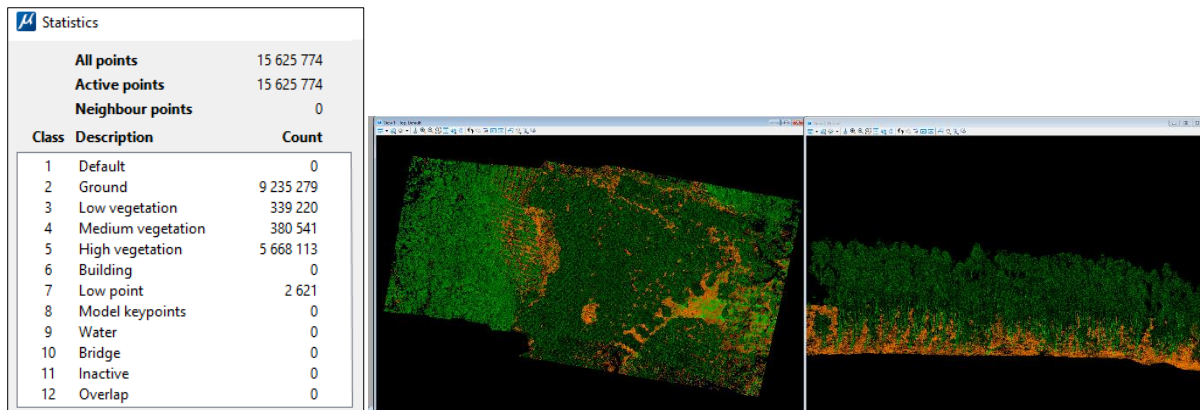


Figura 8. Clasificación Preliminar

Agrupación

Para este procedimiento se identificaron algunos modelos que sirvieron como parámetro para la agrupación automática, así mismo elemento como la densidad mínima de los grupos, área mínima, altura mínima, diámetro máximo y longitud mínima del tronco.

Un grupo de 48 modelos fueron utilizados como datos de entrenamiento los cuales fueron categorizados por altura; un primer grupo de árboles de altura menor a 5 metros, el segundo grupo de alturas entre 5 y 10 metros, el tercer grupo entre 10 y 15 metros por último un grupo para los árboles con altura mayor a 15 metros como se muestra en la *figura 9*.

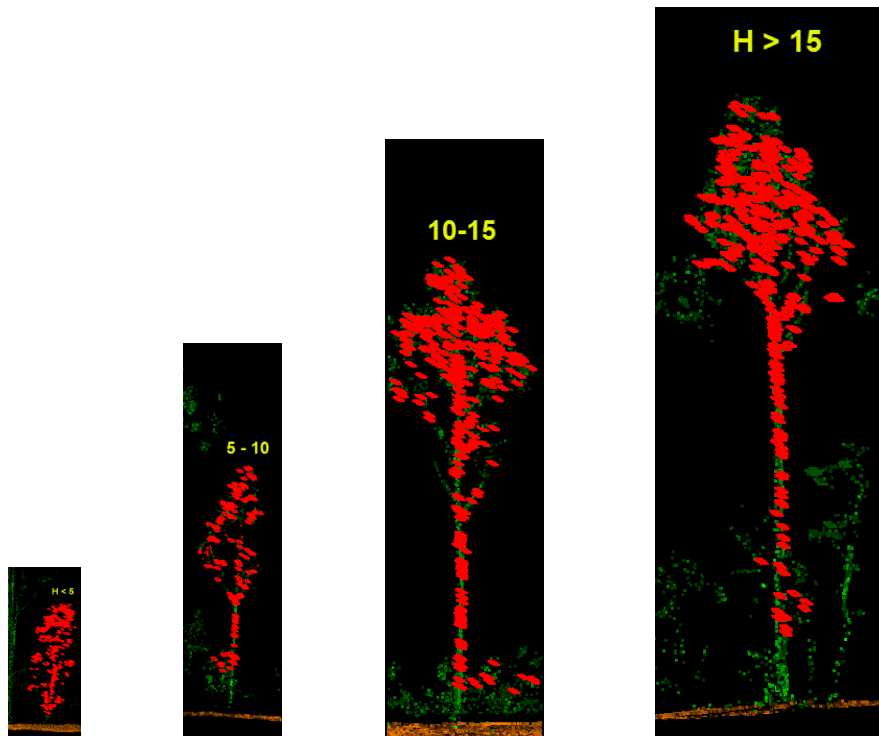


Figura 2. Agrupación de los Individuos

En estudios previos el impacto en la calidad y precisión de la detección se ve afectado por la necesidad de determinar tamaño de grilla para la creación de modelos de altura de la vegetación, seleccionar un tamaño de grilla pequeño mejora la resolución del modelo lo que permite el uso de mayor cantidad de puntos, que incluyen punto debajo de las capas superiores de vegetación lo cual produce un modelo más detallado, pero así mismo induce a la sobre estimación de árboles (sobre- segmentación) (Hyypä et al., 2001).

La segmentación basada en los parámetros como la altura y los demás descritos en el dato de entrenamiento permitió la agrupación de los puntos preclasificados como vegetación en una nueva clasificación de acuerdo con los grupos de puntos resultantes los cuales representan la segmentación de dicha vegetación en una nueva clasificación de árboles individualizados e identificados en grupos.

Se realizó el proceso mediante la estructuración de una rutina diseñada en 4 iteraciones con parámetros determinados a partir de los datos de entrenamiento, esto con el fin de disminuir la sub segmentación es decir la omisión o determinación de dos o más

individuos como uno solo o en caso contrario la sobre segmentación en la cual un individuo es segmentado en diferentes grupos, en ambos casos conluciendo a una disminución en la precisión de la estimación del número de árboles.

Tabla 2. Parámetros

Parámetros					
Iteración	Altura mínima (m)	Area mínima (m ²)	Tolerancia (m)	Altura mínima del tronco (m)	Densidad mínima (punto/ m ²)
1	1	2,5	0,2	1	20
2	5	4,5	0,2	2,5	35
3	10	8	0,2	4,5	45
4	15	10	0,2	6,5	60

Después de ejecutar el procesamiento se aprecia la segmentación en grupos, estos grupos están a su vez clasificados mediante un identificador único que en una primera instancia permite tener la estimación del número de árboles presentes en la nube de puntos como se muestra en la *figura 10*.

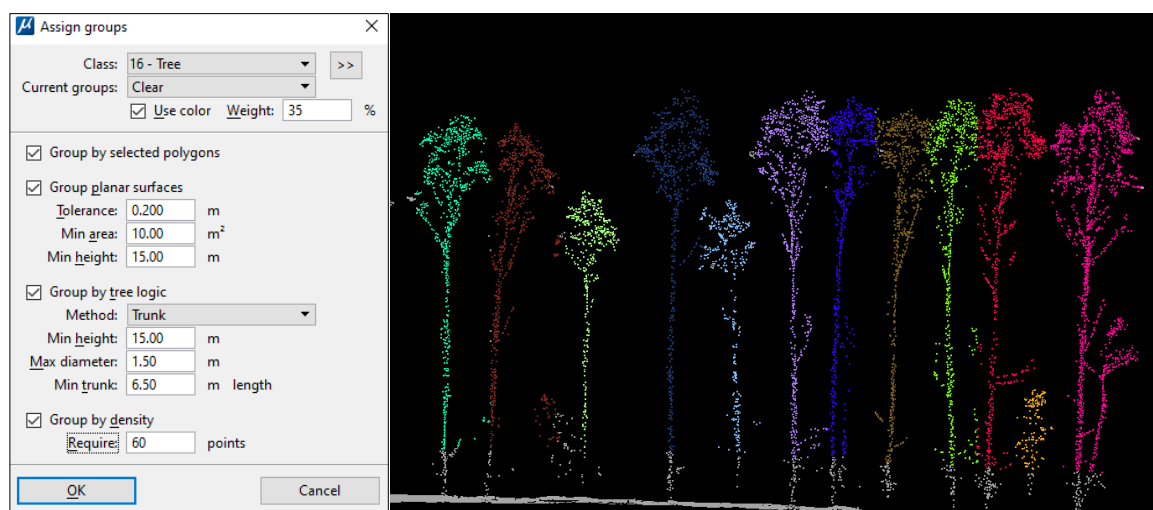


Figura 30. Estimación segmentación en grupos

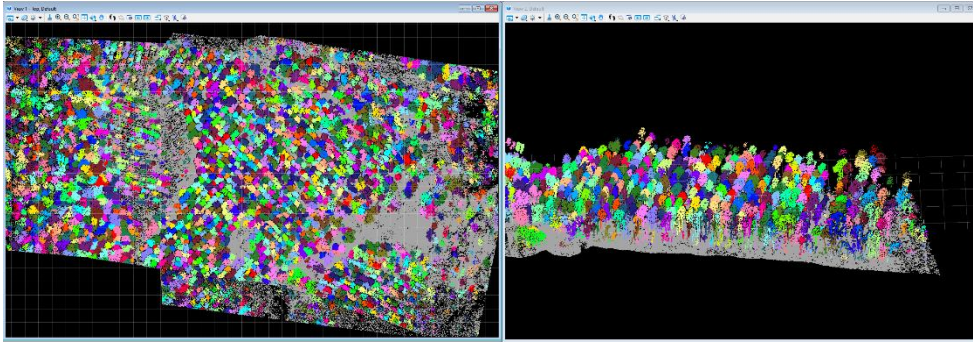


Figura 41. Estimación segmentación en grupos

Con la segmentación realizada como se aprecia en la figura 11. Se definieron los elementos a calcular para cada árbol, de los cual se obtuvo un polígono convexo que muestra las características para cada uno de los puntos agrupados los cuales se consideraron como árboles en la nube de puntos en total 3864 árboles calculados.

Extracción de los resultados de la segmentación

Para la validación de los elementos realizó la extracción en una base de datos formato CSV de los polígonos correspondientes a la nube de puntos segmentada por grupos, en la cual están registrados los atributos de cada uno de los elementos donde se especifica identificador del árbol, posición X, posición Y, altura del árbol, y el diámetro de la corona, como se aprecia en la *figura 12*.

Table						
Deteccion_Arboles						
FID	Shape *	tree_id	x	y	height	crown_diam
0	Polygon	1	970787.892274	2185928.97817	14.864674	4.261772
1	Polygon	2	970642.985277	2185957.42673	17.218239	5.555008
2	Polygon	3	970612.247205	2185935.62496	17.159492	4.633842
3	Polygon	4	970746.68005	2185845.69707	16.2335	5.304741
4	Polygon	5	970784.165158	2185973.44299	16.924234	6.452566
5	Polygon	6	970733.472332	2185963.26951	13.139948	3.028437
6	Polygon	7	970830.276322	2185814.72115	15.262962	5.372962
7	Polygon	8	970751.721504	2185999.69852	17.167461	4.914424
8	Polygon	9	970684.876007	2185951.42664	7.355546	2.765189
9	Polygon	11	970610.424186	2185896.50424	12.923723	5.480144
10	Polygon	13	970872.227821	2185819.06009	13.849783	3.19495
11	Polygon	15	970612.182346	2185875.15934	14.356428	5.848976
12	Polygon	16	970832.673839	2185910.6673	13.502228	3.811096
13	Polygon	17	970855.00427	2185923.76405	13.572222	4.517844
14	Polygon	18	970784.493655	2185961.65136	16.186161	3.515101
15	Polygon	19	970879.572903	2185922.5505	14.473517	3.740237
16	Polygon	20	970959.231276	2185889.42136	11.391773	3.919551
17	Polygon	21	970906.630934	2185803.84114	13.123871	3.652225

Figura 52. Extracción datos csv

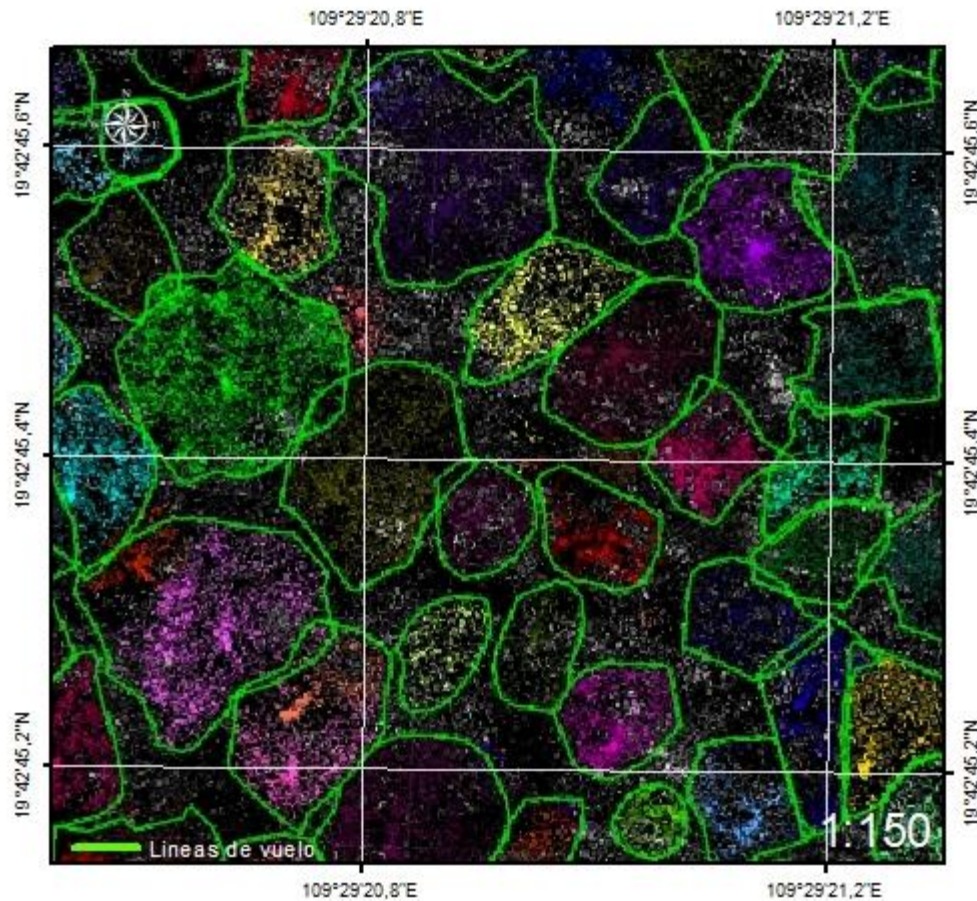


Figura 13. Generación de Polígonos

Evaluación de la Precisión y Calidad

Generado el cálculo de los árboles se evalúa la precisión y calidad de la metodología anteriormente descrita de acuerdo al muestreo realizado.

Áreas de muestreo

Para el control de calidad se tomaron 3 áreas de control, con las siguientes dimensiones área (1) 0.4532 ha, área (2) 0.2979 ha y área (3) 0.2230 ha, teniendo en cuenta la densidad de la muestra recolectada y para garantizar la calidad de la información.

Dentro de cada una de las muestras tomadas se hizo el conteo manual de los árboles, considerando las siguientes características de cada uno de los individuos, altura con respecto

al terreno, geometría de la corona, donde se identificaron 564 Árboles Control en el área (1), 171 árboles en el área (2) y 161 árboles en el área (3) como se muestra en la *tabla3*.

Tabla 3. Resultados Iniciales

Árboles control	Procesamiento
564	560
171	169
161	166

Uno de los mayores desafíos para realizar el conteo individual de árboles consiste en lograr identificar exitosamente los grupos de árboles con similares alturas y alta densidad ya que pueden fácilmente ser detectados como un único árbol (sub segmentación) (Mongus & Žalik, 2015a).

En una comparación inicial realizada mediante un análisis espacial teniendo en cuenta la posición de los polígonos generados en el conteo manual (Árboles Control) y los generados mediante el procesamiento se evidenció una mínima variación entre los dos (*Tabla 5*). A priori la cantidad de polígonos indicó una alta tasa de acierto sin embargo una revisión más exhaustiva muestra las coincidencias entre los árboles que fueron correctamente detectados, a su vez se lograron identificar los árboles que fueron sub segmentados y los sobres segmentados para cada una de las áreas de control.

Los casos de su segmentación y sobre segmentación en los estudios en los que se usa modelos derivados de LiDAR se producen con pequeñas variaciones en el modelo de vegetación (Hyypä et al., 2001). En el caso del presente estudio al incluir los parámetros de altura mínima y altura mínima del tronco en conjunto con una buena densidad directamente

en la nube de puntos contribuyen al mejor desempeño y la obtención de mejores tasas de acierto. (Mongus & Žalik, 2015a).

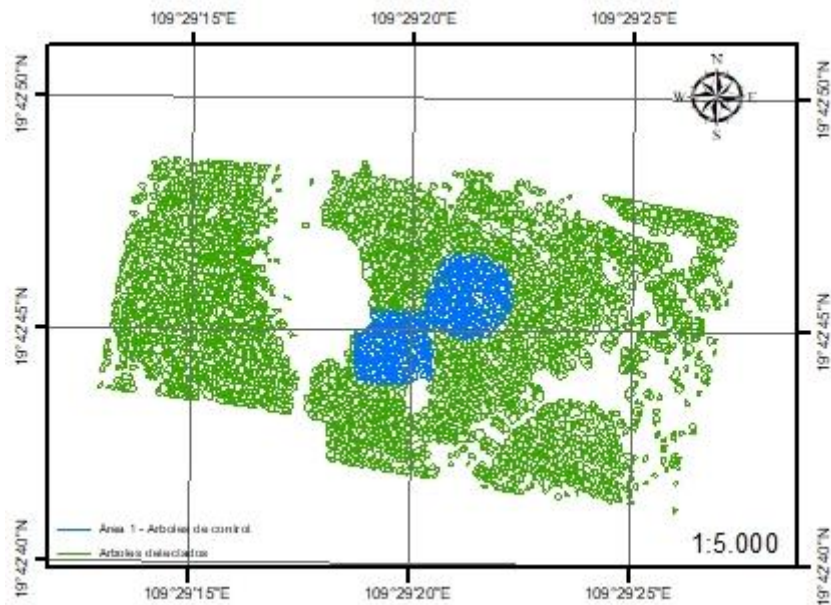


Figura 64. Área 1

Tabla 4. Resultados Área 1

Grupo	Altura (m)	Árboles control	Detectados Correctamente	Sobre-Segmentados	Sub-Segmentados
1	< 5	46	42	3	1
2	5-10	55	40	11	3
3	10-15	232	185	17	29
4	> 15	231	196	13	20
		564	463	44	53

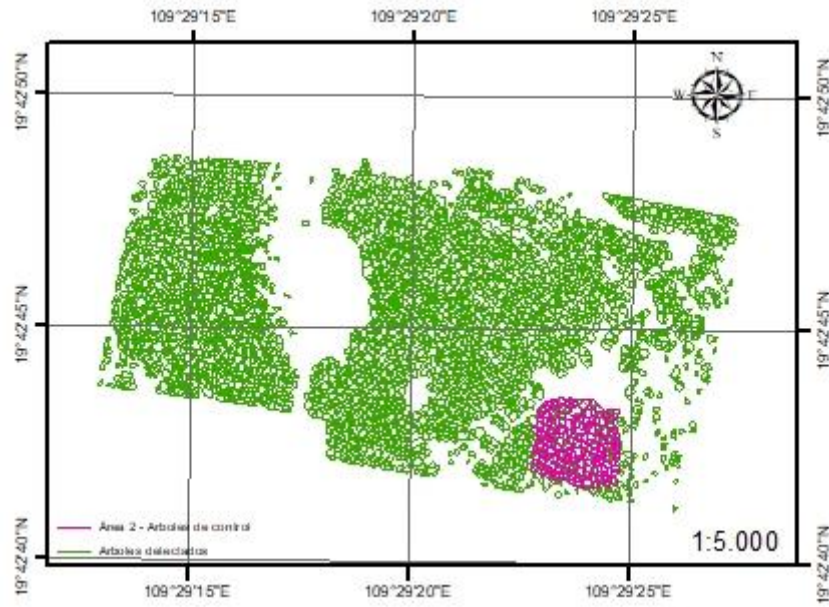


Figura 75. Área 2

Tabla 5. Resultados Área 2

Grupo	Altura (m)	Árboles control	Detectados Correctamente	Sobre Segmentados	Sub Segmentados
1	< 5	2	1	0	0
2	5-10	14	9	3	2
3	10-15	123	88	1	33
4	> 15	32	18	0	14
		171	116	4	49

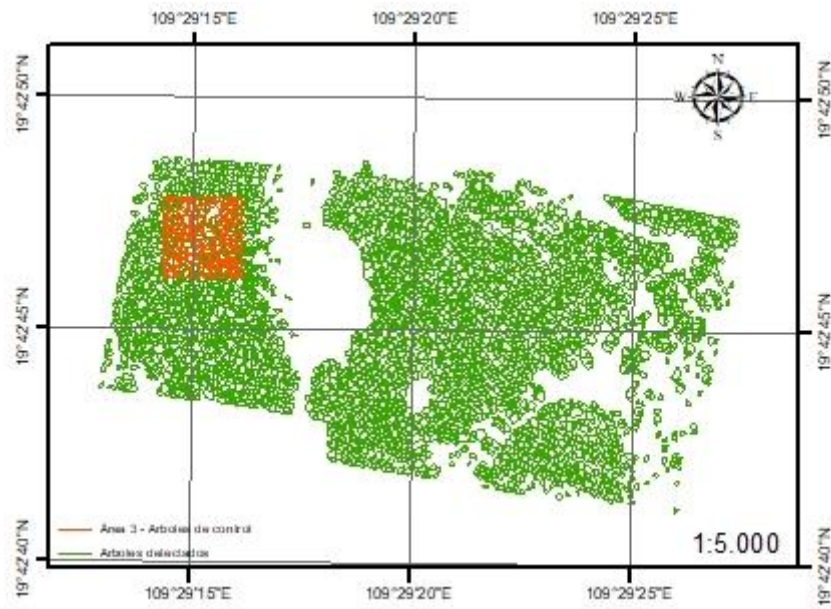


Figura 86. Área 3

Tabla 6. Resultados Área 3

Grupo	Altura (m)	Árboles control	Detectados Correctamente	Sobre Segmentados	Sub Segmentados
1	< 5	0	0	0	0
2	5-10	0	0	0	0
3	10-15	5	3	2	0
4	> 15	156	145	16	0
		161	148	18	0

Sub segmentación y Sobre segmentación

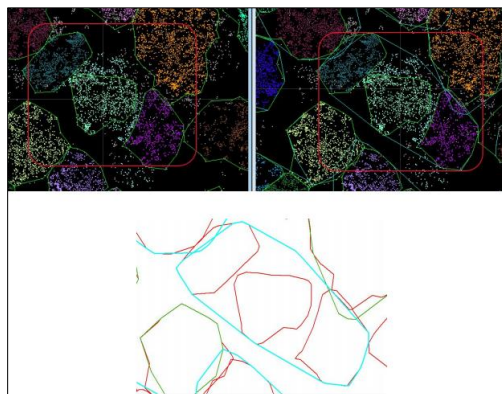


Figura 97. Ejemplo sub segmentación

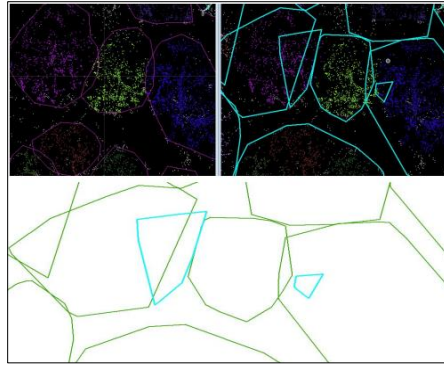


Figura 108. Ejemplo sobre segmentación

Caracterización

Una vez se realizó la caracterización del tipo de segmentación realizada en el procedimiento se obtuvo una clasificación, con la cual se realizó la evaluación de rendimiento de la segmentación mediante la metodología planteada, con el uso de una matriz de confusión para clasificación binaria como se muestra en la *tabla 7*.

Tabla 7. Modelo de Matriz de confusión

		Predicción	
		0	1
Realidad	0	TN	FP
	1	FN	TP

Donde:

0: No es un árbol

1: Es un árbol

True Positive [TP] cuando el árbol ha sido correctamente detectado, False Negative [FN] no identificado como un individuo o más de un árbol fue detectado como uno solo (sub-segmentado); False Positive [FP] los casos donde el árbol no existe o un solo individuo fue segmentado como dos o más (sobre-segmentado).

Tabla 8. Matriz para Área 1

		Predicción	
		0	1
Realidad		0	1
	0	-	44
	1	53	463

Tabla 9. Matriz para Área 2

		Predicción	
		0	1
Realidad		0	1
	0	-	4
	1	49	116

Tabla 10. Matriz para Área 3

		Predicción	
		0	1
Realidad		0	1
	0	-	18
	1	3	161

La exhaustividad (Recall)

Con las matrices de los resultados *Tabla8*, *tabla9* y *tabla10*. De las áreas de control se logró establecer la precisión de la metodología, para lo cual se calculó la precisión (*Precision*) con el fin de visualizar que porcentaje de árboles fueron correctamente identificados mediante la metodología. La exhaustividad (*Recall*) que es la tasa de acierto que se obtuvo la metodología, es decir de los árboles en las áreas de control cuantos realmente fueron detectados. Finalmente, el Valor-F1 (*F1-score*) se calculó para determinar la precisión

general combinando tanto la exhaustividad como la precisión como se muestra en las *tablas 12, 13 y 14*.

Tabla 11. Resultados Área 1 P (Precisión), E (Exhaustividad), V-F1 (Valor-F1)

Grupo	Altura (m)	P	E	V-F1
1	< 5	93%	98%	95%
2	5-10	78%	93%	85%
3	10-15	92%	86%	89%
4	> 15	94%	91%	92%
	sumatoria	91%	90%	91%

Tabla 12. Resultados Área 2 P (Precisión), E (Exhaustividad), V-F1 (Valor-F1)

Grupo	Altura (m)	P	E	V-F1
1	< 5	100%	100%	100%
2	5-10	75%	82%	78%
3	10-15	99%	73%	84%
4	> 15	100%	56%	72%
	sumatoria	97%	70%	81%

Tabla 13. Resultados Área 3 P (Precisión), E (Exhaustividad), V-F1 (Valor-F1).

Grupo	Altura (m)	P	E	V-F1
1	< 5			
2	5-10			
3	10-15	60%	100%	75%
4	> 15	90%	100%	95%
	sumatoria	89%	100%	94%

Se obtuvo una precisión general del 91% para el área de control (1) *tabla 11*; 81% para el área de control (2) *tabla 12*; y 94 % para el área de control (3) *tabla 13*. El desempeño en el área 1 es notable ya que en esta área la densidad de árboles es aproximadamente tres veces (564) más que la densidad de árboles presente en las dos restantes áreas de control (171) y (161). Puede atribuirse a la alta densidad de puntos (441 puntos por metro cuadrado), con lo cual se obtuvo una excelente penetración a nivel de terreno y de la estructura del árbol, dado que la altura del tronco con respecto al terreno fue uno de los parámetros considerados en la determinación de los grupos y la segmentación. Se ha evidenciado que la eficiencia para detección de los troncos de los árboles está relacionada con altas densidad de puntos. (ÖZDEMİR et al., 2021).

Adicionalmente se observó que para el área 1 el 82% eran árboles de más de 10 metros de altura fueron estos los que mayor tasa de éxito tuvieron en la metodología.

Los elementos como el espaciamiento entre los árboles, los ángulos de incidencia del sensor, la penetración al suelo favorece positivamente la clasificación, el agrupamiento y por ende la detección del árbol en este determinado grupo. (Mongus & Žalik, 2015)

En el área de control número 2 se observó que el 91% de los árboles son de más de 10 metros de altura, en el rango de los árboles con alturas entre 10 y 15 metros se presentaron 33 casos de árboles no detectados o sub segmentación este posiblemente a la separación de los árboles, y una menor densidad de puntos.

En el área 3 se mostró una excelente respuesta en la detección de árboles, 97% de estos por encima de los 15 metros de altura, una exhaustividad del 100 % muestra que los modelos y la parametrización para esta zona ha sido la adecuada en conjunto con una alta densidad de la zona por encima de 400 puntos por metro cuadrado.

Conclusiones

La metodología desarrollada describe paso a paso el procedimiento para el cálculo de árboles a gran escala, detallando la utilización de todos los datos tomados y evidenciando que esta metodología utiliza métricas únicas.

El desarrollo de la metodología ha sido exitoso para la detección individual de árboles a gran escala, pero es importante señalar que las buenas prácticas en el manejo en este tipo de datos son fundamentales para alcanzar la precisión esperada, desde la planificación del vuelo, la correcta calibración y corrección de desajustes que guíen a tener a disposición la mejor calidad posible, son elementos indispensables para adelantar procedimientos similares a los presentados en este documento.

De los 896 árboles de control se detectaron correctamente un total de 727, entre los grupos por altura menores a 5 metros se detectaron 43 de 48 presentes en la zona de control. Entre 5 y 10 metros la detección fue de 49 de 69 identificados, el número detectado de la categoría de 10 a 15 metros fue de 276 de 360, finalmente en la categoría de árboles de altura superior a 15 metros se detectaron 359 de 419.

Los resultados en la tasa de detección en las zonas de control obtuvieron una precisión promedio de 89% (valor -F1) con esto se indica que los valores de sub segmentación y los de sobre segmentación fueron bajos abriendo una nueva la posibilidad para el mapeo detallado de árboles individuales sobre grandes áreas, además proporciona información sobre la estructura espacial para la gestión forestal, la estimación de las existencias de carbono.

La densidad de los puntos es un factor que incide de manera importante en la adopción de esta metodología, es de resaltar que como se planteó al inicio del proyecto una densidad superior a los 200 puntos es la base para mejorar la calidad de la detección usando la metodología adoptada.

Para obtener mayor tasas de detección deben considerarse apropiados datos de entrenamiento teniendo en cuenta la variabilidad de las formas, especies y alturas presentes en la zona de estudio y la aplicación de las iteraciones para la clasificación y segmentación basada en estos datos de entrenamiento puede contribuir a menores índices de sub segmentación y sobre segmentación, especialmente en áreas donde se presenta mayor variabilidad en la estructura física de los árboles presentes en el área de estudio.

Los resultados obtenidos mediante la matriz de confusión y las métricas de Precisión, Exhaustividad y el Valor F permitieron estimar que la metodología aplicada presenta muy buenos indicadores para el uso en estudios forestales.

Recomendaciones

El uso de la tecnología LIDAR ha mostrado grandes posibilidades para estudios de gestión en el ámbito forestal, sin embargo para continuar con el desarrollo de estas metodologías se puede considerar para futuros estudios la integración de LIDAR aerotransportado y LIDAR terrestre lo que permitirá tener una mejor estructura en la nube de puntos integrada que permita combinar con mayor precisión parámetros como la máxima altura, la altura del tronco y en general mejorar la descripción geométrica y la estructura observada de los árboles en la zona de estudio, al contar con mayores densidades podrá determinarse parámetros adicionales fenotípicos y biofísicos que incluso permitan llegar a la identificación de las especies observadas.

Bibliografía

- Aguilar, F. J., & Mills, J. P. (2008). Accuracy assessment of lidar-derived digital elevation models. *The Photogrammetric Record*, 23(122), 148-169. <https://doi.org/10.1111/j.1477-9730.2008.00476.x>
- ArcGIS Resource Center. (s. f.). Recuperado 28 de febrero de 2021 de <https://resources.arcgis.com/>
- Coordinate Systems and Conventions. (2017). *Performance of the Jet Transport Airplane: Analysis Methods, Flight Operations and Regulations*, 597-600. <https://doi.org/10.1002/9781118534786.app3>
- Gobakken, Terje & Næsset, Erik. (2009). Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*. 39. 1036-1052. 10.1139/X09-025.
- Goutte, C., and E. Gaussier. (2005). “A Probabilistic Interpretation of Precision, Recall and F-score, with Implication for Evaluation.” In 27th European Conference on IR Research, Santiago De Compostella. Spain, March 21–23. doi:10.1080/09553008714551221.
- Inertial Explorer®. (s. f.). NovAtel. Recuperado 28 de febrero de 2021 de <https://novatel.com/products/waypoint-software/inertial-explorer>
- LASer (LAS) File Format Exchange Activities – ASPRS. (s. f.). Recuperado 1 de febrero de 2021, The American Society for Photogrammetry and Remote Sensing. de <https://www.asprs.org/divisions-committees/lidar-division/laser-las-file-format-exchange-activities>
- Lidar — Light Detection and Ranging. (s. f.). (5 de febrero de 202), National Oceanic and Atmospheric Administration U.S. Department of Commerce., de <https://oceanservice.noaa.gov/facts/lidar.html>
- Martínez Tobón, C. D., Aunta Duarte, J. E., & Valero Fandiño, J. A. (2013). Aplicación de datos LiDAR en la estimación del volumen forestal en el parque metropolitano

bosque San Carlos. *Ciencia e Ingeniería Neogranadina*, 23(1), 7.
<https://doi.org/10.18359/rcin.229>

Mongus, D., & Žalik, B. (2015). An efficient approach to 3D single tree-crown delineation in LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108, 219–233. <https://doi.org/10.1016/j.isprsjprs.2015.08.004>

Ronald E. McRoberts, Warren B. Cohen, Erik Næsset, Stephen V. Stehman & Erkki O. Tomppo (2010) Using remotely sensed data to construct and assess forest attribute maps and related spatial products, *Scandinavian Journal of Forest Research*, 25:4, 340-367, DOI: 10.1080/02827581.2010.497496

Scandinavian Journal of Forest Research, (27 de Abril 2010)25: 4 , 340-367, DOI: 10.1080 / 02827581.2010.497496

SIAC. (s. f.). www.siac.gov.co. Recuperado 20 de febrero de 2021 de <http://www.siac.gov.co/inventario-forestal-nacional>

Terrasolid. (s. f.). *Terrasolid - Software For Point Cloud and Image Processing*. Recuperado 16 de marzo de 2021 de <https://terrasolid.com/>

Zhang, Q., Niu, X., & Shi, C. (2019). Assessment of the effect of GNSS sampling rate on GNSS/INS relative accuracy on different time scales for precision measurements. *Measurement*, 145, 583–593. <https://doi.org/10.1016/j.measurement.2019.05.104>

Anexos

- Autorización de uso de datos.



EWATT = 易瓦特
Diamond Bar, CA 91765 USA

THIS AUTHORIZATION AGREEMENT is made and entered into as of April 20, 2021 between Ewatt Aerospace Inc with offices at 3384 Diamond Bar, CA 91765 Los Angeles, and William Saenz Blanco, with offices at CLL 2e 1a 85 Bogota, Colombia.

Purpose. Authorization format for the use of data or productions in general through this document as the owner of the rights of the works **Ewatt Forestry LIDAR Data** Ewatt Aerospace Inc authorizes to use said works exclusively for use in the study called *Methodology and accuracy assessment of Detection of individual trees using airborne lidar data*. In the same way, Ewatt Aerospace Inc guarantees that is the owner of the rights and therefore Ewatt Aerospace Inc can grant this authorization.


This authorization does not imply transfer of copyright and additionally guarantees, the mention of the author's name. In the event that the work involves the image of a person, the user will be responsible for requesting the required authorizations and will exonerate EWATT AEROSPACE INC from any requirement for eventual claims to the right to the image.


No Obligation. Nothing herein shall obligate either party to proceed with any transaction between them, and each party reserves the right, in its sole discretion, to terminate the discussions contemplated by this Agreement.

No Warranty. ALL CONFIDENTIAL INFORMATION IS PROVIDED "AS IS." NEITHER PARTY MAKES ANY WARRANTIES, EXPRESS, IMPLIED OR OTHERWISE, REGARDING ITS ACCURACY, COMPLETENESS OR PERFORMANCE.

Miscellaneous. Neither party may assign its rights or obligations under this Agreement without the prior written consent of the other party. Any attempted assignment in violation of this Section 11 will be null and void. This Agreement shall be governed by the laws of the State of California, without reference to conflict of laws principles. This document contains the entire agreement between the parties with respect

to the subject matter hereof, and neither party shall have any obligation, express or implied by law, with respect to trade secret or proprietary information of the other party except as set forth herein. The invalidity or unenforceability of any provision of this Agreement, or any of its terms or provisions, will not affect the validity of this Agreement as a whole, which will at all times remain in full force and effect. A failure to enforce any provision of this Agreement will not constitute a waiver thereof or of any other provision. This Agreement may not be amended, nor any obligation waived, except by a writing signed by both parties hereto. Any notices required to be given under this agreement shall be deemed given upon the earlier of receipt of five (5) days after mailing by certified mail, return receipt requested, or hand delivery by messenger or express service, to the addresses stated on the first page, or to such other address as the either party may specify to the other in writing from time to time.

By: 
Name: PETER CHEN
Title: COO
Date: 06-01-2021

By: 
Name: William Saenz Blanco
Title: Engineer
Date: 06-01-2021