NEON L0-TO-L1 DISCRETE RETURN LiDAR ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD)

PREPARED BY	ORGANIZATION	DATE
Keith Krause	AOP	03/20/2015
Tristan Goulden	AOP	03/02/2015

APPROVALS	ORGANIZATION	APPROVAL DATE
Kate Thibault	SCI	03/25/2022

RELEASED BY	ORGANIZATION	RELEASE DATE
Tanisha Waters	CM	03/25/2022

See configuration management system for approval history.

Change Record

REVISION	DATE	ECO #	DESCRIPTION OF CHANGE
A	08/27/2015	ECO-02764	Initial release
B	03/25/2022	ECO-06791	• Revised logo and fine print • Minor formatting updates

TABLE OF CONTENTS

1   DESCRIPTION

1.1   Purpose

This document details the algorithms used in generating the Discrete Return Light Detection and Ranging (LiDAR) point cloud NEON Level 1 data product NEON.DOM.SITE.DP1.30003 from Level 0 data, and ancillary data (such as calibration data), obtained via instrumental measurements made by the Airborne Observation Platform (AOP) full waveform LiDAR instrument. It includes a detailed discussion of measurement theory and implementation, appropriate theoretical background, data product provenance, quality assurance and control methods used, approximations and/or assumptions made, and a detailed exposition of uncertainty resulting in a cumulative reported uncertainty for this product.

1.2   Scope

One of NEON's goals is to provide detailed aerial data about regional landscapes and vegetation. This will be accomplished via the remote sensing arm of NEON, called the Airborne Observation Platform (AOP). NEON AOP will fly a suite of integrated remote sensing instruments consisting of an imaging spectrometer, a waveform light detection and ranging (wLiDAR) instrument and a high-resolution digital camera. See the AOP Overview Document (RD[03]) and AOP Payload Design Document (RD[04]) for more information about AOP, the instruments and the complete payload as mounted on the aircraft. As part of the AOP, the airborne remote sensing instruments will collect measurements that characterize ecosystem properties within a 100-300 km^2 region surrounding NEON sites, once per year during a period of peak greenness. The remote sensing instruments are mounted in a common rigid frame and installed in a DeHavilland DHC-6 Twin Otter aircraft flying at 1,000-2,000 m above ground level (AGL) at an average speed of 100 knots. The AOP instrument suite will collect raw data that will be used to derive science data products to provide regional characterization of: land use, vegetation structure, biochemical and biophysical properties of vegetation, and ecosystem responses to changes in land use, climate and the movement of invasive species. Figure 1 shows a high-level overview flowchart of the AOP Level 1 data processing.

This document describes the theoretical background and entire algorithmic process for creating the NEON Level 1 Discrete Return LiDAR Point Cloud data product NEON.DOM.SITE.DP1.30003, corresponding to the green "Discrete LiDAR Processing" box in **Figure 1**. The algorithm reads in a processed aircraft trajectory (RD[05]) and raw discrete return LiDAR data and produces discrete return LiDAR data products in the form of 3D point clouds. As seen in the figure, the discrete return LiDAR data products are used as inputs to the waveform LiDAR (RD[06]), digital camera (RD[07]) and NEON imaging spectrometer (NIS) data processing algorithms (RD[08], RD[09], and RD[10]). This document does not provide computational implementation details, except for cases where these stem directly from algorithmic choices explained herein.
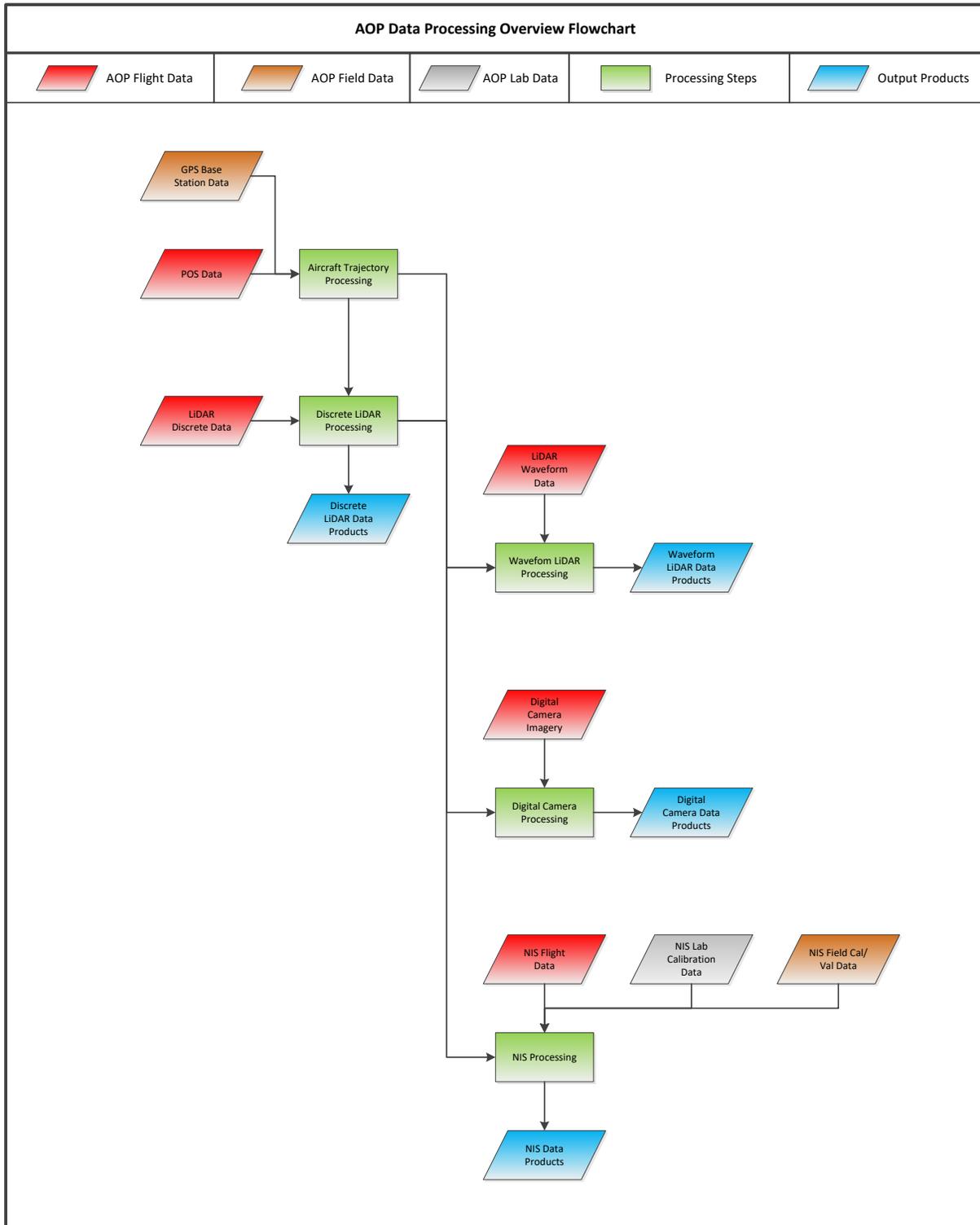
Figure 1. Overview Flowchart of AOP Level 1 Data Processing for data collected by the wLiDAR, the NEON Imaging Spectrometer (NIS) and the Digital Camera.

2   RELATED DOCUMENTS AND ACRONYMS

2.1   Applicable Documents

Applicable documents contain information that shall be applied in the current document. Examples are higher level requirements documents, standards, rules and regulations.

AD[01]	

2.2   Reference Documents

Reference documents contain information complementing, explaining, detailing, or otherwise supporting the information included in the current document.

RD[01]	NEON.DOC.000008	NEON Acronym List
RD[02]	NEON.DOC.000243	NEON Glossary of Terms
RD[03]	NEON.DOC.002236	AOP Overview Document
RD[04]	NEON.DOC.001547	AOP Payload 1 Definition Document
RD[05]	NEON.DOC.00XXXX	PosPac Processing Procedure
RD[06]	NEON.DOC.001293	Waveform Lidar ATBD
RD[07]	NEON.DOC.001211	Digital Camera L0 to L1 ATBD
RD[08]	NEON.DOC.001210	Spectrometer DN to Radiance ATBD
RD[09]	NEON.DOC.001290	Spectrometer L1B to L1G (Geolocation of spectrometer data) ATBD
RD[10]	NEON.DOC.001288	Spectrometer Atmospheric Correction ATBD
RD[11]	NEON.DOC.001557	Flight Planning Standards and Configurations
RD[12]	NEON.DOC.00XXXX	Discrete LiDAR Processing Procedure

2.3   External References

External references contain information pertinent to this document, but are not NEON configuration-controlled. Examples include manuals, brochures, technical notes, and external websites.

ER[01]	ASPRS LAS Specification Version 1.3 – R11 October 24, 2010 http://www.asprs.org/a/society/committees/standards/LAS_1_3_r11.pdf Accessed 10/20/2014
ER[02]	The National Geodetic Service GEOID12A Page http://www.ngs.noaa.gov/GEOID/GEOID12A/ Accessed 10/26/2014
ER[03]	Weitkamp, Claus, "Ch1. Lidar: Introduction," p 5, "Laser Remote Sensing," edited by Takashi Fujii and Tetsou Fukuchi, CRC Press, Taylor & Francis Group, Boca Raton, Florida, 2005
ER[04]	"Topographic Laser Ranging and Scanning: Principles and Processing," edited by Jie Shan and Charles K. Toth, CRC Press, Taylor & Francis Group, Boca Raton, FL, 2009
ER[05]	Chu, Xinzhao and George C. Papen, "Ch5. Resonance Fluorescence Lidar for Measurements of the Middle and Upper Atmosphere," pp. 192-195, "Laser Remote Sensing," edited by

	Takashi Fujii and Tetsou Fukuchi, CRC Press, Taylor & Francis Group, Boca Raton, Florida, 2005
ER[06]	Mahafza, Bassem, R., "Radar Systems Analysis and Design Using MATLAB: Second Edition", Chapman & Hall/CRC Press, Taylor & Francis Group, Boca Raton, Florida, 2005
ER[07]	Optech Gemini Brochure
ER[08]	Optech ALTM Gemini Summary Specification Sheet 120605
ER[09]	Joe Hutton and Dr. Mohamed MR Mostafa, "Georeferencing: State of the Art and New Trends," Workshop #1 at the ASPRS 2013 Annual Conference, Baltimore, Maryland, March 24, 2013.
ER[10]	Liadsky, J. 2013. Personal communication and unpublished Geo-Referencing White Paper
ER[11]	Snyder, J. P. (1987), *Map projections–A working manual*, 1395, USGPO
ER[12]	Applanix POSAV Specifications http://www.applanix.com/media/downloads/products/specs/posav_specs_1212.pdf Accessed 10/20/2014

2.4 Acronym List

AGL	Above Ground Level
AOP	Airborne Observation Platform
ASPRS	American Society for Photogrammetry and Remote Sensing
CHM	Canopy Height Model
DEM	Digital Elevation Model
DGPS	Differential Global Positioning System
DSM	Digital Surface Model
DTM	Digital Terrain Model
ECEF	Earth-Centered Earth-Fixed frame
ENU	East, North, Up
FOV	Field Of View
FWHM	Full Width at Half Maximum
GIS	Geographic Information Systems
GPS	Global Positioning System
IFOV	Instantaneous Field Of View
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
IRTF2000	International Terrestrial Reference Frame 2000
LAS	LASer (LAS) file format exchange
LCP	Laser Calibration Parameters file
LiDAR	Light Detection and Ranging
LMS	Optech LiDAR Mapping Suite
NAVD88	North American Vertical Datum 1988

NIS	NEON Imaging Spectrometer
NEON	National Ecological Observatory Network
PPM^2	Points Per Meter Squared
PRF	Pulse Repetition Frequency
SBET	Smoothed Best Estimate of Trajectory
TBL	Intensity tables file
TOF	Time-Of-Flight
UTM	Universal Transverse Mercator
WGS84	World Geodetic System 1984
wLiDAR	Waveform Light Detection and Ranging

3   DATA PRODUCT DESCRIPTION

The NEON Level 1 Discrete Return LiDAR Point Cloud data product, NEON.DOM.SITE.DP1.30003, is a discrete return LiDAR point cloud in the American Society for Photogrammetry and Remote Sensing (ASPRS) LASer (LAS) file format exchange. The full waveforms are analyzed in the instrument hardware using a proprietary signal detection algorithm and any detected objects are saved as discrete returns (more discussion of target detection is found in section 4.2.1). A point cloud is a collection of point data records containing the X, Y and Z position plus relative signal intensity for each laser return point. The contents of the LAS format are described in detail in Section 3.1.1. The X and Y coordinates are reported in the output horizontal datum and projection and the Z values are reported in absolute elevation in the output vertical datum. The discrete return LiDAR point cloud is a geolocated product given the map coordinates of each laser point are known. An example point cloud returned from a group of trees is shown in **Figure 2**:
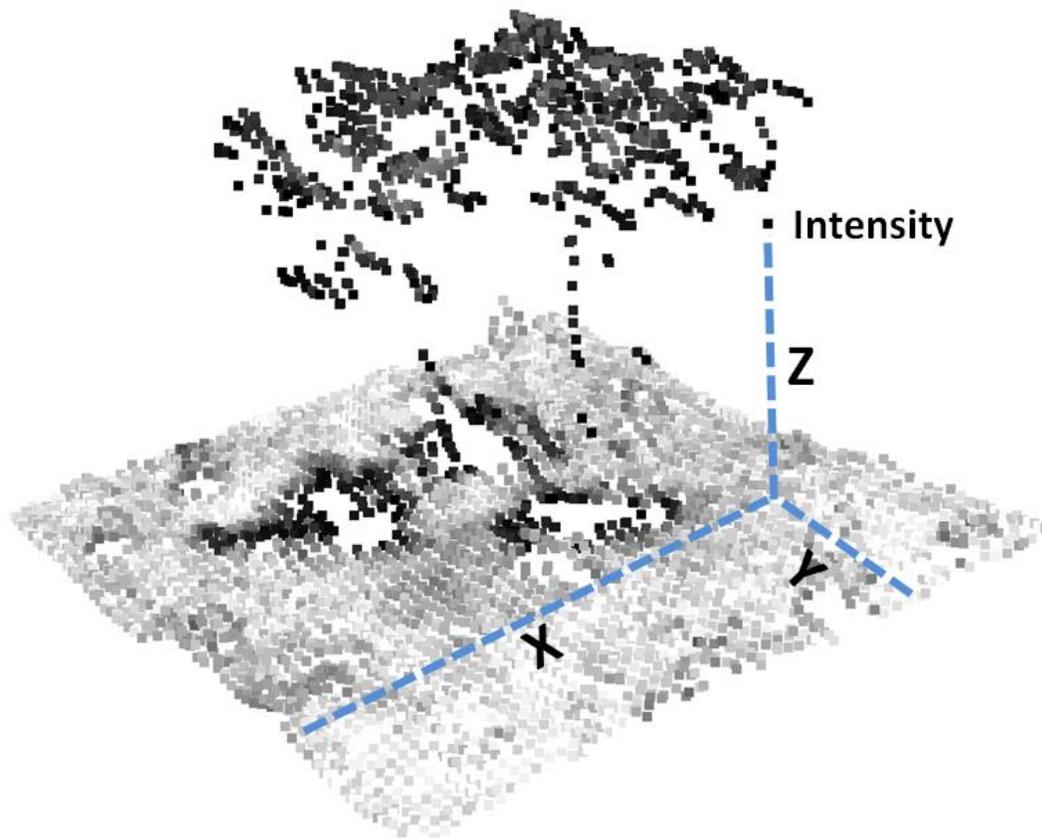
Figure 2. Example point cloud of a group of trees illustrating that each point has X, Y, Z and intensity values.

In discrete return LiDAR, laser pulses of light that reflect off of single return objects such as the ground have a much higher intensity than pulses that must penetrate the canopy (which loose photon energy with each interaction with objects). The discrete return LiDAR point cloud is different than the full waveform data as it is considerably smaller in file size and condenses the information into a small number of points per laser shot rather than many data bins (roughly 100 bins per laser shot in the full waveform). Any information about the shape of the waveform return, which might provide important information about the scattering properties of the object, is lost. Each AOP flight line is saved as an individual LAS file. **Figure 3** shows an example of a discrete LiDAR point cloud for a flight line collected by AOP at Harvard Forest in August 2012. The left side of the figure shows the entire flight line and the right shows a close up view of the forest as viewed in 3D from an oblique side angle. The point clouds are colored by absolute elevation, where the red points represent the highest elevations within the scene and blue represents the lowest elevations.
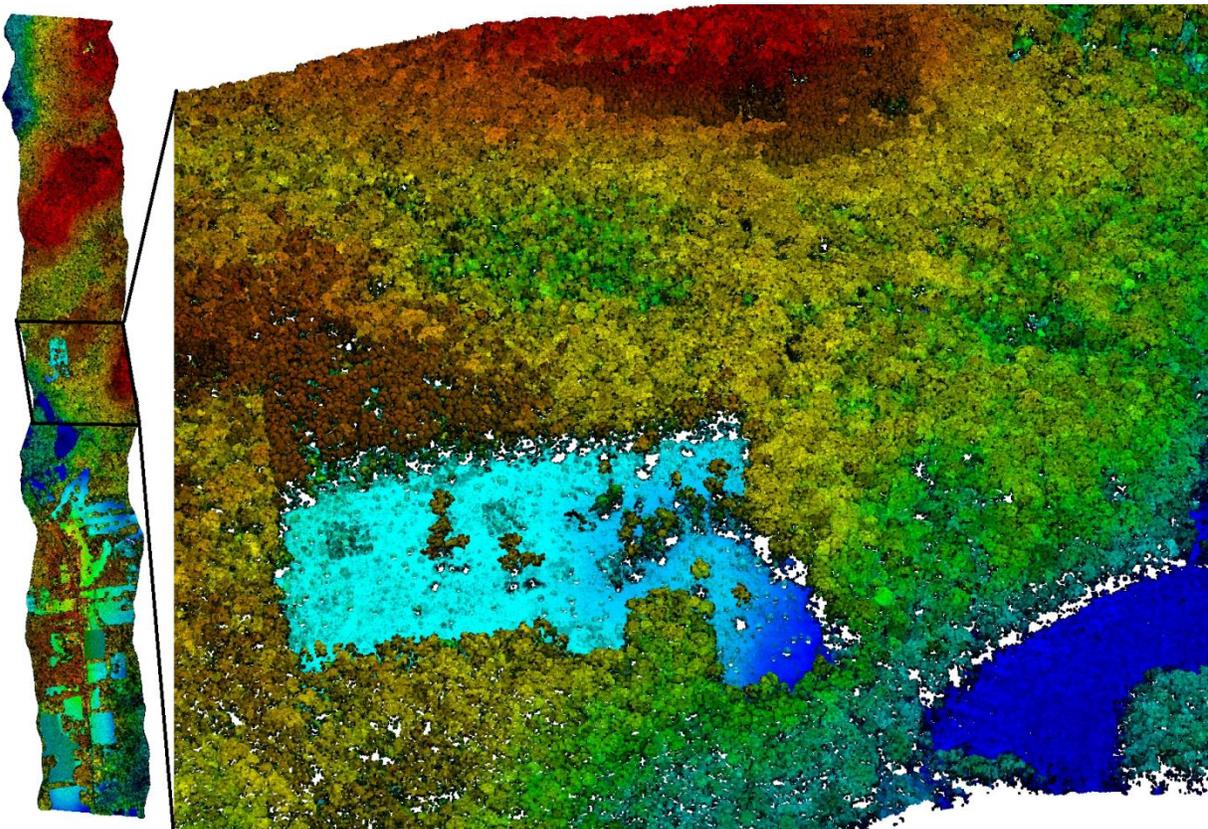
Figure 3. Example Discrete Return LiDAR Level 1 Point Cloud Product colored by absolute elevation. The left side of the figure shows the entire flight line and the right side shows a close up view of the forest as viewed in 3D from an oblique side angle.

A nominal 10 km long flight line flown at a speed of around 100 knots will take about 200 seconds to collect. At a pulse repetition frequency (PRF) value of 100 kHz, the resulting product will contain approximately 20 million laser shots. The wLiDAR instrument can detect multiple returns from a single laser shot so the product will contain greater than > 20 million point records. The size of a nominal LAS file will be approximately 534 MB.

Discrete return LiDAR have many uses: 3D visualization; generation of surface models such as bare-Earth digital elevation models (DEM) also referred to as digital terrain models (DTM), digital surface models (DSM), and canopy height models (CHM); generation of vertical distributions of points (pseudo waveforms); analysis of vegetation structure; and watershed analysis.

3.1 Variables Reported

The NEON Level 1 data product NEON.DOM.SITE.DP1.30003 is a discrete return LiDAR point cloud in ASPRS LAS format consisting of X, Y, Z coordinates and a relative signal intensity for each detected return. The X and Y output coordinates will be reported as Easting and Northing values in a Universal Transverse Mercator (UTM) projection with the World Geodetic System 1984 (WGS84) International Terrestrial Reference Frame 2000 (ITRF 2000) ellipsoid horizontal datum with units of meters. The Z coordinates will be reported in a North American Vertical Datum 1988 (NAVD88) using the National Geodetic Survey Geoid12A height model with units of meters. The UTM zone will vary depending on the latitude/longitude of the specific NEON site. Other information about each return is also provided such as the return number of the object if there is more than one return in the laser pulse, the total number of returns in the laser pulse, the point classification (if a classification algorithm has been run to determine if the point is ground, vegetation, a man-made object, etc.), the scan angle rank, and the Global Positioning System (GPS) time of the laser shot. Each flight line will be saved as an individual output file. While the primary output for this product is a LAS format file, ASCII output files are also required as intermediate products and are used as inputs to the Waveform LiDAR processing algorithm.

3.1.1 LAS 1.3 Format

The output files will be in ASPRS LAS 1.3 format following the ASPRS LAS Specification Version 1.3 – R11 (ER[01]). The LAS format is a binary file and many remote sensing and Geographic Information Systems (GIS) software packages have LAS file readers. The LAS 1.3 format has three sections: public header block, variable length records, and point data records. Versions 1.3 and higher have the potential to include full waveform data, however, this option will not be utilized as this is a discrete return LiDAR product only.

3.1.1.1 Public Header Block

The public header block contains high level file information such as the LAS version number, generating software, creation day and year, header size, offset to point data and number of variable length records. In order to compress LAS file sizes, the X, Y, and Z coordinates are scaled from double precision floating

point numbers into a long word integer using the scale factor and offset values in the public header block. Specific to the point data records, the LAS Specification ER[01] lists:

- "Point Data Format ID: The point data format ID corresponds to the point data record format type. LAS 1.3 defines types 0 through 5".
- "Point Data Record Length: The size, in bytes, of the Point Data Record".
- "Number of point records: This field contains the total number of point records within the file".
- "Number of points by return: This field contains an array of the total point records per return. The first unsigned long value will be the total number of records from the first return, and the second contains the total number for return two, and so forth up to five returns".
- "X, Y, and Z scale factors: The scale factor fields contain a double floating point value that is used to scale the corresponding X, Y, and Z long values within the point records. The corresponding X, Y, and Z scale factor must be multiplied by the X, Y, or Z point record value to get the actual X, Y, or Z coordinate. For example, if the X, Y, and Z coordinates are intended to have two decimal point values, then each scale factor will contain the number 0.01".
- "X, Y, and Z offset: The offset fields should be used to set the overall offset for the point records. In general these numbers will be zero, but for certain cases the resolution of the point data may not be large enough for a given projection system. However, it should always be assumed that these numbers are used".
- "Max and Min X, Y, Z: The max and min data fields are the actual unscaled extents of the LAS point file data, specified in the coordinate system of the LAS data".

See the LAS 1.3 Specification for full public header block details and data formats.

3.1.1.2   Variable Length Records

One or more variable length records can be stored in the LAS file. Typically these records store information such as the data geographic projection. See the LAS 1.3 Specification for full variable length record details and data formats.

3.1.1.3   Point Data Records

The point data records contain specific data for each discrete return detected by the instrument. The product will use LAS 1.3 Point Data Record Format 1 which contains:

- "X, Y, and Z: The X, Y, and Z values are stored as long integers. The X, Y, and Z values are used in conjunction with the scale values and the offset values to determine the coordinate for each point".
- "Intensity: The intensity value is the integer representation of the pulse return magnitude. This value is optional and system specific. However, it should always be included if available".
- "Return Number: The Return Number is the pulse return number for a given output pulse. A given output laser pulse can have many returns, and they must be marked in sequence of

return. The first return will have a Return Number of one, the second a Return Number of two, and so on up to five returns".

- "Number of Returns (for this emitted pulse): The Number of Returns is the total number of returns for a given pulse. For example, a laser data point may be return two (Return Number) within a total number of five returns".

- "Scan Direction Flag: The Scan Direction Flag denotes the direction at which the scanner mirror was traveling at the time of the output pulse. A bit value of 1 is a positive scan direction, and a bit value of 0 is a negative scan direction (where positive scan direction is a scan moving from the left side of the in-track direction to the right side and negative the opposite)".

- "Edge of Flight Line: The Edge of Flight Line data bit has a value of 1 only when the point is at the end of a scan. It is the last point on a given scan line before it changes direction".

- "Classification: This filed represents the "class" attributes of a point. If a point has never been classified, this byte must be set to zero. There are no user defined classes since all point formats 0 supply 8 bits per point for user defined operations". See the LAS 1.3 Specification for more details on the possible values for classification.

- "Scan Angle Rank: The Scan Angle Rank is a signed one-byte number with a valid range from -90 to +90. The Scan Angle Rank is the angle (rounded to the nearest integer in the absolute value sense) at which the laser point was output from the laser system including the roll of the aircraft. The scan angle is within 1 degree of accuracy from +90 to –90 degrees. The scan angle is an angle based on 0 degrees being nadir, and –90 degrees to the left side of the aircraft in the direction of flight".

- "User Data: This field may be used at the user's discretion".

- "Point Source ID: This value indicates the file from which this point originated. Valid values for this field are 1 to 65,535 inclusive with zero being used for a special case discussed below. The numerical value corresponds to the File Source ID from which this point originated. Zero is reserved as a convenience to system implementers. A Point Source ID of zero implies that this point originated in this file. This implies that processing software should set the Point Source ID equal to the File Source ID of the file containing this point at some time during processing".

- "GPS Time: The GPS Time is the double floating point time tag value at which the point was acquired. It is GPS Week Time if the Global Encoding low bit is clear and Adjusted Standard GPS Time if the Global Encoding low bit is set (*see Global Encoding in the Public Header Block description)".*

To convert a given point record from scaled X, Y, and Z values to absolute coordinates,

$X_{coordinate} = (X_{record} * X_{scale}) + X_{offset}$

$Y_{coordinate} = (Y_{record} * Y_{scale}) + Y_{offset}$

$Z_{coordinate} = (Z_{record} * Z_{scale}) + Z_{offset}$

See the LAS 1.3 Specification (ER [01]) for full point data record details and data formats. A future version of the product might include RGB values derived using the NEON high resolution RGB imagery. This will require use of Point Data Record Format 3 which includes values for red, green and blue.

3.1.2   ASCII Format

ASCII output files, which contain the scan mirror angle position versus time are also required as intermediate products. The scan mirror angle position information within these files are used as inputs to the Waveform LiDAR processing algorithm (RD[06]) . The full required contents of the ASCII files will be determined upon completion of the Waveform LiDAR processing algorithm. Possible information required by the waveform algorithm includes:

- Laser shot GPS time
- Laser shot scan angle
- Laser shot number of ranges
- Laser range
- Laser range intensity
- Laser point X-coordinate
- Laser point Y-coordinate
- Laser point Z-coordinate

In future versions of this algorithm, ASCII files may be required to include standard deviation information about each of the calculations and may be used for quality assurance and control.

3.2   Input Dependencies

The algorithm requires several user defined inputs, calibration files, and raw and intermediate data files as listed in **Table 1** and described in the following sections.

Table 1. Input Dependencies.

Input	Data Category	Processing Level
Laser Calibration Parameters (LCP) File	Calibration Data	Calibration
RES File	Calibration Data	Calibration
Intensity Tables (TBL) file	Calibration Data	Calibration
Range File (.range)	Flight Mission Data	Raw

Smooth Best Estimate of Trajectory (SBET.out) File	Flight Mission Data	Intermediate
Trajectory residual precision values (Smrmsg.out) File	Flight Mission Data	Intermediate
Temperature (T)	Meteorological Data	Ancillary
Pressure (P)	Meteorological Data	Ancillary
Ellipsoid to Geoid Model	Geographic Data	External

3.2.1 Instrument Calibration Files

The Optech Gemini waveform LiDAR instrument requires three calibration files: the laser calibration parameters (LCP) file, the RES file, and intensity tables (TBL file). The RES file, intensity tables, and an initial LCP file are generated after calibration of the instrument by Optech. The LCP file is updated by running a self-calibration within the Optech LMS software using data collected during a boresight calibration flight (RD[11]).

3.2.1.1 LCP File

The LCP (laser calibration parameters) file contains the scan mirror scale and offset factors plus the boresight alignment angles Ex, Ey and Ez. An example LCP file is shown in **Figure 26** in Appendix A. The LCP file is in XML format.

3.2.1.2 RES File

The RES file contains the range offset factors.

3.2.1.3 Intensity Table

The proprietary signal processing algorithms that detect objects in the received signal may introduce a bias in the range estimation. The intensity tables contain range adjustment values as a function of detected return intensity for each of the PRF settings and for 1st, 2nd, 3rd, and 4th returns. These tables are generated during instrument calibration at the factory.

3.2.2 Flight Mission Data

The flight mission data includes the raw Level 0 discrete return LiDAR data, the trajectory files, and meteorological data for the flight.

3.2.2.1 Raw Discrete LiDAR Data

The raw Level 0 discrete return LiDAR data comes in the form of a binary range file. The range file is proprietary so the exact contents are not known to NEON, however, the processing algorithm, i.e. the Optech LMS software, can read in these files.

3.2.2.2 Trajectory Files

The trajectory files include the smoothed best estimate of trajectory (SBET) file (SBET.out) and a set of residual precision values (smrmsg.out). The SBET is a binary file that provides aircraft longitude, latitude, height, roll, pitch and heading versus GPS time. The smrmsg file is also a binary file that gives residual trajectory precision values that can be used to calculate laser point accuracy. The SBET and smrmsg files are generated by processing the raw aircraft GPS and inertial measurement unit (IMU) with the Applanix POSPAC MMS software (RD[05]). Figure 8 shows an example of the trajectory in latitude and longitude of a nominal engineering flight.

3.2.3 Meteorological Data

Input values for the meteorological data must be set, specifically the temperature (deg C) and pressure (mBar or hPa). Temperature and pressure values are recorded at the beginning of the flight at the airport, however it is recommended to use the average of the ground and aircraft altitude values if known. Also, the algorithm allows the input of a meteorological file that would provide temperature and pressure versus time. These values are used to calculate the index of refraction of air in order to calculate the correct laser range based on the time of flight between the laser pulse and the received signal. An example of temperature and pressure from a nominal engineering flight is listed in **Table 2**:

Table 2. Example of temperature and pressure from a nominal engineering flight.

Parameter	Value
Temperature	29.0° C
Pressure	1015.92 mbar

3.2.4 Geographic Data

The georeferencing equations are designed to perform the calculations in ellipsoid space, which is mathematically simpler. The lat/lon geodetic coordinates are converted to the output projection, in this case as Easting and Northing values in a Universal Transverse Mercator (UTM) projection with the World Geodetic System 1984 (WGS84) International Terrestrial Reference Frame 2000 (ITRF 2000) ellipsoid horizontal datum with units of meters. The Z coordinates will be reported in a North American Vertical Datum 1988 (NAVD88) using the National Geodetic Survey Geoid12A height model with units of meters. The NOAA National Geodetic Survey maintains the Geoid12A model (ER[02]). This model is a series of

geometric harmonics that describes where the mean sea height would be at a given location on Earth based on Earth's gravity and rotation. In a practical application of the model, look up tables are generated to convert from ellipsoid to geoid for a given spatial coordinate on the ground.

3.3   Product Instances

Currently there is only one instance of this product, specifically a discrete return LiDAR point cloud in ASPRS LAS 1.3 format (ER[01]). See Section 9 for future modifications and plans which will describe potential future sub-products.

3.4   Temporal Resolution and Extent

The high cost of aircraft operations will limit the frequency of standard airborne observations of each site to once per year to detect inter-annual trends. Flights will be planned to occur at each site during a period of peak greenness to minimize the risk of acquiring data too early in the season (during leaf green up in the spring) or too late in the season (when vegetation becomes senescent) in which case the instruments would be measuring physical differences in vegetation due to single year phenology and not long term trends. The discrete return LiDAR algorithm is applied on each AOP flight line, which typically measure between 5 and 20 km in length. Flight speeds are typically around 100 knots (185.2 km/hour), and therefore, the time required to acquire flight lines of the lengths stated will range from 1.6 to 6.5 minutes. A nominal flight lasts about 4-5 hours including the survey of the site, turns between flight lines, and transit to and from the airport. The full waveform LiDAR instrument uses a pulsed laser, which has a pulse repetition frequency (PRF) of 33-167 kHz where the nominal PRF value is around 100 kHz. The time between laser pulses is between 6-30 μs with a nominal time of 10 μs. A laser pulse takes 6.67 μs to travel to a target at the nominal range of 1000 m and then return back to the instrument receiver. The LiDAR instrument employs a scanning mirror to sweep out the laser pulses in the across-track direction on the ground which can operate up to 70 Hz but is typically set to 33.3-50 Hz to optimize the point spacing with the aircraft speed. The outgoing pulses are near Gaussian in shape and have a full width at half maximum (FWHM) value of about 10-15 ns. The return waveform is digitized to 1 ns bins.

3.5   Spatial Resolution and Extent

AOP will collect information that characterizes ecosystem properties with a 100-300 km^2 region surrounding NEON sites. Each flight line is typically between 5 and 20 km in length. The aircraft will fly between 500-2000 m above ground level (AGL) with a nominal flight altitude of 1000 m AGL. The LiDAR has a narrow beam divergence of 0.25 mrad and a wide divergence of 0.8 mrad, which will produce a beam spot size on the ground (instantaneous field of view IFOV) of 0.25 m and 0.8 m respectively at 1000 m AGL. The field of view of the LiDAR is set to 37.0-40.0 degrees to image a slightly wider swath on the ground than the imaging spectrometer, giving a 670-728 m wide swath at 1000 m AGL. With the nominal PRF of 100 kHz, a scan mirror frequency of 50 Hz and a flight speed of about 100 knots, the spatial sampling is around 0.524m in the across-track and 0.5 m in the along-track direction. The actual ground resolution will vary with flight altitude, topography, and cross-track field angle. While it is

NEON's goal to collect remote sensing data using consistent methods, even when flying straight and level the aircraft experiences changes in roll, pitch, and yaw due to turbulence. These changes can yield inconsistencies in ground sampling. Finally, an outgoing pulse width of 10 ns gives a vertical range resolution of about 1.5 meters. This means that two objects that are closer than 1.5 m in height cannot be resolved as individual objects. The vertical range accuracy of single objects is much better than 1.5 m given the waveform is sampled at 1 ns bins (0.15 m resolution) and many discrete detection algorithms such as leading edge detection rely on fitting shapes to a higher sampling interval.

4   SCIENTIFIC CONTEXT

4.1   Theory of Measurement/Observation

4.1.1   Basic LiDAR Theory for Airborne Laser Mapping

Laser distance measurements using the light detection and ranging (LiDAR) technique have been in practice since the 1960s and were developed shortly after the invention of the laser. However, the concept was demonstrated much earlier in the late 1940's to detect cloud heights using searchlight mirrors ER[03]. At a high level, the concept is fairly simple: the laser emits a pulse of light (referred to as the outgoing T0 pulse); the pulse of light travels through the atmosphere; the light interacts with objects causing some of the photons to scatter; the scattered photons travel through the atmosphere a second time; and the photons that are scattered back to the instrument are measured by a receiver. This concept is illustrated for the AOP case in **Figure 4** where the laser is mounted on board an aircraft and the scattering/reflecting objects are targets on the Earth's surface such as vegetation, bare ground and man-made objects.
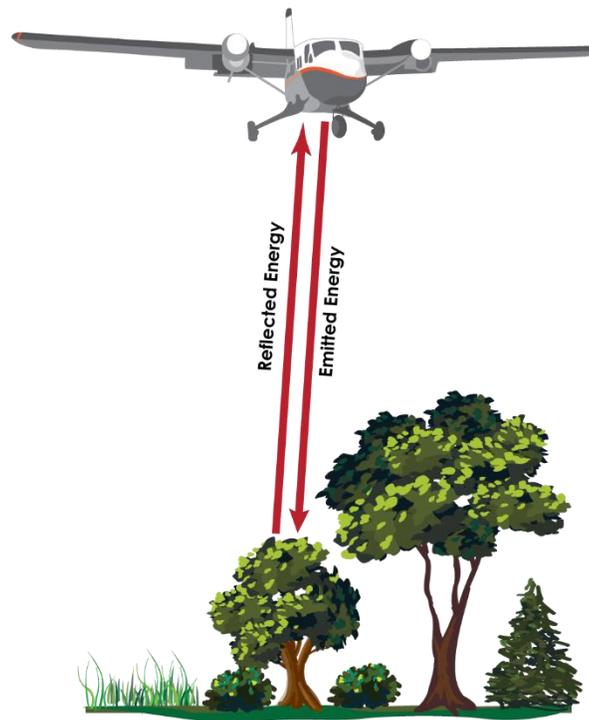
Figure 4. Basic Airborne LiDAR Concept.

The range (distance) to the target can be calculated by measuring the amount of time after the outgoing T0 pulse is emitted pulse to when the receiver measures a return signal and then applying the speed of light in air to convert the elapsed time to distance traveled by the pulse of light. If the position and

orientation of the LiDAR instrument are known, then the absolute position of the scattering object can also be calculated. For more information on basic LiDAR concepts and ranging, refer to Chapter 1 in ER[04]. **Figure 5** shows an example of an outgoing T0 pulse and the resulting received return waveform from vegetation objects on the ground. As can been seen in the left side of the figure, the outgoing pulse is near Gaussian in shape. The time interval for the waveform segment along with target detection time adjustments are used in the range calculation (see section 4.2.1 for further discussion on calculation of time-of-flight and range).
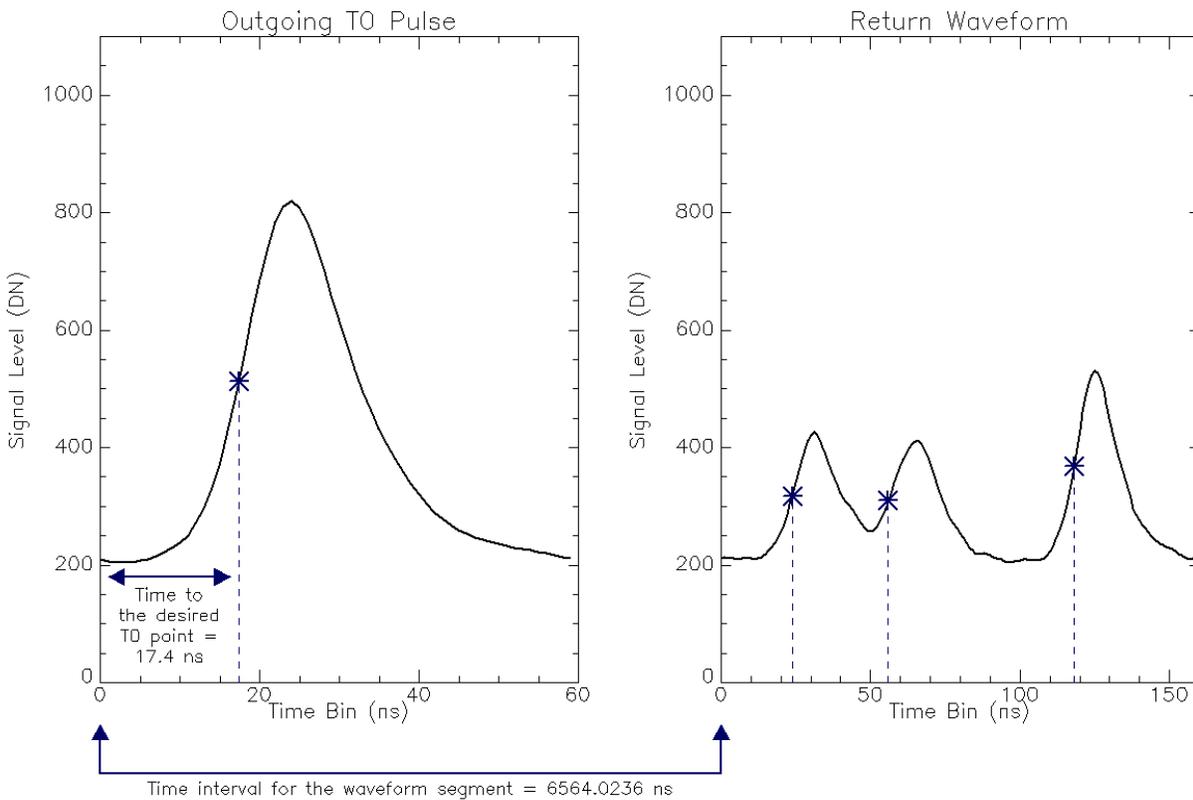
Figure 5. An example of an outgoing T0 pulse and the resulting received return waveform from vegetation objects on the ground

4.1.1.1   LiDAR Equation

A generalized LiDAR equation describes how many photons are expected to be measured by the instrument. For the purposes of nominal terrestrial airborne laser scanning, fluorescence will be ignored so the measured/received wavelength is the same wavelength as the instrument outgoing laser wavelength. Modified from Chu and Papen, 2005 (ER[05]), the LiDAR equation can be written as:

$$N_S(\lambda, R) = N_L(\lambda) \cdot [\beta(\lambda, R)\Delta R] \cdot \frac{A}{R^2} \cdot [T^2(\lambda, R)] \cdot [\eta(\lambda)G(R)] + N_B$$

Where:

- λ is the instrument laser wavelength
- R is the range from the instrument
- $N_S(\lambda, R)$ is the expected number of photon counts detected at wavelength λ from a target a distance R from the instrument
- $N_L(\lambda)$ is the number of photons emitted by the laser (transmitted photons)
- $\beta(\lambda, R)$ is the volume backscatter coefficient at range R for wavelength λ. This is a combination of the scattering properties of any medium at a distance R including its physical size and reflectance (if it is a solid object)
- ΔR is the size of the range bin or range interval and is determined by the sampling of the receiver signal
- A is the receiving telescope aperture area
- $T^2(\lambda, R)$ is the light transmission through the medium for the transmitted laser and return signal photons or the two-way transmission (both the outgoing and incoming transmission at range R and wavelength λ)
- $\eta(\lambda)G(R)$ is the overall system efficiency including the optical efficiency and a geometric factor for the transmitter/receiver combination
- And N_B is the background noise and detector dark count.

Here:

- $\beta(\lambda, R)\Delta R$ can also be interpreted as the probability that a transmitted photon is scattered by an object into a unit solid angle
- and A/R^2 can also be interpreted as the probability that a scattered photon is collected by the receiving telescope or the solid angle subtended by the receiver.

Furthermore, $N_L(\lambda)$, the number of transmitted laser photons, can be written as

$$N_L(\lambda) = \left(\frac{P_L(\lambda)\Delta t}{\frac{hc}{\lambda}} \right)$$

Where:

- $P_L(\lambda)$ is the peak laser power (assuming a top hat temporal pulse shape)
- Δt is the outgoing pulse width
- h is Planck's constant
- c is speed of light
- And λ is the wavelength.

4.1.1.2   Laser Ranging

The range between two locations can be measured accurately by timing how long it takes light to travel from one location to the other. In LiDAR systems, the laser pulse must travel from the instrument to an object and then scatter back to the receiver, so it is the two-way travel time (or 'time of flight') that is measured. Range is calculated as

$$R = c \cdot \frac{\tau}{2}$$

Where:

- R is the range in m
- c is the speed of light (m/s)
- and τ is the time interval between the outgoing pulse and the received signal.

Unless the LiDAR is being operated in a vacuum (for instance space vehicle docking systems) the index of refraction of air, *n*, should be taken into account, modifying the equation to

$$R = \frac{c}{n} \cdot \frac{\tau}{2}$$

Where:

- n is the index of refraction of air which varies with temperature, pressure, and humidity.

4.1.1.3 Range Resolution and Range Precision

Mahafza defines range resolution as "the ability to detect targets in close proximity to each other as distinct objects" (ER[06]). The outgoing pulse is not an instantaneous delta function and has a temporal shape as seen in **Figure 5**. Most laser pulses can be characterized as being near-Gaussian with a pulse width value defined by the full-width-at-half-maximum (FWHM). An outgoing pulse with an ideal delta function would theoretically be able to differentiate two semi-transparent thin hard objects that are infinitesimally close to each other. However, those same objects will blend together in the received signal with a Gaussian shape pulse with a finite pulse width. The receiver detector and signal processing chain will also increase the range resolution beyond the actual laser pulse width. A practical value for the range resolution can be calculated as:

$$\Delta R = c \cdot \frac{\Delta \tau}{2}$$

Where

- ΔR is the range resolution in meters
- And $\Delta \tau$ is the pulse FWHM.

For example, an outgoing pulse width of 10 ns gives a range resolution of about 1.5 m. This means that two objects that are closer than 1.5 m cannot be resolved as individual objects.

Range precision describes the precision in the measurement of absolute range to a target. In practice, the range precision of single objects is much better than 1.5 m since many discrete detection algorithms use analog signal processing techniques that can detect targets with a fine temporal sampling (better than the outgoing pulse width).

4.1.1.4 Range Gate

A range gate is employed to block window reflections of T0 outgoing pulse and potential bright atmospheric returns that are close in range to the instrument that might saturate the detectors or take up data recorder space. The instrument does not record any data before the range gate is opened. Typically the range gate is set to block signals within 150 m of the instrument.

4.1.1.5   Range Ambiguity

Range ambiguity occurs when there are multiple pulses in the air. Specifically, ambiguity occurs if the range to a target is too large where the time of flight is greater than 1/PRF (pulse repetition frequency). For a single pulse in air at a time, the maximum target range is

$$R_{max} = \frac{c}{2} \cdot \frac{1}{PRF}$$

Many newer LiDAR systems have proprietary technology to handle range ambiguity and multiple pulses in the air at the same time. If a system cannot simultaneously track multiple pulses in the air, then the return signal from a previous pulse may be recorded one time interval later resulting in a significant error in the recorded time-of-flight.

4.1.1.6   Beam Divergence

The outgoing laser beam is not collimated and has a beam divergence angle. The size of the laser spot on the ground is proportional to the aircraft altitude and can be calculated as

$$D_S = 2h \cdot tan\left(\frac{\gamma}{2}\right)$$

Where:

- D_s is the beam spot diameter
- h is the height or range to the target
- and γ is the beam divergence angle.

4.1.1.7   Swath Width

The swath width on the ground of the laser scan is constrained by the mirror scan angle or field of view and can be calculated as

$$S_W = 2h \cdot tan\left(\frac{\theta}{2}\right)$$

Where:

- S_W is the data cross track swath width and θ is the mirror full scan angle.

4.1.1.8 Ground Sampling and PPM-Squared (PPM^2)

The spacing of laser points on the ground is determined by the aircraft altitude, aircraft velocity, laser PRF, scan mirror frequency, and mirror full scan angle. A common metric is to calculate how many points fall within a 1 m squared area or the points per meter squared (PPM^2). For a fixed aircraft altitude and velocity, the LiDAR point spacing on the ground can be adjusted by setting the PRF, scan mirror frequency, and the scan half angle. For a chosen scan mirror half angle, the scan mirror frequency will increase (lower frequency) or decrease (higher frequency) the spacing between points (in adjacent mirror scans) in the along-track direction. Also for a chosen scan mirror half angle, changes to the PRF will increase (higher PRF) or decrease (lower PRF) the spacing between points in the across-track direction. It is desirable to set the LiDAR parameters so that the laser points on the ground are evenly spaced in the cross-track and down-track directions (at NADIR). Note: The Optech Gemini uses a saw-tooth mirror scanning pattern so the along-track spacing between points on the edges of the swath is slightly larger than at NADIR. **Figure 11** shows the LiDAR point spacing on the ground for a nominal collection.

4.1.1.9 No Return Situation

Situations occur where little or no light is scattered back to the receiver and no return is detected. An example of this is when the laser spot illuminates water at an off-NADIR angle and any of the scattered light is specularly reflected away from the instrument.

4.1.2 AOP Waveform LiDAR Instrument Description

The AOP is currently flying the Optech Incorporated Airborne Laser Terrain Mapper (ALTM) Gemini as the full waveform LiDAR instrument (ER[07]). The Gemini is pictured in Figure 6 and the nominal instrument specifications are given in **Figure 7** (ER[08]).

Figure 6. Optech Gemini System Components.

The Gemini contains nominally a 10 Watt pulsed Nd:YAG laser that operates at a wavelength of 1064 nm and emits 250 µJ energy per pulse max with a pulse width on the order of 10 ns. The laser pulse repetition frequency (PRF) can be configured to one of seven values: 33, 50, 70, 100, 125, 142 or 166 kHz. The outgoing pulse is approximately Gaussian with and the pulse width or FWHM varies with PRF. The Gemini uses a scanning mirror to sweep out the laser pulses in the across-track direction on the ground. The mirror scan width angle can be configured to set the field of view (FOV). The scan frequency of the mirror can also be set in order to speed up or slow down the locations of the laser pulses on the ground in both the across-track and down-track directions. The Gemini also has foreoptics which establish the instantaneous field of view (IFOV) and can be set to a narrow divergence value of 0.25 mrad (1/e) or a wide value of 0.8 mrad (1/e). The pulse sampling on the ground can be specifically configured by tuning the PRF, the mirror scan width, the mirror scan frequency, the beam divergence and the aircraft speed (RD[11]). Finally, the Gemini has full waveform capability, but in the case of the discrete return, the return waveform signal is analyzed and targets are identified. The system can record up to four range measurements (4 returns). In the case where the system detects more than four objects, it records the 1st, 2nd, 3rd and last return ranges. The intensity of each return signal is also recorded.

Parameter	Specification
Operational envelope[1]	150-4000 m AGL, nominal
Laser wavelength	1064 nm
Horizontal accuracy[2]	1/5, 500 x altitude (m AGL)
Elevation accuracy[2]	<5-35 cm; 1 σ
Effective laser repetition rate	Programmable, 33-167 kHz
Position and orientation system	POS AV™ AP50 (OEM); 220-channel dual frequency GPS/GNSS/Galileo/L-Band receiver
Scan width (FOV)	Programmable; 0-50°
Scan frequency	Programmable; 0-70 Hz
Sensor scan product	1000 maximum
Beam divergence	Dual divergence: 0.25 mrad (1/e) and 0.8 mrad (1/e), nominal
Roll compensation	Programmable; ±5° at full FOV
Range capture	Up to 4 range measurements, including 1st, 2nd, 3rd, and last returns
Intensity capture	Up to 4 intensity returns for each pulse, including last (12 bit)
Video camera	Internal video camera (NTSC or PAL)
Image capture	Compatible with full Optech camera line (optional)
Full waveform capture	12-bit IWD-2 Intelligent Waveform Digitizer (optional)
Data storage	Removable solid state disk SSD (SATA II)
Power requirements	28 V; 900 W; 35 A (peak)
Dimensions and weight	Sensor: 260 mm (w) x 190 mm (l) x 570 mm (h); 23 kg Control rack: 650 mm (w) x 590 mm (l) x 530 mm (h); 53 kg
Operating temperature	-10°C to +35°C (with insulating jacket)
Relative humidity	0 – 95% non-condensing

1 10% reflective target
2 Accuracy dependent on selected operational parameters using nominal 50° FOV in standard atmospheric conditions

DANGER

US FDA 21 CFR 1040.10 and 1040.11; IEC/EN 60825-1

Optech Incorporated
300 Interchange Way, Vaughan ON, Canada L4K 5Z8
Tel: +1 905 660 0808   Fax: +1 905 660 0829
www.optech.com
© Optech Incorporated. E&OE. Information subject to change without notice. Printed in Canada. 120805

Optech

Figure 7. Optech Gemini Instrument Specifications.

4.1.3   AOP Nominal LiDAR Measurement Observation

4.1.3.1   Concept of Operations

See the flight planning standards and configurations document (RD[11]).

4.1.3.2   Example of Nominal Flight and Instrument Configuration

An example of a nominal flight is presented for a flat site of ~150 km² size. The nominal collection parameters are:

- Coverage: 10 km x 15 km box
- Aircraft Speed: 50 m/sec (97.2 kts)
- Altitude: 1,000 m AGL
- Flight Line Overlap: 30%
- Discrete return plus full 12-bit waveform data
- PRF: 100 kHz
- Scan Frequency: 50 Hz
- Half Scan Angle: 18.5 deg
    - Desire to have wider LiDAR swath width wider than spectrometer swath width
- Beam Divergence: Wide

Based on these example parameters, a flight plan can be designed that includes 24 North-South production flight lines and two East-West cross strips (one at the beginning and the other at the end of the flight). The flight path is shown in **Figure 8** for a data collection over Harvard Forest in August 2012.
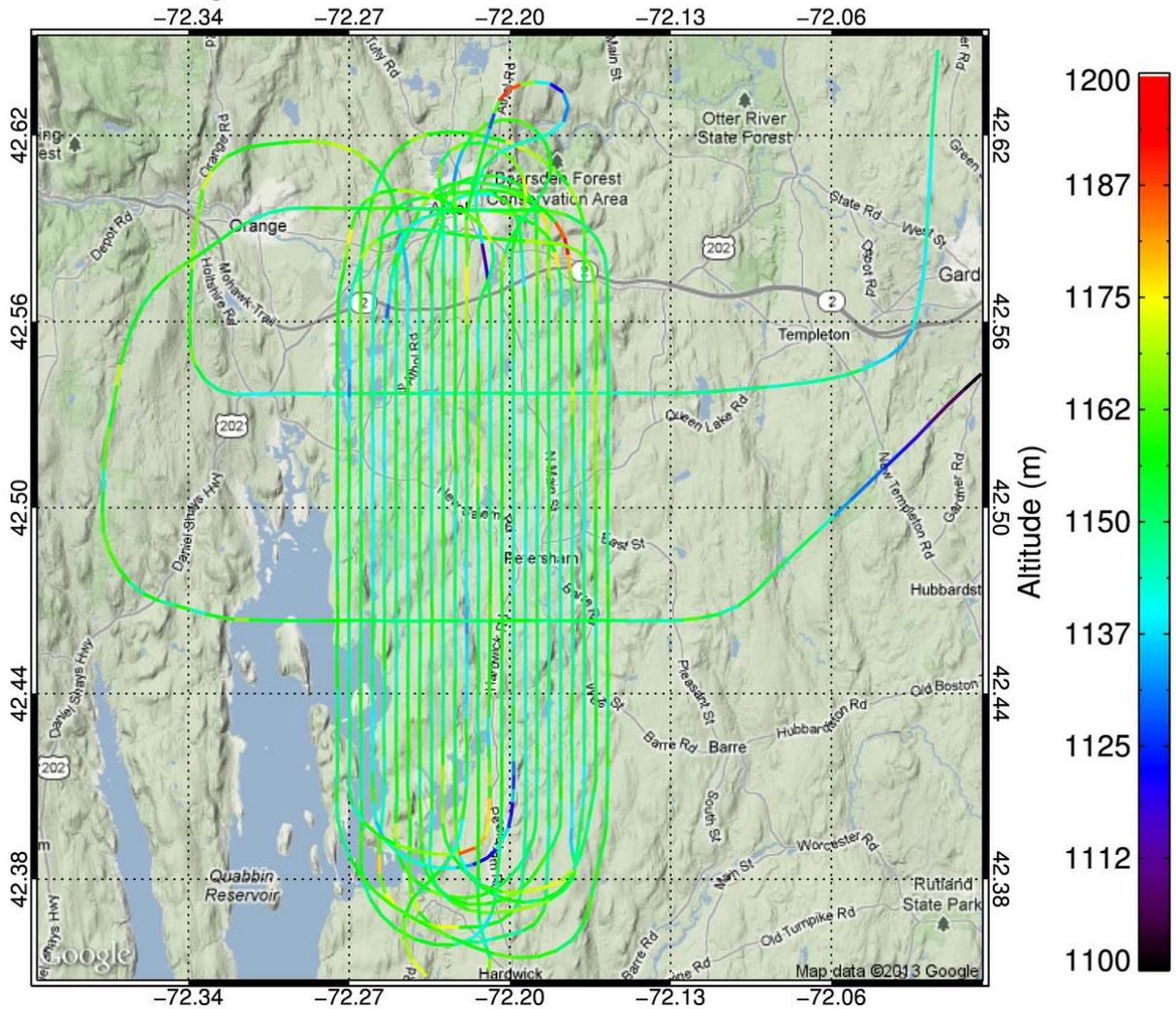
Figure 8. Flight Track of a Nominal AOP Collection.

At NADIR, these flight and instrument parameters achieves 0.8 m spots with 3.82 points per m^2 (PPM^2) and a point spacing of 0.524 m in the across-track direction and 0.5 m in the along-track direction. The LiDAR ground swath width is 670 m. **Figure 9** shows a trace of the scan pattern on the ground for seven full cycles of the scan mirror. The mirror uses a saw-tooth scan pattern which combined with the aircraft speed leaves a slightly wider along-track spacing of ~1 m on the edges of the flight lines.
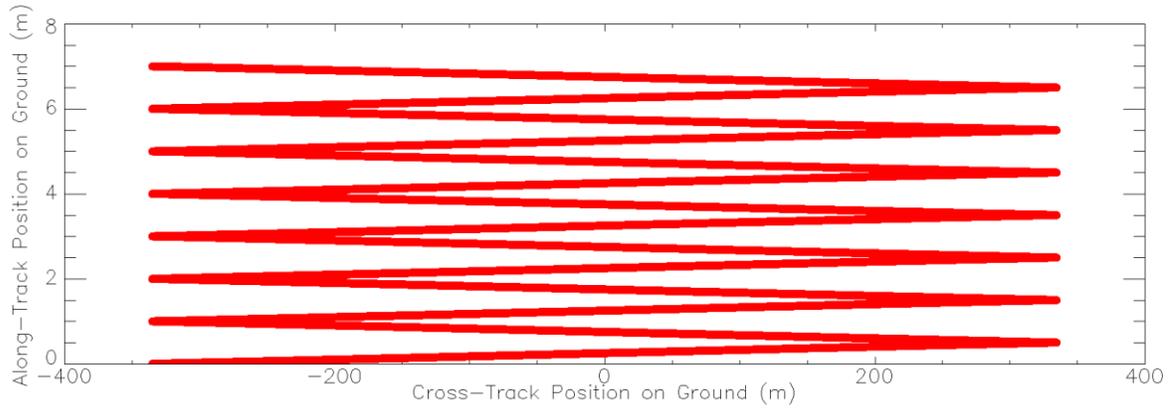
Figure 9. LiDAR Scan Pattern on the Ground for a Nominal Collection.

Figure 10 shows the LiDAR point spacing on the ground for three different view angles including the center, quartile and edge of the field of view. With 0.8 m spots spaced ~0.5 m, wall-to-wall coverage is achieved in the center with some oversampling. Moving off-nadir to the quartile, some sampling gaps start to appear in the corners. Finally larger sampling gaps are seen at the edge of the FOV, especially in the along-track direction.
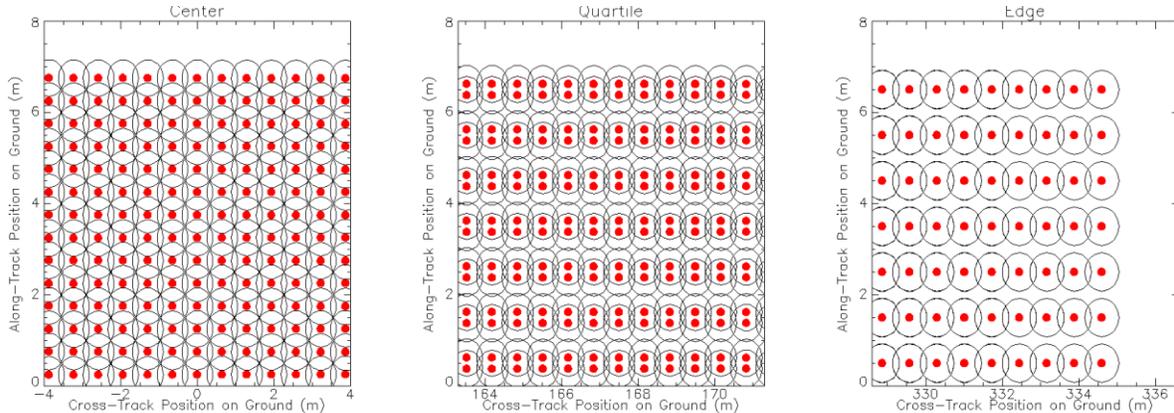
Figure 10. LiDAR Point Spacing on the Ground for a Nominal Collection.

Note that the flight and instrument parameters will change depending on the size and topography of each NEON site. For instance, sites with large terrain relief might require a larger percent flight line overlap to avoid data gaps and might require operating the LiDAR at a slower PRF in order to avoid range ambiguities.

4.2 Theory of the Algorithm

The discrete return LiDAR algorithm has several components: target detection, calculation of time-of-flight (TOF), calculation of range, and geolocation. A high level processing flow chart is shown in **Figure 11**:

Figure 11. High-level discrete LiDAR processing flow.

In the following sections, target detection, calculation of time-of-flight, and calculation of range will be combined in a single discussion.

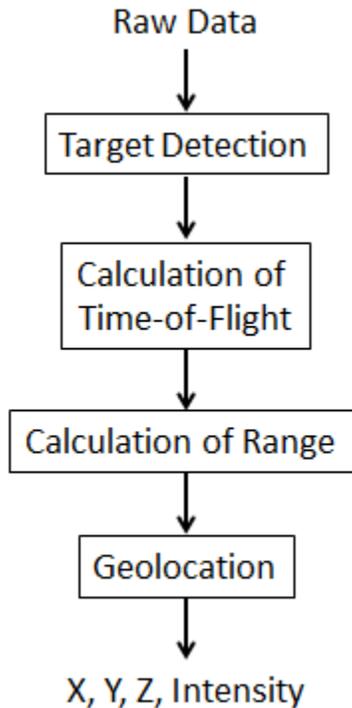### 4.2.1   Target Detection, Calculation of Time-of-Flight, and Calculation of Range

Discrete return LiDAR algorithms attempt to identify targets in the returned signal scattered back to the receiver by calculating the time elapsed between the generation of the T0 pulse and the arrival of the received signal. Several different types of algorithms exist (ER[04]), but one commonly used simple method is leading edge detection using the 50% intensity point. This is illustrated in **Figure 12** which shows the outgoing T0 pulse on the left and the return waveform on the right for what appears to be a single hard target (possibly the ground or other hard surface) collected over Harvard Forest in August 2012. For each of the two curves, the leading edge 50% point is identified. The time of flight is calculated:

$$\tau = t_s + t_{1-}t_0$$

Where:

- τ is the time of flight (ns)
- t_s is the time interval for the waveform segment (ns)
- t_1 is the time to the desired 50% leading edge point on the return waveform
- and t_0 is the time to the desired 50% leading edge point on the T0 outgoing pulse.

For the example in the figure, the time of flight is 6563.724 ns.

Figure 12. Example of Leading Edge Detection for Time of Flight Calculation.

Before range is determined, the index of refraction in air must be calculated. At 1064 nm this can be approximated by

$$n = 1 + 78.7 \times \frac{P}{(273.15 + T)} \times 10^{-6}$$

Where:

- T is the mean atmospheric temperature in deg C
- and P is the mean pressure in millibars (mbar) or hectopascals (hPa).

For this example, T = 29.0° C and P = 1015.92 mbar giving n = 1.0002646.

Finally, range is calculated as

$$R = \frac{c}{n} \cdot \frac{\tau}{2}$$

Arriving at a value of 983.617 m. Note that the actual discrete return LiDAR algorithm as implemented in the Optech Gemini instrument detects discrete return targets from the analog waveform signal of the receiver using a proprietary algorithm. **Figure 13** shows a second example with multiple returns, where three significant objects are identified. In this example the first return (vegetation) has a range of 984.621 m, the second return has a range of 989.431 m and the last return (which might be the ground) has a range of 998.753 m.
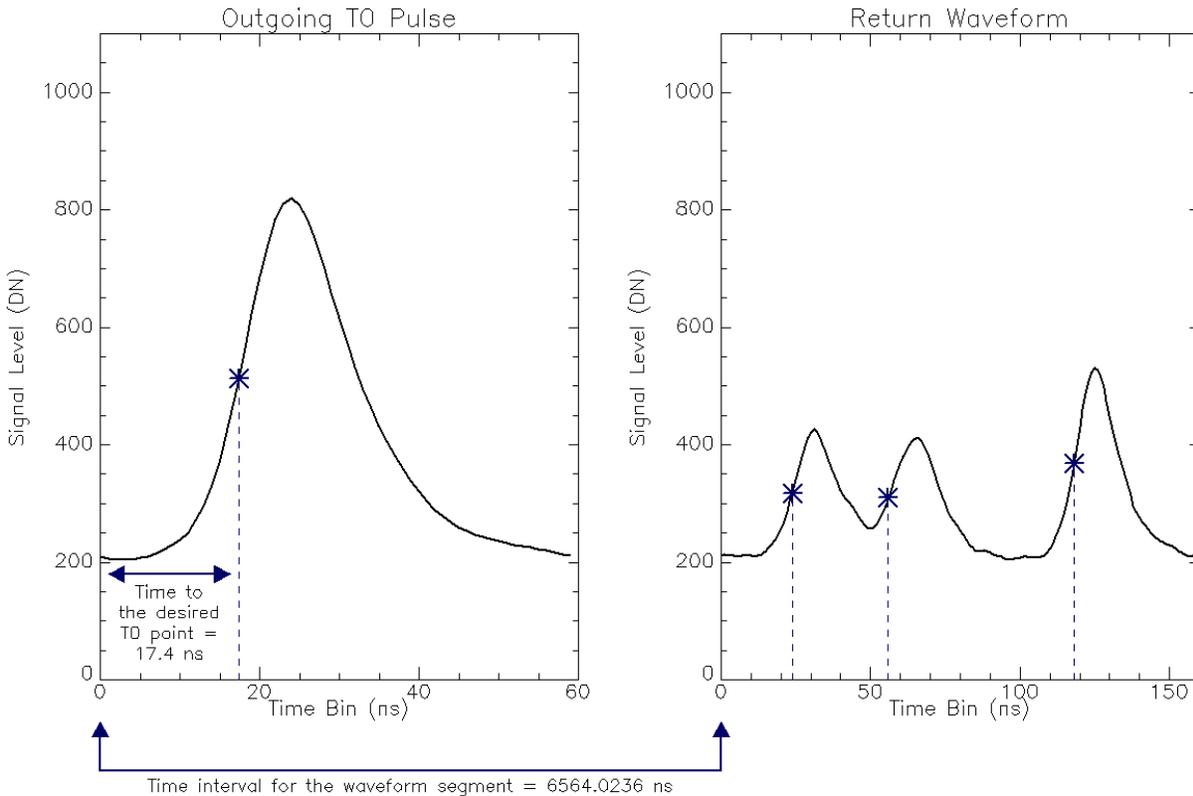
Figure 13. Example of Leading Edge Detection for Time of Flight Calculation for Multiple Returns.

Error sources in the range calculation include ground slope, beam divergence, beam incidence angle, intensity bias, vegetation and atmosphere.

4.2.2   Geolocation

The previous section discusses the calculation of the range between the instrument and a given object on the ground. This is not an absolute range calculation but a relative position in space between the instrument and the object. The data are more useful if those relative positions are transformed into absolute coordinates on a global grid. In order to generate the Level 1 discrete return LiDAR product, it is necessary to calculate the absolute 3D position of the object on the Earth relative to some reference coordinate system. The next processing step is to geolocate the range values to determine the easting, northing and height of each of the discrete returns. Direct georeferencing is used to perform these calculations where the position of an object on the ground is directly computed. For LiDAR processing, five pieces of information are required to georeference each laser shot: GPS data (the location of the instrument), IMU data (the orientation of the aircraft and instrument), the instrument mirror angle (the pointing direction of the laser shot), the range to the object (as discussed in previous section) and calibration information including lever arms and boresight misalignments. The following sets of

equations are adapted from Hutton and Mostafa (ER[09]). The vectors and matrix rotations are illustrated in **Figure 14**:
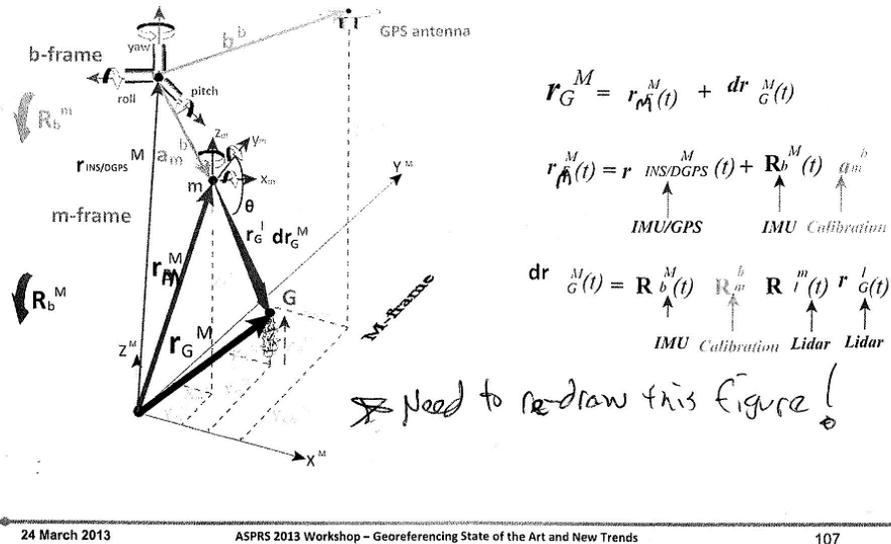
Figure 14. Computation of Ground Coordinates with Direct Georeferencing.

4.2.2.1 Computation of ground coordinates

Computation of ground coordinates is performed in the output mapping frame. Consistent with the four pieces of information required, the computation includes four coordinate frames: the ground coordinate output mapping frame, the IMU body frame, the laser instrument mirror frame and the laser range frame. The direct georeferencing equation is the sum of two vectors

$$r_G^M = r_m^M(t) + dr_G^M(t)$$

Where:

- *G* is an object point on the ground
- *M* is the output mapping frame
- *m* is the instrument reference mirror
- *t* is time

- $r_G{}^M$ is the vector to the object on the ground in the mapping frame (i.e. the coordinates of the ground object)
- $r_m{}^M(t)$ is the vector to the instrument reference mirror in the mapping frame (the coordinates of the instrument) as a function of time
- $dr_G{}^M(t)$ is the vector from the instrument to the object on the ground in the mapping frame (the relative difference in mapping coordinates between the instrument and the object of the ground) as a function of time

The vector to the instrument reference mirror in the mapping frame as a function of time is written as

$$r_m^M(t) = r_{\frac{INS}{DGPS}}^M(t) + \mathbf{R}_b^M(t)\,a_m^b + \mathbf{R}_b^M(t)\,b^b$$

Where:

- INS/DGPS is the inertial navigation system (consisting of an IMU and differential GPS receiver)
- b is the IMU body frame
- $r_m{}^M(t)$ is the vector to the instrument reference mirror in the mapping frame (the coordinates of the instrument) as a function of time
- $r_{INS/DGPS}{}^M(t)$ is the vector to the inertial navigation system/differential GPS in the mapping frame (the coordinates of the IMU body and GPS receiver) as a function of time
- $\mathbf{R}_b{}^M(t)$ is the rotation matrix from the IMU body frame to the output mapping frame (roll, pitch and heading) as a function of time
- $a_m{}^b$ is a set of calibrated lever arms (x, y and z) from the IMU reference point to the laser instrument reference mirror
- b^b is a set of calibrated lever arms (x, y and z) from the IMU reference point to the GPS antenna phase center on the top of the aircraft

Values for b^b are measured during instrument integration into the aircraft and then refined during trajectory processing. Since the typical INS/DGPS processing takes into account the GPS level arms, the resulting smoothed best estimate of trajectory (SBET) file gives the position coordinates for the IMU body/GPS receiver location, reducing the equation to

$$r_m^M(t) = r_{\frac{INS}{DGPS}}^M(t) + \mathbf{R}_b^M(t)\,a_m^b$$

Furthermore, the lever arms from the INS/DGPS to the laser instrument reference mirror may also be included in the trajectory processing directly giving $r_m{}^M(t)$.

The second vector in the direct georeferencing equation is the vector from the instrument to the object on the ground in the mapping frame as a function of time and can be written as

$$dr_G^M(t) = \mathbf{R}_b^M(t)\mathbf{R}_m^b\mathbf{R}_l^m(t)r_G^l(t)$$

Where:

- l is the laser range
- $dr_G^M(t)$ is the vector from the instrument to the object on the ground in the mapping frame (the relative difference in coordinates between the instrument and the object of the ground) as a function of time
- $\mathbf{R}_b^M(t)$ is the rotation matrix from the IMU body frame to the output mapping frame (roll, pitch and heading) as a function of time
- $\mathbf{R}_m^b$ is the rotation matrix from the lidar instrument reference mirror frame to the IMU body frame (the boresight calibration alignment angles as discussed in Appendix A)
- $\mathbf{R}_l^m(t)$ is the rotation matrix from the laser range frame to the laser instrument reference mirror frame (the scanning mirror encoder angle) as a function of time
- $r_G^l(t)$ is the vector from the laser instrument to the ground target (laser range from the mirror to point G) as a function of time

In more detail, the four components of the vector can be defined as

$$r_G^l(t) = \begin{bmatrix} 0 \\ 0 \\ \rho \end{bmatrix}^l$$

Where ρ is the range from the mirror to point G on the ground

$$R_l^m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\cos\omega & \sin\omega \\ 0 & -\sin\omega & -\cos\omega \end{bmatrix}$$

Where ω is the mirror encoder angle

$$R_m^b = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\Theta_x & \sin\Theta_x \\ 0 & -\sin\Theta_x & \cos\Theta_x \end{bmatrix} \begin{bmatrix} \cos\Theta_y & 0 & -\sin\Theta_y \\ 0 & 1 & 0 \\ \sin\Theta_y & 0 & \cos\Theta_y \end{bmatrix} \begin{bmatrix} \cos\Theta_z & \sin\Theta_z & 0 \\ -\sin\Theta_z & \cos\Theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Where Θ_x, Θ_y and Θ_z are the boresight alignment angles determined during calibration

$$R_b^M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Where φ is the roll, θ is the pitch and ψ is the heading.

The desire is to calculate the ground point in the mapping frame. To convert to the mapping frame, another matrix rotation is required to convert from the local level frame (East North Up or ENU frame) to Earth-centered Earth fixed frame (ECEF):

$$R_{ENU}^{ECEF} = \begin{bmatrix} -sin\phi cos\lambda & -sin\lambda & -cos\phi cos\lambda \\ -sin\varphi sin\lambda & cos\lambda & -cos\phi sin\lambda \\ cos\phi & 0 & -sin\phi \end{bmatrix}$$

Where ϕ is the latitude and λ is the longitude of the GPS positions. The ECEF coordinates can be translated into latitude and longitude to put r_G^{ECEF} into a geographic frame using well known relationships (Jekeli, 2001, pg.24), as follows

$$\begin{bmatrix} \phi \\ \lambda \\ h \end{bmatrix}_G = \begin{bmatrix} \tan^{-1}\left(\dfrac{z_{ECEF}}{\sqrt{x_{ECEF}^2 + y_{ECEF}^2}} \left(1 + \dfrac{e^2 N \sin\phi}{z_{ECEF}} \right) \right) \\ \tan^{-1}\left(\dfrac{y_{ECEF}}{x_{ECEF}} \right) \\ \dfrac{\sqrt{x_{ECEF}^2 + y_{ECEF}^2}}{\cos\phi} - N \end{bmatrix}$$

where h is the ellipsoidal height measured along the ellipsoidal normal, N is the radius of curvature of the reference ellipsoid in the prime vertical plane, determined as

$$N = \frac{a}{\sqrt{1 - e^2 \sin^2 \phi}}$$

where a is the semi-major axis of the reference ellipsoid, and e is the eccentricity which is determined as

$$e^2 = 2f - f^2$$

and f is the flattening for the specific desired reference ellipsoid. For reference, the values for the semi-major axis and flattening of the GRS80 ellipsoid, which is used in the WGS84 datum, are 6378137 m and 1/298.257222101 respectively. The solution for ϕ in (18) is not a closed form solution, and an iterative procedure must be employed to achieve a result of sufficient accuracy. This requires an initial estimate of ϕ, where h is assumed to be zero, determined as follows

$$\phi_0 \equiv \phi(h = 0) = \tan^{-1}\left(\frac{z_{ECEF}}{\sqrt{x_{ECEF}^2 + y_{ECEF}^2}}\left(\frac{1}{1 - e^2}\right)\right)$$

and through each successive iteration, (18) and (19) must be updated.

Error sources in the geolocation can come from a systematic bias or noise in determination of the lever arm offsets, boresight angles, laser range, mirror scan angle, GPS positions, IMU observations. A quantification of how the noise in each of these components contributes to the overall uncertainty is provided in Section 6.

4.3   Special Considerations

None.

5   ALGORITHM IMPLEMENTATION

The algorithm that produces the Discrete Return LiDAR Point Cloud NEON Level 1 data product NEON.DOM.SITE.DP1.30003 is implemented in a third party software package provided by the LiDAR instrument vendor: Optech LMS LiDAR Mapping Suite 2.4. **Figure 15** shows a high-level flowchart for the discrete return LiDAR processing.
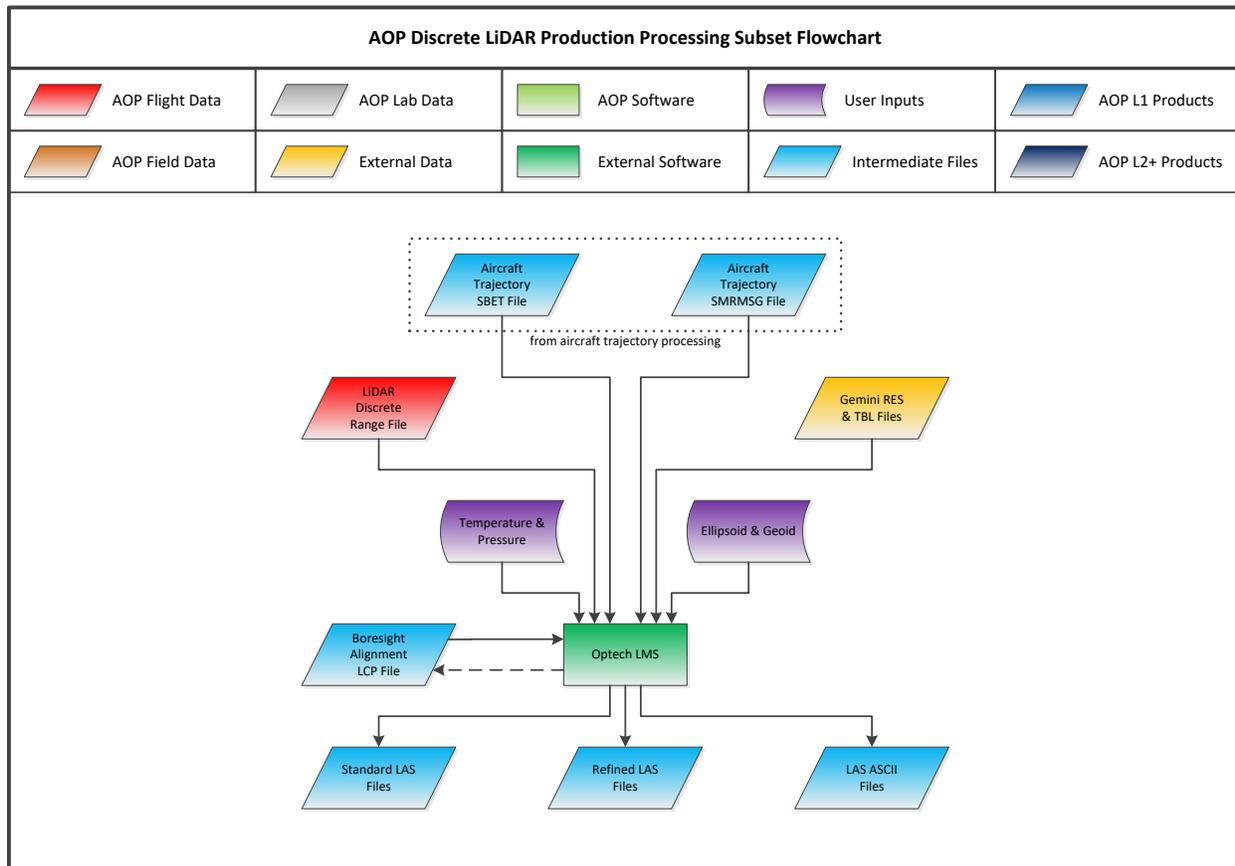
Figure 15. High-Level Discrete Return LiDAR Processing Flowchart.

The inputs include the trajectory (SBET) file, the trajectory precision smrmsg file, the discrete return LiDAR raw range file, the Gemini calibration LCP, res and intensity table files, the temperature and pressure during the data collection and the ellipsoid to geoid height model. The outputs include the standard LAS discrete return point cloud files, the refined LAS files and the ASCII files for each flight line. Also, using special test flight data, the boresight alignment angles in the LCP file can be calibrated and updated as described in Appendix A. This calibration shall be performed once per aircraft installation and should be double checked at the end of the flight season. The LMS workflow is described in the following sections but can be broken down into three sections: standard processing, self-calibration, and refined processing. Specific details for how to set up and run LMS are given in the LMS procedure document (RD[12]). The standard processing generates a discrete return LiDAR point cloud output.

However, points covering the same area on the ground but collected on different flight lines may not be perfectly registered, for instance if the boresight calibration is incorrect or has changed with time or physical installation in a different payload mount. The self-calibration procedure in LMS will use roof lines and planar surfaces (extracted from the standard output) and a least squares regression in order to solve for new calibration parameters that will produce a more self-consistent and registered set of flight lines. The refined processing generates new output files based on the self-calibration improvement.

5.1 Standard Processing

Standard processing reads in the raw data, performs the laser point computation, and generates the output files. The calibration values in the LCP file are used in the laser point computation.

5.1.1 Read Calibration Files

The LMS software reads in the three calibration files: the laser calibration parameters (LCP) file, the RES file, and intensity tables (TBL file). The LCP (laser calibration parameters) file contains the scan mirror scale and offset factors plus the boresight alignment angles Ex, Ey and Ez. The proprietary signal processing algorithms that detect objects in the received signal may introduce a bias in the range estimation. The intensity tables contain range adjustment values as a function of detected return intensity for each of the PRF settings and for 1st, 2nd, 3rd, and 4th returns.

5.1.2 Read and Decode Range File

The LMS software reads in the raw range (.range) file, which is in an Optech proprietary format. The file is parsed to determine the start and stop GPS times for each flight line plus collect other instrument configuration data for each flight line. LMS then generates a flight line trajectory report with the following values for each flight line:

- Strip ID
- Start [s]
- Stop [s]
- Duration [s]
- PRF [KHz]
- Scan Frequency [Hz]

Other discrete return point data is read in such as gps time of the laser shot, time-of-flight, mirror scan angle, return intensity, and return number.

5.1.3 Read Trajectory Files

The LMS software reads in the trajectory files. The trajectory files include the smoothed best estimate of trajectory (SBET) file (SBET.out) and a set of residual precision values (smrmsg.out). The SBET is a binary file that provides aircraft longitude, latitude, height, roll, pitch and heading versus GPS time. The

smrmsg file is also a binary file that gives residual trajectory precision values that can be used to calculate laser point accuracy. LMS appends the flight line trajectory report with the following values for each flight line:

- Speed Avg [m/s]
- Height Avg [m]
- Roll [deg]: min, avg, max
- Pitch [deg]: min, avg, max
- Heading [deg]: min, avg, max
- Roll Std.dev [deg]: min, avg, max
- Pitch Std.dev [deg]: min, avg, max
- Heading Std.dev [deg]: min, avg, max
- East Std.dev [deg]: min, avg, max
- North Std.dev [deg]: min, avg, max
- Height Std.dev [deg]: min, avg, max

5.1.4 Read Meteorological Data

The LMS software reads in the values for the meteorological data, specifically the temperature (deg C) and pressure (mBar or hPa). Temperature and pressure values are recorded at the beginning of the flight at the airport, however it is recommended to use the average of the ground and aircraft altitude values if known. Also, the algorithm allows the input of a meteorological file that would provide temperature and pressure versus time. These values are used to calculate the index of refraction of air in order to calculate the correct laser range based upon the time of flight between the laser pulse and the received signal.

5.1.5 Laser Point Computation

This step of the algorithm takes all of the shot data and calculates the range in air to the target, performs a geolocation of the laser range (the point of reflection of the laser pulse), and converts the latitude, longitude, and height of the return into the output projection and datum. The inputs for the geo-referencing are R_a (Range in air), θ_Z (Scanner angle) and all the parameters in the SBET (Smooth Best Estimate Trajectory) file interpolated to match the GPS times of the lidar's repetition rate. The mathematics of the algorithm are provided by the LiDAR instrument vendor (ER[10]) and are described in the following sections.

5.1.5.1 Calculate range

Before range is determined, the index of refraction in air must be calculated. At 1064 nm this can be approximated by

$$n = 1 + 78.7 \times \frac{P}{(273.15 + T)} \times 10^{-6}$$

Where T is the mean atmospheric temperature in deg C and P is the mean pressure in millibars (mbar) or hectopascals (hPa). The laser range in air is calculated as

$$R_a = \frac{c}{n} \cdot \frac{\tau}{2}$$

Where c is the speed of light and τ is the time-of-flight of the return. The range values are also slightly adjusted using the intensity table correction values from the TBL file.

5.1.5.2   Calculate XYZ in the sensor body frame

$$X = R_a \sin(\theta_y)$$
$$Y = R_a \sin(\theta_z)\cos(\theta_y)$$
$$Z = R_a \cos(\theta_z)\cos(\theta_y)$$

Where, θ_z is the scanner angle and θ_y is the off-nadir angle ($0°$ in Gemini), R_a is the laser range in air.

θ_y is $0°$ in the standard AOP payload integration mount. If the LiDAR instrument was mounted to look at the Earth at an oblique angle, then this angle would need to be used in the calculations.

5.1.5.3   Laser IMU lever arm offset correction

$$X' = X + \Delta x$$
$$Y' = Y + \Delta y$$
$$Z' = Z + \Delta z$$
$$X = X', Y = Y', Z = Z'$$

Where, Δx, Δy, Δz are the IMU lever arm offsets. Note that this correction is not required because the IMU is programmed to record navigation information with respect to the sensor body frame.

5.1.5.4 Laser-IMU Misalignment correction

Correct for the IMU misalignment correction (boresight angles between the instrument optical axis and the IMU body frame) by rotating through the misalignment angles to yield corrected [X, Y, Z]:

θ_x:

$$R(\theta_X) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_X & -\sin\theta_X \\ 0 & \sin\theta_X & \cos\theta_X \end{bmatrix}$$

θ_Y:

$$R(\theta_y) = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix}$$

θ_z:

$$R(\theta_Z) = \begin{bmatrix} \cos\theta_Z & -\sin\theta_Z & 0 \\ \sin\theta_Z & \cos\theta_Z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

5.1.5.5 Roll, pitch, heading correction

Correct for the roll (θ_x), pitch (θ_y) and heading (θ_z) by multiplying the resulting [X,Y,Z] in section 5.1.5.4 to yield [X, Y, Z] in local East, North, Up (ENU) frame:

θ_X :

$$R(\theta_X) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_X & -\sin\theta_X \\ 0 & \sin\theta_X & \cos\theta_X \end{bmatrix}$$

θ_Y :

$$R(\theta_y) = \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix}$$

θ_Z :

$$R(\theta_Z) = \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

5.1.5.6   ENU frame of reference to Geodetic

This is a two set process: a. Transforming the resultants [X,Y,Z] in ENU frame in section 5.1.5.5 to Earth-Centered Earth-Fixed (ECEF) frame; b: Transforming the [X,Y,Z] in ECEF frame to geodetic coordinates (WGS-84) yielding latitude (φ), longitude (λ) and height (h) for the laser shot

5.1.5.6.1 ENU to ECEF

$X_{ECEF} = -X\sin(\varphi)\cos(\lambda) - Y\sin(\lambda) - Z\cos(\varphi)\cos(\lambda)$

$Y_{ECEF} = -X\sin(\varphi)\sin(\lambda) + Y\cos(\lambda) - Z\cos(\varphi)\sin(\lambda)$

$Z_{ECEF} = X\cos(\varphi) - Z\sin(\varphi)$

5.1.5.6.2 ECEF to Geodetic

a = 6378137; [Semi-major earth axis]

b = 6356752.31424518; [Semi-minor earth axis]

$$e = \sqrt{\frac{a^2 - b^2}{a^2}} \; ; \text{[Eccentricity]}$$

$$f = \frac{a - b}{a}$$

$$N = \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi}} \; ; \text{[Radius of curvature]}$$

$$X_{GPS} = (N + H)\cos(\varphi)\cos(\lambda)$$
$$Y_{GPS} = (N + H)\cos(\varphi)\sin(\lambda)$$
$$Z_{GPS} = \left(\left(\frac{b^2}{a^2}\right) \cdot N + H\right)\sin(\varphi)$$

$X = X_{ECEF} + X_{GPS}$

$Y = Y_{ECEF} + Y_{GPS}$

$Z = Z_{ECEF} + Z_{GPS}$

$$P = \sqrt{X^2 + Y^2}$$

$$\lambda = \tan^{-1}(Y, X)$$

$$\varphi = \tan^{-1}\left(\frac{Z}{P \cdot (1 - e^2)}\right)$$

Iterate the next steps until two consecutive iterations result in negligible difference for φ and h:

$$N = \frac{a}{\sqrt{1 - e^2 \sin^2(\varphi)}}$$

$$h = \frac{P}{\cos(\varphi)} - N$$

$$\varphi = \tan^{-1}\frac{Z}{P\left(1 - e^2 \cdot \dfrac{N}{N + h}\right)}$$

The resulting $[\varphi, \lambda, h]$ for every discrete lidar return are the necessary geodetic coordinates with reference to WGS-84 ellipsoid, for a given laser range in air and pointing vectors

5.1.5.7 Convert to Output projection and Geoid Height

The georeferencing equations are designed to perform the calculations in ellipsoid space, which is mathematically simpler. The lat/lon geodetic coordinates are converted to the output projection, in this case as Easting and Northing values in a Universal Transverse Mercator (UTM) mapping projection using the International Terrestrial Reference Frame 2000 (ITRF 2000) ellipsoid as a horizontal datum in units of meters. Conversion from a geographic coordinates to UTM can be performed with formulas detailed in ER [11]. The Z coordinates will be reported in a North American Vertical Datum 1988 (NAVD88) using the National Geodetic Survey Geoid12A height model with units of meters. The NOAA National Geodetic Survey maintains the Geoid12A model (ref). This model is a series of geometric harmonics that describes where the mean sea height would be at a given location on Earth based on Earth's gravity and rotation. In a practical application of the model, look up tables are generated to convert from ellipsoid to geoid for a given spatial coordinate on the ground.

5.1.6 Generate Output Files

LMS reformats the data and generates output files in ASPRS LAS 1.3 format (described in Section 3.1) with GPS time in seconds of week. ASCII output files are also required as intermediate products and are used as inputs to the Waveform LiDAR processing algorithm (RD[06]). The primary reason for generating

ASCII files is to extract the scan mirror angle position versus time. The full required contents of the ASCII files are TBD, but possible information required by the waveform algorithm is:

- Laser shot GPS time
- Laser shot scan angle
- Laser shot number of ranges
- Laser range
- Laser range intensity
- Laser point X-coordinate
- Laser point Y-coordinate
- Laser point Z-coordinate

LMS has the option to include standard deviation values in the ASCII output that could be used for accuracy assessment and quality control/assurance in future versions of the waveform algorithm.

5.2 Self-Calibration

During standard processing, points covering the same area on the ground but collected on different flight lines may not be perfectly registered, for instance if the boresight calibration is incorrect or has changed with time or physical installation in a different payload mount. The self-calibration procedure in LMS uses roof lines and planar surfaces (extracted from the standard output) and a least squares regression in order to solve for new calibration parameters that will produce a more self-consistent and registered set of flight lines. The user can set the line type as production (default) or cross. Cross strips are used in the self-calibration processing step by extracting planar surfaces and roof lines and bundle adjusting the production lines with a least-squares regression. In this case, a cross strip provides more area of data overlap between flight lines.

5.2.1 Planar Surface Extraction

LMS will extract planar surfaces from the standard output using a proprietary feature extraction algorithm. **Figure 21** in Appendix A shows an example of some roof tops that were extracted. The user can adjust several parameters used in the extraction such as the plane length, plane width, and surface roughness.

5.2.2 Roof Line Extraction

LMS will extract roof lines from the standard output using a proprietary feature extraction algorithm. **Figure 21** in Appendix A shows an example of some roof lines that were extracted. The user can adjust several parameters used in the extraction such as the minimum roof slope, maximum roof slope, max. azimuth difference between roof planes, max. slope difference between roof planes, max. distance between roof plane centers, and the shortest acceptable roof line length.

5.2.3   Self-Calibration Block Adjustment

The self-calibration step takes the planar surface and roof line extraction data and compares the same object as seen by two flight lines. The differences in position and orientation of those objects are used to block adjust the flight lines and determine new self-calibration values (refined LCP values) and other corrections. Ideally the mean differences between flight lines after self-calibration should be zero and the rms values should decrease (see **Figures 22** and **23** in Appendix A for an example). Based on the default correction settings, LMS will update the scan angle scale and boresight angle Ez for the sensor and apply a Z position correction for each flight line and also a roll and pitch correction for each flight line.

5.2.3.1   Sensor Corrections

Sensor corrections solve for errors in the instrument scan mirror angle value. LMS will use the least-squares regression to solve for the best scan angle scale factor which converts the raw mirror encoder angle into a calibrated scanner angle (θ_z from section 5.1.5.2). One value is determined for the entire flight.

5.2.3.2   Boresight Corrections

Boresight corrections solve for errors in the angular alignment of the lidar optical boresight with the IMU body frame. LMS will use the least-squares regression to solve for the best values of $\Delta\Theta_x$, $\Delta\Theta_y$, and $\Delta\Theta_z$. Note that during production processing, only $\Delta\Theta_x$ and $\Delta\Theta_y$ are updated for the refined processing. All three angles are solved for during a boresight calibration as described in Appendix A. One value is determined for the entire flight.

5.2.3.3   Position Corrections

Position corrections solve for errors in the X, Y and Z GPS position of the instrument. Note that during production processing, only Z is updated for the refined processing. Position correction values are determined for each individual flight line.

5.2.3.4   Orientation Corrections

Orientation corrections solve for errors in the roll, pitch and heading attitude of the instrument. Note that during production processing, only heading is updated for the refined processing. Orientation correction values are determined for each individual flight line.

5.3   Refined Processing

The new self-calibrated LCP values plus flight line position and orientation corrections are used to re-process the data to arrive at a refined output that will produce a more self-consistent and registered set

of flight lines. The refined processing repeats the laser point computation using the new sensor parameters and formats the output files.

5.3.1 Laser Point Computation

Same as section 5.5 but with the new sensor parameters.

5.3.2 Generate Output Files

Same as section 5.1.6.

5.4 Analyze Results

After LMS has processed the data, a series of summary reports and figures are generated. It may be necessary to adjust a few parameters and re-process the data. In the case of forests, LMS may have trouble extracting planar surfaces. This problem is typically solved by increasing the surface roughness. It is also desirable for at least 10% of all tie planes to be selected for the self-calibration. LMS reports the number of selected tie planes and total number of identified tie planes. If the percentage used is too low, then the total number of points for self-calibration must be increased under the block processing planar surface extraction parameters.

6   ANALYSIS OF UNCERTAINTY

Uncertainty in LIDAR point cloud observations is affected by several factors, both internal and external to the sensor. Several studies have been conducted which empirically assessed the errors in LIDAR point clouds and derived products (for example, Huising and Gomes Pereira, 1998; Hodgson and Bresnahan, 2004; Hyyppa et al. 2005; Goulden and Hopkinson, 2010; Joerg et al., 2012), isolating several identifiable error sources. Based on these assessments, Goulden and Hopkinson (2014) broadly defined the sources of uncertainty of airborne LiDAR observations into three categories

> 1) individual hardware components of the sensor (GPS, IMU, laser ranger, laser scanner),
>
> 2) terrain effects such as terrain slope and reflectivity,
>
> 3) land cover such as vegetation.

Given that this document pertains to the production of LIDAR point clouds directly from raw sensor data, and the state of the terrain and land cover is unknown, the uncertainty in this product can be confined to contributions from category one; errors in the individual sensor hardware components.

The geolocation of airborne LiDAR observations requires time-varying measured data streams from the following individual LiDAR sensor subsystems (detailed in Section 4.2.2.1); 1) the aircraft position from the GPS antenna and receiver, 2) aircraft orientation from the IMU, 3) ranges to objects from the laser ranger and 4) a scan angle from an angular encoder. These data streams are used in the direct geo-referencing equation (Section 4.2.2.1) to determine the three-dimensional coordinate (x,y,z) of objects intercepted by a laser pulse. Each of the measurements made by the sensor sub-systems has an inherent error that will affect the uncertainty in final coordinates. The measurement error of the hardware sub-system observations can be propagated through the LIDAR direct geo-referencing equation to produce an uncertainty in the final point coordinate. Details of the algorithm and success of this approach for quantifying observed LiDAR point coordinate uncertainties has been demonstrated in Glennie (2008) and Goulden and Hopkinson (2010).

The approach taken by Goulden and Hopkinson (2010) to quantify the uncertainty in LIDAR point positions utilized the General Law of Propagation of Variances (GLOPOV, Wolf and Ghilani, 1997). The GLOPOV states that the sum of the squares of the partial derivative of each observable quantity, multiplied by the observation's predicted error (in the form of a variance or covariance), will result in the total propagated error. The GLOPOV methodology can be generally written as

$$\mathbf{C}_{Target} = \mathbf{A}\mathbf{C}_{Obs}\mathbf{A}^T \tag{6.1}$$

where

$$\mathbf{A} = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \dfrac{\partial F_1}{\partial x_2} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \dfrac{\partial F_2}{\partial x_1} & \dfrac{\partial F_2}{\partial x_2} & \cdots & \dfrac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & \dfrac{\partial F_m}{\partial x_2} & \ddots & \dfrac{\partial F_m}{\partial x_n} \\ & & \cdots & \end{bmatrix} \tag{6.2}$$

$$\mathbf{C}_{obs} = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \cdots & \sigma_{x_1 x_n} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & \cdots & \sigma_{x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_n x_1} & \sigma_{x_n x_2} & \cdots & \sigma_{x_n}^2 \end{bmatrix} \tag{6.3}$$

and

$$\mathbf{C}_{Target} = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \cdots & \sigma_{x_1 x_m} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & \cdots & \sigma_{x_2 x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_m x_1} & \sigma_{x_m x_2} & \cdots & \sigma_{x_m}^2 \end{bmatrix} \tag{6.4}$$

In **A**, *F* represents the function which calculates our desired variable. For the LiDAR case, *F* is the direct geo-referencing equation as formulated in Section 4.2.2.1. The different enumerations (1 to *m*) of *F* in **A** represent the different output variables for which we are establishing uncertainty. For each intercepted object in the LiDAR point cloud we are interested in the uncertainty in the *x*, *y* and *z* coordinates (*m*=3).

The enumerations (1 to *n*) for *x* in **A** and $\mathbf{C}_{Obs}$ represent the different variables in the direct geo-referencing equation which have known error. Measured quantities with known error include roll, pitch, yaw, the aircraft position (*x*, *y*, *z*), the laser range and the scan angle. Therefore, there are eight measured quantities with an associated error (*n*=8), leaving **A** with dimensions of three by eight for the LiDAR case presented here.

C represents a covariance matrix, which contains either the uncertainty in the final output coordinates ($\mathbf{C}_{Target}$, 3 x 3 matrix) or the measurement errors ($\mathbf{C}_{Obs}$, 8 x 8 matrix). We assume the errors in the measured sensor observations are statistically independent, therefore, only the main diagonal of $\mathbf{C}_{Obs}$ is populated, and the remaining entries are zero. If there was a known covariance between sensor hardware components the appropriate non-diagonal elements would require quantification. The assumption of statistical independence is valid for the range and scanner errors, however, is not valid for the position and orientation errors due to the tightly coupled processing algorithm implemented in the generation of the airborne trajectory. Violating the independence assumption will likely lead to an overestimation of uncertainty estimates, which is more desirable than an underestimation, therefore the assumption of statistical independence of the GPS/IMU positioning results is maintained. $\mathbf{C}_{Target}$ will generally be fully populated in operational scenarios with uncertainty in the individual coordinates along the main diagonal, and the associated covariance in the off-diagonal elements.

Table 3 provides approximations for the errors of sensor sub-components for the Optech ALTM Gemini system in-use at NEON at standard confidence (≈68.4), which populate the main diagonal of $\mathbf{C}_{Obs}$. The listed values for the GPS and IMU were obtained from commercial literature for the POS AV 510 system ER [12]; however these errors vary temporally and are dependent on the GPS constellation conditions and baseline lengths to the nearest GPS basestation. To achieve a rigorous estimate of the uncertainty in LiDAR derived coordinates, time-varying error estimates of the aircraft trajectory are obtained from Applanix's PosPAC software (**Figure 16**). The value for the scan angle error is obtained from Goulden and Hopkinson (2010), who tested the scan angle error listed in Table 6-1 for an Optech ALTM 3100 EA, a similar sensor configuration to the Optech Gemini. Goulden and Hopkinson (2010) also determined an error in the laser ranger of 1.5 cm, based on the precision of the time counter in typical laser ranging systems of ±0.1 ns (Baltsavias, 1999). However, testing of NEON's Gemini system has shown a consistent ranging precision over hard flat surfaces of 4-5 cm when operated under nominal operational configurations (100 kHz, wide beam divergence); therefore, the empirically derived uncertainty in ranging is used for the ranging error.
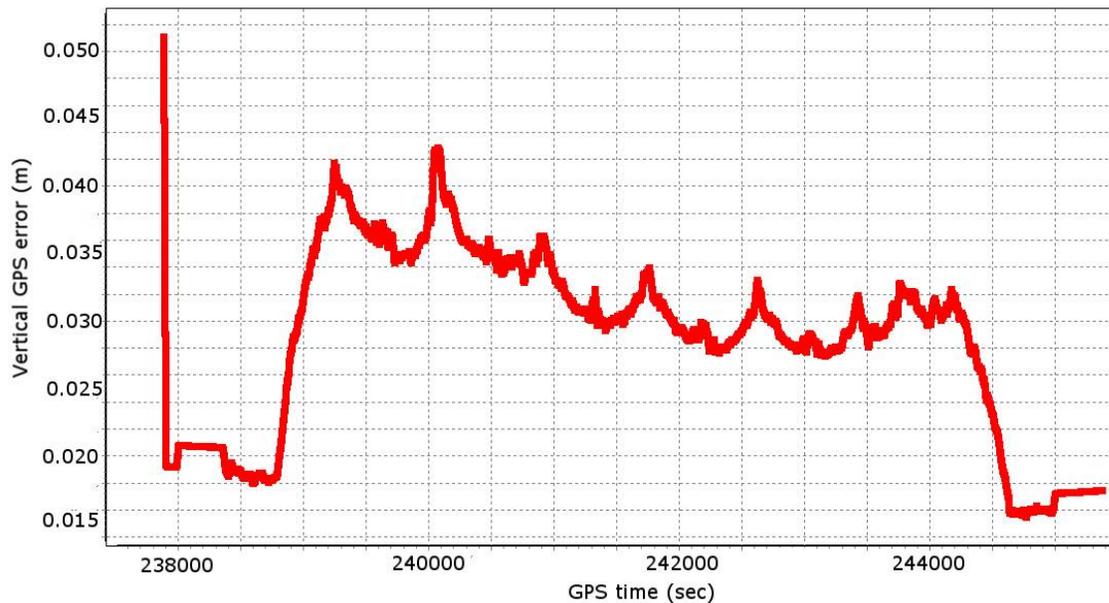
Table 3. Error estimates used to populate C*obs.*

Observation	Error Magnitude
GPS[†]	Horizontal: 0.03-0.05 m
	Vertical: 0.05-0.10 m

IMU[†]	Roll / pitch: $0.005 - 0.01°$
	Yaw: 0.01-0.02°
Scan angle	0.003°
Beam divergence	0.8 mRad
Laser Range	0.04-0.05 m

† Typical values given by for the POS AV 510 GPS / IMU system

Figure 16. Example of time-varying error in the observed vertical GPS coordinate as produced by Applanix's POSPac software.

Once $\mathbf{C}_{Obs}$ is appropriately populated, equation 6.1 can be used to determine $\mathbf{C}_{Target}$ for each individual LiDAR point observation in the survey. To demonstrate the quantification of uncertainty using equation 6.1 on a NEON example case, the error simulation routine implemented in Goulden and Hopkinson (2010) was applied to a NEON survey of the city of Boulder conducted on October 8th. The flight parameters of the survey are listed in **Table 4**, and the survey was conducted in a north-south flight line pattern. Results of the error simulation algorithm show a vertical striping pattern in the spatial distribution of the z-coordinate uncertainty which corresponds to the flight line locations (**Figure 17**). Goulden and Hopkinson (2010) previously identified that the uncertainty in LiDAR point coordinates introduced by instrument sources is directly proportional to scan angle, roll, pitch, and flying height. This leads to a tendency for the largest errors to exist at swath edges, and the minimum errors to exist at

nadir, also indicating that the selected flight pattern will exert some control over the spatial distribution of point coordinate uncertainty. The dependence of the spatial distribution of uncertainty to flight-line orientation was also empirically observed in Li et al. (2011).

Table 4. Flight parameters of Oct. 8th City of Boulder survey.

Parameter	Value
Mean flight altitude	~1800 m a.g.l.
Pulse repetition frequency	70 kHz
Scan frequency	33 Hz
Beam divergence	0.8 mRad
Half scan angle	18°
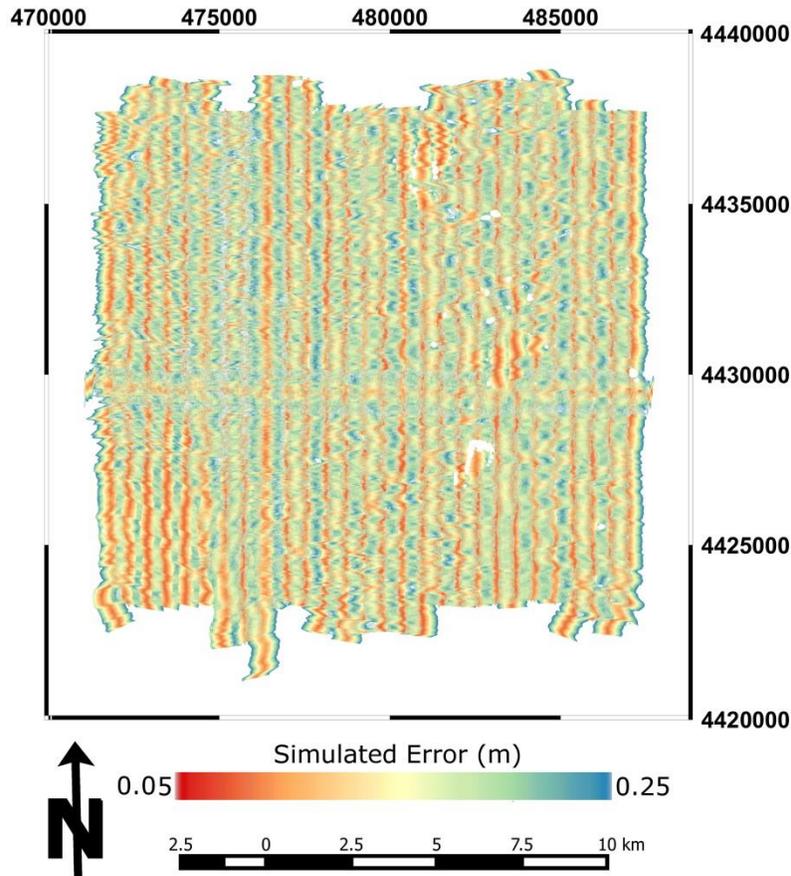
Figure 167. Simulated error from lidar sensor hardware sources for the October 8th 2013 survey of the city of Boulder, Colorado.

The minimum uncertainty occurs at the center of the swaths (red areas in **Figure 17**), and increases as the scan angle increases to the scan line edges (blue areas in **Figure 17**). At a minimum, the uncertainty in the z coordinate $\left(\sigma_{z_{Target}} \right)$ was approximately 0.04 m, as a result of a low GPS vertical error of approximately 1 cm. The upper limit of ε_{z_T} was approximately 0.36 m, where 95% of ε_{z_T} fell below this value. The Boulder survey was flown higher than typical NEON science collects, indicating errors under nominal survey configurations will be slightly lower, especially on near swath edges where the scan angle reaches its maximum. The minimum potential vertical uncertainty can be easily quantified with a simplification of the GLOPOV is which includes only the combination of the vertical error in the GPS coordinate and the error in the observed range. Using the error conditions observed in the Boulder survey of a vertical GPS error of ~0.01 cm and a laser range error of 0.04 cm, an approximation of the minimum possible error in the vertical coordinate is

$$\min\left(\sigma_{z_{Target}} \right) = \sqrt{\left(\varepsilon_{GPS_z}{}^2 + \varepsilon_{LaserRange}{}^2 \right)} = \sqrt{(0.01^2 + 0.04^2)} \approx 0.041 \text{ m} \qquad 6.5$$

where $\varepsilon_{z_{GPS}}$ is the error in the vertical component of the observed coordinate and ε_ρ is the error in the laser range. Equation 6.5 assumes the pulse is transmitted at an instant when roll, pitch and scan angle are equal to zero, the pulse intercepts a hard flat surface and all systematic biases have been identified and removed. The approximation generally holds valid for the nadir portions of the swath in the Boulder survey example case (**Figure 17**).

The horizontal uncertainty will occur with the same spatial pattern as the vertical uncertainty, but with different magnitude. Horizontal error will be most identifiable in sloped terrain or where breaklines exist in the landscape. Although horizontal error exists on flat surfaces, it should not be considered a concern because it will not lead to additional vertical error, and will not propagate further into downstream processing. A simplification of the GLOPOV can also be used to quantify the minimum horizontal uncertainty which is a combination of the error in the scan angle and GPS horizontal error. If we consider a pulse which intercepts a distinct breakline in the landscape, such as a building edge, the **minimum** horizontal error is achieved if the pulse is transmitted at an instant when roll, pitch, and scan angle are zero, and the beam transmission direction is in the same vertical plane as the building edge (beam divergence can be ignored). Under these circumstances, if the plane is flying at an altitude of 1000 m, the error in the scan angle will lead to a horizontal error of

$$\tan(\varepsilon_\theta) * (FlyingHeight) = \tan(0.003°) * 1000 \, m \approx 0.052 \text{ m} \qquad 6.6$$

where ε_θ is the error in the scan angle. The error due to the scan angle can be combined with the GPS error similarly to equation 6.5 to produce

$$\min\left(Error_{xy_T}\right) = \sqrt{\left(\varepsilon_{xy_{GPS}}{}^2 + \varepsilon_\theta{}^2\right)} = \sqrt{(0.03^2 + 0.052^2)} \approx 0.06 \text{ m} \qquad 6.7$$

However, if the beam direction does not lie in the same vertical plane as the building edge an additional error with magnitude approximately equal to the width of the laser beam footprint can exist (Goulden and Hopkinson, 2010). As NEON typically operates the Optech Gemini in a wide beam divergence mode, the associated footprint diameter at 1000 m flying height leads to a beam footprint diameter of 0.8 m, indicating **maximum** horizontal errors in excess of 0.4 m (half beam divergence) could be observed in the presence of breaklines in the landscape. For uncertainty quantification, no assumption about the placement of the beam centerline is made. A worst-case scenario approach is taken which assumes the maximum case of horizontal error which includes the half beam width, therefore,

$$\left(Error_{xy_T}\right) = \sqrt{\left(\varepsilon_{xy_{GPS}}{}^2 + \varepsilon_\theta{}^2 + \varepsilon_\gamma{}^2\right)} = \sqrt{(0.03^2 + 0.052^2 + 0.4^2)} \approx 0.402 \text{ m} \qquad 6.8$$

where ε_γ is the error due to beam divergence. It should be noted that horizontal errors can only be combined in the manner displayed in equation 6.7 and 6.8 if they have the same direction in the horizontal plane. If $\varepsilon_{xy_{GPS}}$, ε_θ and ε_γ are in different directions, the combined error will be slightly reduced. Similar to equation 6.5, the relationship in 6.8 assumes the pulse is transmitted at an instant when roll, pitch and scan angle are equal to zero and is provided only as a reasonably equivalent simplification of the actual propagation of horizontal uncertainty with the GLOPOV.

7   VALIDATION AND VERIFICATION

7.1   Algorithm Validation

TBR. Use regression testing with known inputs and outputs.

7.2   Data Product Validation

TBR. Could verify the geolocation accuracy by surveying ground control points and then identifying those points in an intensity image derived from the LiDAR data. Both horizontal and vertical accuracy could be validated.

7.3   Data Product Verification

TBR

8   SCIENTIFIC AND EDUCATIONAL APPLICATIONS

3D visualization

Colorized point clouds

Digital surface model, bare-Earth digital elevation model and canopy height model generation

High resolution terrain generation

Vegetation structure, Cover, Density, spatial configuration

Watershed analysis – flow routing, watershed delineation

Change detection: landcover, terrain, landuse, vegetation

Wildlife Habitat analysis

Riparian cover analysis

9   FUTURE MODIFICATIONS AND PLANS

The following items are being considered for future inclusion in the algorithm for generating this data product:

1. Add product metadata file
2. Add KML flight line boundary file
3. Add geolocated browse image
4. Classification of points
5. Noise filtering
6. LAZ compression
7. Colorized point clouds
8. ASCII output standard deviation values

10   CHANGELOG

N/A

11   APPENDIX A   BORESIGHT ALIGNMENT CALIBRATION

The boresight alignment angles between the LiDAR instrument reference mirror frame and the IMU body frame are calibrated in the aircraft using a specific flight configuration. This configuration includes fly seven flight lines, all at the nominal flying height. Three parallel flight lines are flown with 60% side overlap between lines and the center line has a heading 180 degrees from the other two lines (opposite direction). The next three lines follow the same configuration except they are perpendicular headings (+90 degrees) to the first three. Finally a single line is flown on the diagonal with a heading of 45 degrees apart from the other lines. This flight is conducted over an area with lots of roof targets (roof lines and planar surfaces) such as a suburb. An example boresight alignment calibration flight plan is shown in **Figure 16**.

Figure 17. Example boresight alignment calibration flight plan over Grand Junction, CO.

The flight trajectories are loaded into the LMS software and the boresight alignment calibration area is cropped out from the rest of the data as shown in **Figure 17**:
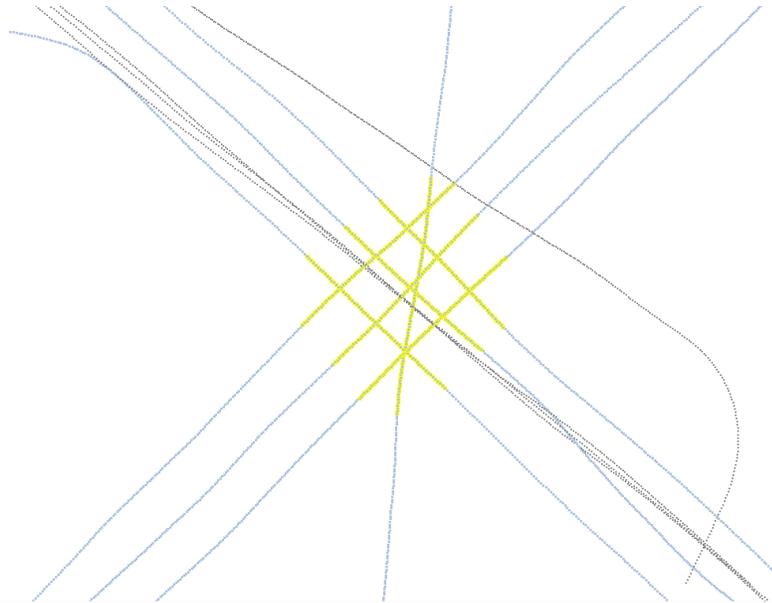
Figure 18. Data from boresight alignment calibration area is cropped.

A 3D view of the cropped point cloud data is shown in **Figure 18** and a close up view in **Figure 19** where blue areas have lower elevations and red areas have higher elevations.

Figure 19. 3D point cloud data of the boresight calibration area.

Figure 20. 3D close up view of the boresight calibration area.

During standard processing, the Optech LMS software extracts roof lines and planar surfaces from the 3D point clouds of the boresight calibration area. These lines and surfaces are shown in **Figures 20** and **21**:
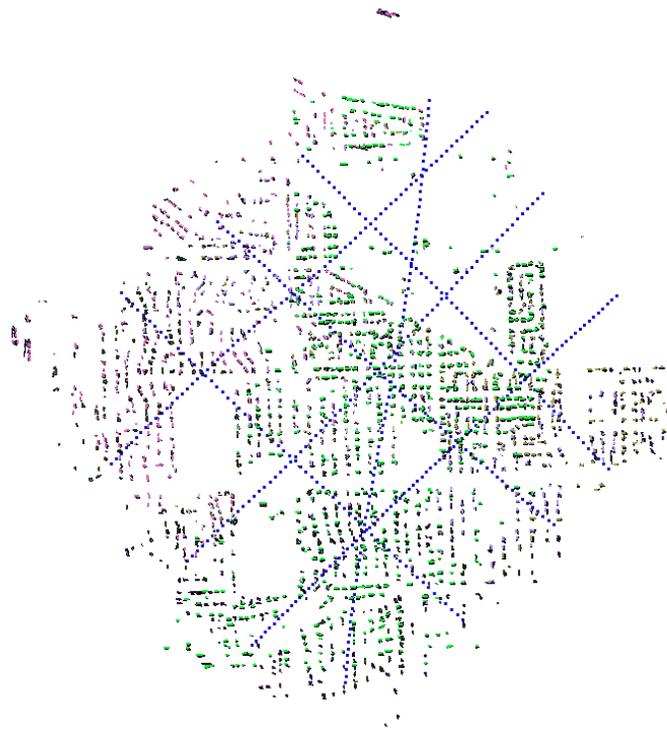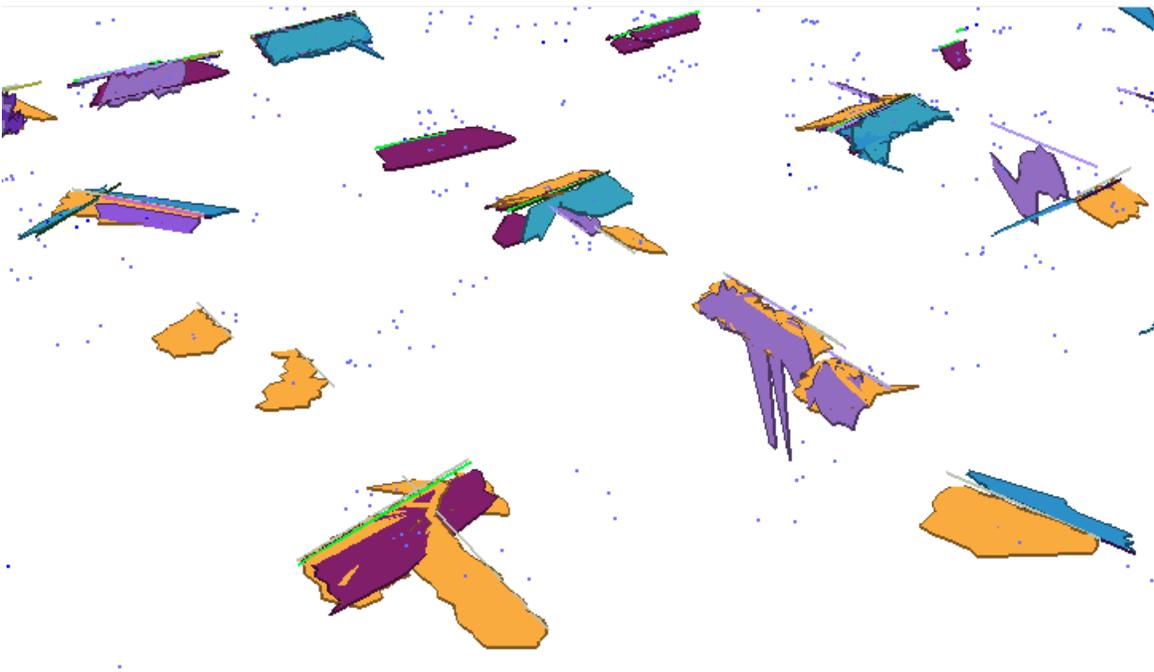
**Figure 21**. Roof line and surface extraction in LMS.

Figure 22. A close up view of the roof line and surface extraction in LMS.

The LMS software uses a least squares regression to solve for boresight alignment angles Ex, Ey and Ez plus a scan mirror scan angle scale factor. These values are updated in the LCP file during boresight calibration. **Figure 22** shows the vertical difference between a tie plane identified in 2 different flight lines (hopefully the same planar surface) versus scan angle. Here the differences have a slope versus scan angle where tie plane differences are negative for negative scan angles and positive for positive scan angles. After performing the boresight alignment calibration, this same plot has less of a trend versus scan angle and most tie plane differences are close to 0.0 m, as shown in **Figure 23**.
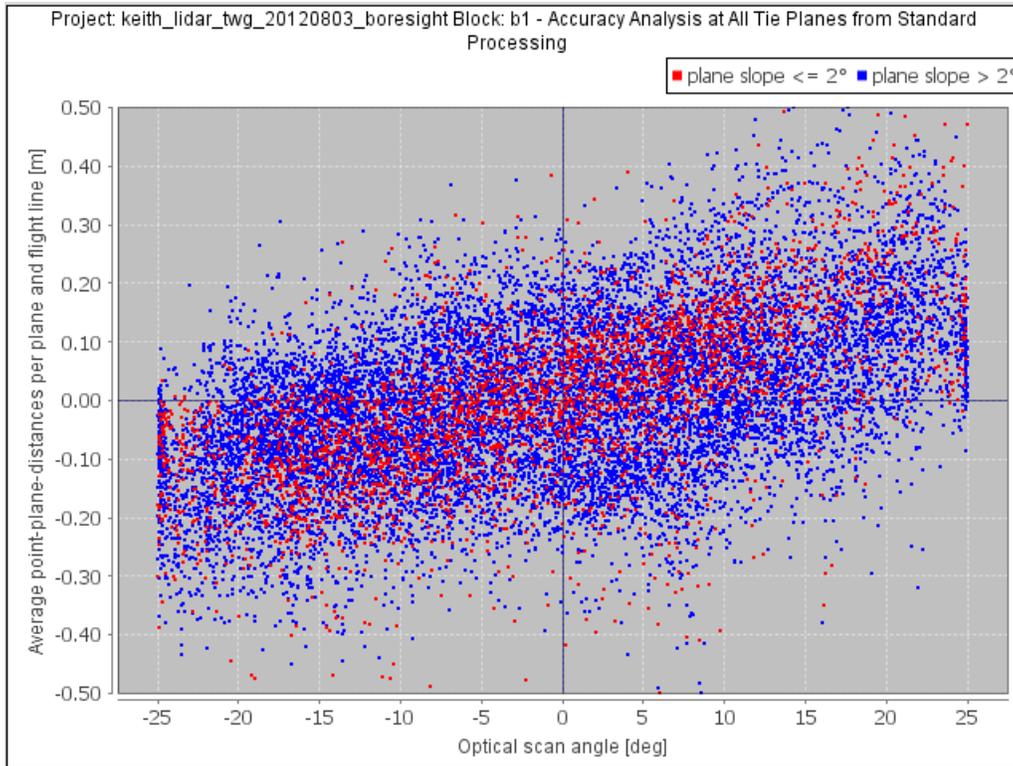
Figure 23. Difference in tie planes versus scan angle before boresight alignment calibration.
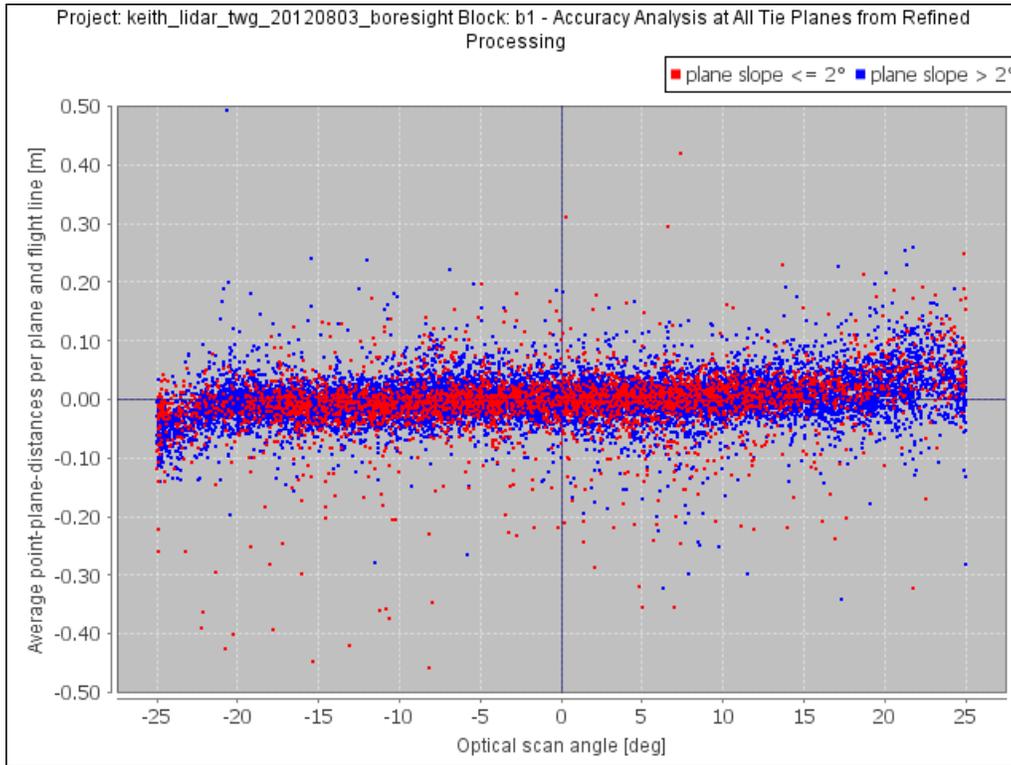
Figure 24. Difference in tie planes versus scan angle after boresight alignment calibration.

The LMS software reports the results of the boresight alignment calibration as updates to the LCP values of scan angle scale, boresight angle Ex, Ey and Ez as shown in **Figure 24**:

Figure 25. Example of value updates from a boresight alignment calibration.

Operationally the self-calibration process is repeated 10 times, updating the LCP values each time, and then the average value (of each parameter) is calculated using the last seven processing runs. This gives LMS a chance to converge to the best answer and averages out any noise in the calculation from a single calibration. The results of a series of 10 boresight alignment calibrations are shown in **Figure 25**:
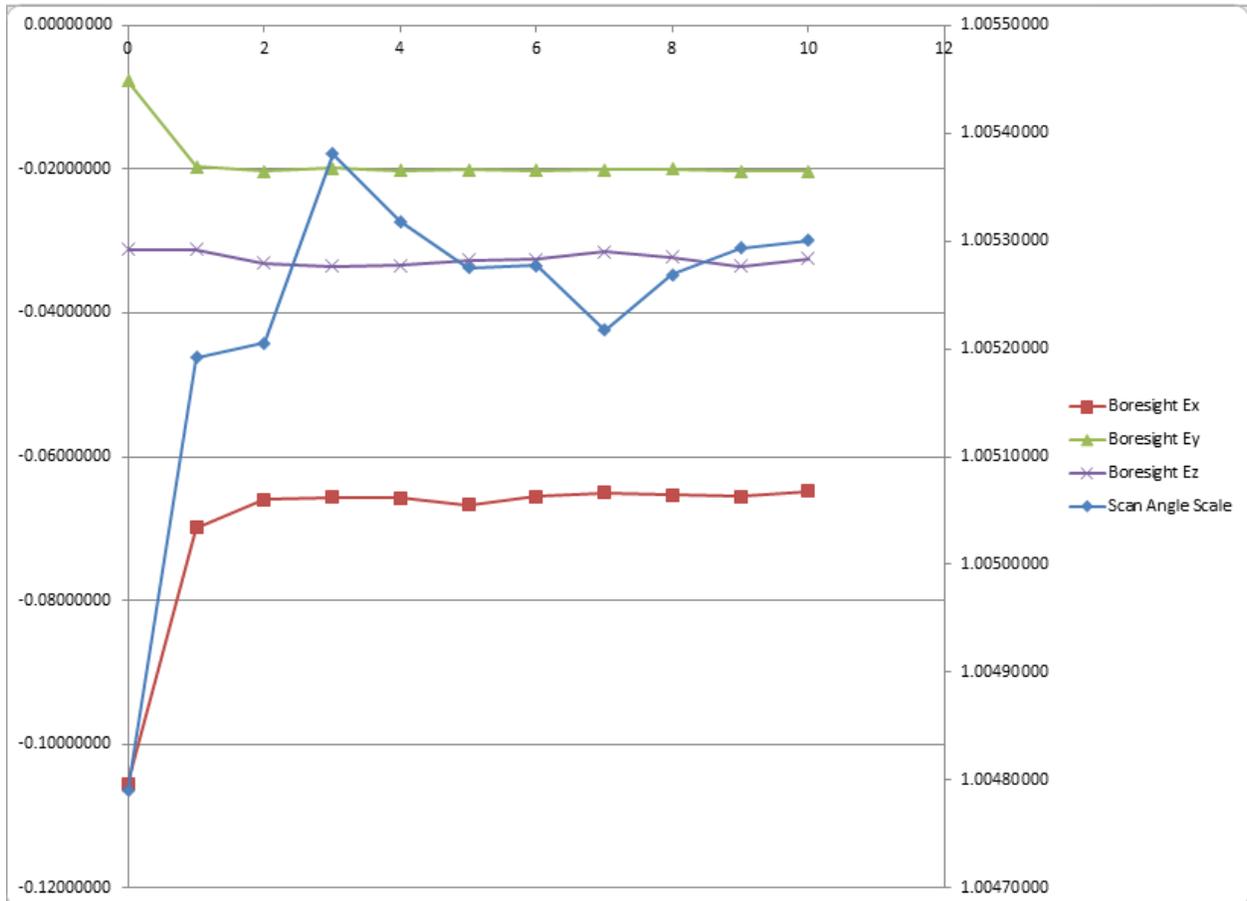
Figure 26. Boresight alignment calibration LCP values for a series of 10 consecutive self-calibrations with parameter updates for each iteration.

The boresight alignment calibration described here is performed for each aircraft installation and a single final LCP file is generated as shown in **Figure 26**:

```xml
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<boost_serialization xmlns="http://lcp" version="3" signature="serialization::archive">
    <calibration_parameters>
        <type>altm</type>
        <lasers>
            <count>0</count>
        </lasers>
        <oscillating-scanners>
            <count>1</count>
            <item>
                <id>1</id>
                <scan-offset>0.0</scan-offset>
                <scan-scale>1.0047785</scan-scale>
            </item>
        </oscillating-scanners>
        <cylindric-scanners>
            <count>0</count>
        </cylindric-scanners>
        <boresights>
            <count>1</count>
            <item>
                <id>1</id>
                <imu_ex>-0.10724933</imu_ex>
                <imu_ey>-0.00748166</imu_ey>
                <imu_ez>-0.03544567</imu_ez>
                <pos_dx>0.0</pos_dx>
                <pos_dy>0.0767</pos_dy>
                <pos_dz>-0.0279</pos_dz>
            </item>
        </boresights>
    </calibration_parameters>
</boost_serialization>
```

Figure 27. LCP file with the final boresight alignment calibration values.