Aalto University
School of Engineering

Arttu Järvinen

Airborne laser scanning data comparison based on roof features

Thesis submitted for examination for degree of Master of Science in Technology.

Espoo 31.12.2019
Supervisor: Prof. Matti Vaaja, Aalto University
Advisors: Prof. Harri Kaartinen & Prof. Antero Kukko, FGI

Aalto University
School of Engineering

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Arttu Järvinen

Title of thesis Airborne laser scanning data comparison based on roof features

Master programme Geoinformatics **Code** ENG22

Thesis supervisor Prof. Matti Vaaja

Thesis advisor(s) Prof. Harri Kaartinen & Prof. Antero Kukko

Date 31.12.2019 **Number of pages** 65 + 8 **Language** English

Abstract

Laser scanning is nowadays one of the most important technology in geospatial data collection. The technique has developed together with the other technologies and sciences, and the systems can be used with many different platforms on land, in the ocean and in the air. Airborne laser scanning (ALS) started right after the invention of the laser in 1960's and the usage grew in 1990's, when the first commercial system was released. The development has augmented the ways of surveying and the systems have new features and more options to collect as accurate data as possible. Several wavelengths and higher frequencies able thousands or even millions of measurements per second. The multispectral systems enable the characterization of the targets from the spectral information which helps for example in the data classification. Single photon technique provides higher imaging capability with lower costs and is used in the extensive topographic measurements. The processing of the point clouds are more important when the densities grow and the amount of noise points is higher. The processing usually includes preprocessing, data management, classification, segmentation and modeling to enable the analyzing of the data.

The goal of the thesis is to compare and analyze the datasets of five different airborne laser scanners. The conventional LiDAR datasets are collected from low altitude helicopter with the Riegl's VUX-1HA and miniVUX-1UAV systems. The state-of-the-art sensors, Titan multispectral LiDAR (Teledyne Optech) and SPL100 single photon LiDAR (Leica), are used in the data collection from the aircraft. The data is collected from the urban area of Espoonlahti, Finland, and the comparison is based on the roof features. Other land cover classes are left out from the investigation. From the roof features are investigated the differences, accuracies and qualities between the datasets. The urban environment was selected because the lack of ALS research done for the built environment, especially in Finland. The thesis introduces the background of the airborne laser scanning, theories and literature review, materials and methods used in the project.

The laser scanners used in the work produce dense point clouds, where the most dense is up to 80 pts/m². Based on the results the accuracies vary mainly between 0 and 10 cm. The scanners with infrared wavelengths produce better than 10 cm accuracies for the outlines of the roofs, unlike the green wavelength scanners. The differences in the corner coordinates are between 1 and 8 cm with a few exceptions. SPL100 system has the best height accuracy of 4.2 cm and otherwise the accuracies vary between 5 and 10 cm. The largest deviation compared to the roof planes occurs in the miniVUX-1UAV data (over 5 cm). For the surface areas the infrared frequencies produce differences of 0 to 2 percent from the reference data, whereas the differences of the green wavelength are mainly 1 to 7 percent. For the inclinations no significant differences were observed.

Keywords ALS, multispectral, single photon, LiDAR, 3D, point cloud, data comparison

Aalto-yliopisto
Insinööritieteiden
korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Diplomityön tiivistelmä

Tekijä Arttu Järvinen

Työn nimi Ilmalaserkeilausaineistojen vertailu perustuen kattojen ominaisuuksiin

Maisteriohjelma Geoinformatics **Koodi** ENG22

Työn valvoja Prof. Matti Vaaja

Työn ohjaaja(t) Prof. Harri Kaartinen & Prof. Antero Kukko

Päivämäärä 31.12.2019 **Sivumäärä** 65 + 8 **Kieli** Englanti

Tiivistelmä

Laserkeilaus on nykyään yksi tärkeimmistä tekniikoista geospatiaalisen tiedon keräämisessä. Tekniikka on kehittynyt yhdessä muiden teknologioiden ja tieteiden kanssa, ja järjestelmiä voidaan käyttää monilla eri alustoilla maassa, meressä ja ilmassa. Ilmalaserkeilaus (ALS) alkoi heti laserin keksimisen jälkeen 1960-luvulla ja käyttö kasvoi 1990-luvulla ensimmäisen kaupallisen järjestelmän julkaisun jälkeen. Kehitys on lisännyt mittaustapoja ja järjestelmien ominaisuuksien parantuessa on enemmän vaihtoehtoja kerätä tarkkaa aineistoa. Useilla aallonpituuksilla ja korkeammilla taajuuksilla pystytään tekemään tuhansia tai jopa miljoonia mittauksia sekunnissa. Monispektriset järjestelmät mahdollistavat kohteiden tunnistamisen spektritietojen (aallonpituuksien jakauman) mukaan, jota voidaan hyödyntää esimerkiksi aineistojen luokittelussa. Yksifotoni–tekniikka mahdollistaa suuremman mittauskyvyn pienemmällä kustannuksella (energiankulutus) ja sitä käytetään laajojen alueiden mittauksissa. Pistepilvien käsittely on entistä tärkeämpää kun tiheydet kasvavat ja virhepisteiden määrä on suurempi. Prosessointiin kuuluu yleensä esikäsittely, tiedonhallinta, luokittelu, segmentointi ja mallinnus, ennen aineiston analysointia.

Tämän opinnäytetyön tavoitteena on vertailla ja analysoida viiden eri ilmalaserkeilaimen tuottamia aineistoja. Ns. tavanomaiset LiDAR–aineistot on kerätty matalalla lentävästä helikopterista Rieglin VUX-1HA ja miniVUX-1UAV –keilaimilla. Viimeisintä tekniikkaa edustavat Titan monispektri LiDAR (Teledyne Optech) ja SPL100 single photon LiDAR (Leica) -aineistot on kerätty lentokoneesta. Aineistot on kerätty Espoonlahden alueelta ja vertailu perustuu kattojen ominaisuuksiin. Muut maanpinnan kohteet jätetään tarkastelun ulkopuolelle. Pistepilvien perusteella tutkitaan aineistojen välisiä eroja, tarkkuuksia ja muita ominaisuuksia. Kaupunkiympäristö valittiin kohteeksi vähäisen rakennetun ympäristön ALS–tutkimuksen takia etenkin Suomessa. Opinnäytetyössä esitellään ilmalaserkeilauksen taustaa, teoriaa ja tehdään kirjallisuuskatsaus aiheeseen liittyen, sekä käydään läpi projektissa käytetyt aineistot ja menetelmät.

Työssä käytetyt keilaimet tuottavat tiheitä pistepilviä, joista tihein on jopa 80 pistettä/m². Tulosten perusteella tarkkuudet vaihtelevat pääosin 0 – 10 cm välillä. Kattolinjojen kohdalla infrapuna-aallonpituutta käyttävät keilaimet pääsevät alle 10 cm, toisin kuin vihreän aallonpituuden keilaimet. Kattojen kulmakoordinaattien erot ovat 1 – 8 cm välillä muutamaa poikkeusta lukuun ottamatta. Korkeuksissa paras tarkkuus on SPL100 laserkeilaimella 4.2 cm, ja muuten ollaan 5 – 10 cm tarkkuuksissa. Suurimmat hajaumat tasoon verrattaessa syntyy miniVUX-1UAV aineistoon (yli 5 cm). Pinta-aloissa infrapunataajuudet tuottavat 0 – 2 prosentin eroja vertailuaineistoon, kun taas vihreällä aallonpituudella erot ovat pääosin 1 – 7 prosenttia. Kaltevuuskulmissa ei havaittu merkittäviä eroja.

Avainsanat ilmalaserkeilaus, multispektri, yksifotoni, LiDAR, 3D, pistepilvi, aineistojen vertailu

Foreword

Strategic Research Council at the Academy of Finland project "Competence Based Growth Through Integrated Disruptive Technologies of 3D Digitalization, Robotics, Geospatial Information and Image Processing/Computing - Point Cloud Ecosystem (293389 / 314312)" and Academy of Finland project ""Multi-spectral personal laser scanning for automated environment characterization (300066)" are acknowledged for financial support.

I would like to thank my supervisor Matti Vaaja, who guided me during the thesis and gave me a hint to ask from the FGI (Finnish Geospatial Research Institute) about the possibility to work on laser scanning in the research environment. I would like to thank my advisors Harri Kaartinen and Antero Kukko, who set the goal, gave the feedback and advised me during the project. Thanks to Juha Hyyppä for the possibility to write this thesis, and also to the people in the FGI and NLS (Finnish National Land Surveying).

A special thanks to the boys with whom I managed to complete the courses during the studies and with whom we had a great time together: Niko, Joonas, Jussi, Jani and Mathew.

Finally, thanks to my family and friends who have been curious about the studies, who have always supported me, and who have been great examples in my life!

Espoo 31.12.2019

Arttu Järvinen

Table of Contents

Symbols

c	Speed of light
e_1	Observed value
e_2	Reference value
F	Size of the covered area
h	Survey height
n	Number of samples (data points)
Pm	Mean point density
PRF	Pulse repetition frequency
R	Range
RMSE	Root mean squared error
sw	Swath width
Std	Sample standard deviation
t	Certain time interval
t^f	Time of flight
v	Flight speed
x_1	X –coordinate of a point 1
x_2	X –coordinate of a point 2
xi	Measured value
ẍ	Mean value
y_1	Y –coordinate of a point 1
y_2	Y –coordinate of a point 2
θ	Scan angle
λ	Wavelength
ϕ	Phase resolution
π	Pi

Abbreviations

ALS	Aerial Laser Scanning
BC	Before Christ
CAD	Computer-aided Design
CW	Continuous Wave
DSM	Digital Surface Model
DTM	Digital Terrain Model
FGI	Finnish Geospatial Research Institute
FoV	Field of View
GIS	Geographical Information System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
IMU	Inertial Measurement Unit
LiDAR	Light Detection and Ranging
MLS	Mobile Laser Scanning
MMS	Mobile Mapping System
NASA	National Aeronautics and Space Administration
NIR	Near-infrared
NOAA	National Oceanic and Atmospheric Administration
POS	Position and Orientation System
PRF	Pulse Repetition Frequency
PRR	Pulse Repetition Rate
PS	Phase Shift
RMS(E)	Root Mean Squared (Error)
RTK	Real Time Kinematic
SWIR	Short-wave Infrared
TLS	Terrestrial Laser Scanning
ToF	Time of Flight
UAV	Unmanned Aerial Vehicle

m	Meter
m^2	Square meter
cm	Centimeter
mm	Millimeter
µm	Micrometer
nm	Nanometer
ns	Nanosecond
ppm	Parts per million

1 Introduction

1.1 Background

Land surveying and its techniques have developed during the centuries together with the other technologies and sciences. The historic records say that the surveying goes back in time of the Egyptians in 2700 BC and 1400 BC, when the pyramids were built and lands were plotted for the purpose of taxation. Babylon, Greece and Rome are also mentioned as places where the surveying have been evolved [1]. Two parts of the surveying process are the documentation and analyzing of the data, which are strongly related to the cartography. Cartography is called as the art and science of map making. The first artifacts were drawings into the rocks or wooden plates, and they didn't based on measurements but imagined evocation [2, 3].

If we move back to the present time, the development is obvious. There are new highly equipped devices, advanced algorithms and knowledge thanks for the history and technological development. Photogrammetric, terrestrial, aerial and satellite systems can confidently reach centimeter accuracy and soon even millimeters may be standard. Map making and analyzing of the datasets are based on accurate measurements with big amount of data, which requires some experience with GIS software [4]. Also the planning and analyzing have changed from 2D to 3D world.

At the moment one of the most important technology in geospatial data acquisition and in cartography is laser scanning. In modern surveying the meaning of the laser scanning corresponds to LiDAR (Light Detection and Ranging) because of the combination of the laser light and 3D scanning. The invention of the laser happened in 1960's and after that it has been used in numerous applications [5]. Because of the lack of the proper technologies in 1960's, it took some time before the laser systems started to be in universal use in topographic applications. In the field of surveying, after the period of development and research, the laser systems became to commercial use in the mid 1990's. NASA is mentioned as one of the major pioneer in the development of these systems through their actions in the mapping of the Arctic areas [5]. Also the NOAA and the University of Stuttgart played the key role in the pioneer work of the development [6].

Laser scanning is typically divided into three categories: terrestrial laser scanning (TLS), airborne laser scanning (ALS) and mobile laser scanning (MLS). Terrestrial lasers were used in the mid 1960's onwards to measure distances and ranges. Around 1970's lasers started to replace other instruments in distance measuring, even without possibility to measure angles. Angles were measured with other devices, for example with theodolites [5]. Nowadays so called total stations are combination of the distance measuring laser instrument and the angle measuring instrument which were merged together after 1970's. Later, these devices went through a lot of improvements, for example in the features of lasers, quality of distance measurement and automation, which lead to the development of the present terrestrial laser scanners [5]. First generation of terrestrial laser scanners was in the mid 1990's, when all systems were pulse based, i.e. the distance is determined by the by the laser pulse's time-of-flight between the sensor and the target.

One reason for the early progress of ALS was the curiosity to test the lasers from the airborne platform. Different kind of aerial lasers were tested and developed between 1960's and 1990's. The first airborne platform laser altimeters were flown in 1965 to measure continuous profiles of the terrain, but the features of the applications were very limited. For the military use were the GPS and IMU attached to the laser scanner in 1980's, but after that, in the early 1990's the number of aerial platforms grew. The first commercial ALS system (with GPS and IMU) were available in the mid 1990's [5]. The first UAV platform ALS system came in 2010 [7] and the first multispectral airborne system in 2014 [8]. The first commercial single photon ALS system was released in 2017 [9]. Nowadays the sensor devices can be attached, not only to steady terrestrial or flying airborne platforms, but also to mobile mapping systems (MMS) for example to backpacks, cars or boat-based platforms [10].

1.2 Objective and structure

The previous part of the work was kind of a background for this thesis. The technologies and ways of surveying are evolving all the time but maybe the quickest improvements in the next few years will happen in laser scanning. New platforms and possibilities to scan and profile on the ground, in the air and in the water are even more developed. The systems include more options for example in the number of beams or wavelengths, or they co-operate together with other sensors. In this work is used, among other things, multi-wavelength and single photon data, together by now traditional one channel laser scanned data.

The objective of the thesis is to investigate differences of the point clouds and 3D models of the airborne laser scanned data collected with four sensors (Riegl VUX-1HA, Riegl miniVUX-1UAV, Leica SPL100 and Teledyne Optech Titan). The comparison is performed by using the building roof features. The idea is to examine differences between the point clouds and to investigate and analyze the quality and accuracy of the datasets. The test site is located in Espoonlahti, an urban area in Espoo, Finland, which is one of the permanent test sites of the Finnish Geospatial Research Institute (FGI). The lack of airborne laser scanning research done for built environment, which includes also multispectral and single photon data, was one of the key reasons for selecting the topic of the thesis. Especially in Finland are done lot of research for vegetation because of the interest and optimal circumstances. In this work, the comparison includes

- point densities,
- surface areas,
- outlines (lengths and widths),
- planimetric (x and y –coordinates) differences of the corner points,
- heights,
- inclinations (angles),
- deviations of the points.

The work is limited not only with the boundaries of the test site or the number of the selected buildings, but also in a way that the research is based on the differences concerning the roofs of the buildings. Vegetation, roads, walls etc. are left out because is wanted to focus specially on the roofs.

This chapter described some background and set the objective of the thesis. Chapter two is the literature review, which describes the theory behind the airborne systems and the laser scanning, describes some mathematical concepts and previous works related close to the objective of the thesis. Chapter three presents the materials and methods for the roof feature measurements and comparison. Chapter four shows the outcome of the investigation. Chapter five presents the overview of the research, discloses the possible error sources, and takes a look to the ideas how the work could be continued. The last chapter concludes the work.

2 Theory and literature review

2.1 Airborne laser scanning

Section 2.1.1 clears the idea of the equipment and principles in the airborne laser scanning. The basic devices are described and the purpose for the use is examined. In the section are also listed the general things that have influence to the point cloud determination. Manufacturers may have their own components or methods as well, but the idea is to go through the fundamentals of the ALS. Section 2.1.2 describes some methods of the calibration and registration. Section 2.1.3 explains what the multi-wavelength means in the case of the ALS and what the benefit is. Last section 2.1.4 presents the idea of single photon technique.

2.1.1 Principles of ALS

The airborne laser scanning can be done from the fixed wing aircraft, from the helicopter or with the UAV platform. The same components are used in all the platforms and the working mechanism is usually similar. On-board there are two main components which are the laser scanner and a POS (GNSS/IMU) system (Fig. 1). The GNSS (Global Navigation Satellite System) in the POS system receives the positioning and timing data to define the location. The satellites of the European Galileo, American GPS, Russian GLONASS and Chinese BeiDou systems provide the global coverage of signals for the GNSS receiver [11], but all the satellite systems can be used alone as well to receive position information. Laser scanner measures the distance of the points on the scene, and the GNSS together with the IMU (Inertial Measurement Unit) handles the information about the exact position and orientation of the system [12]. Together with the main components are working the computer and the data recording unit (Fig. 1).

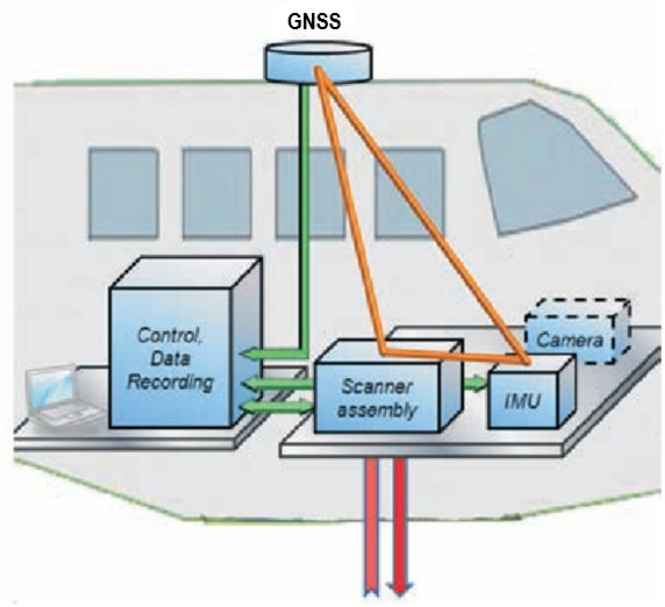
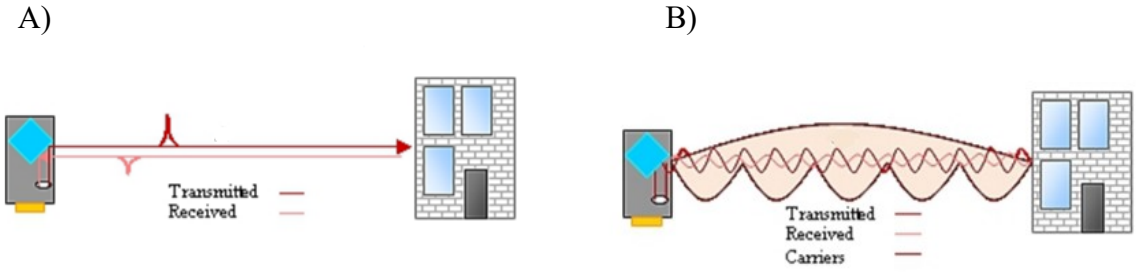


Figure 1. The main components of an airborne laser scanning system (modified from Figure 1.13 from the book "Airborne and Terrestrial Laser Scanning" by Vosselman G. and Maas H-G., page 22) [12].

- **Scanner** keeps the laser, optics and mechanics inside and works from the hole on the bottom of the aircraft. It can also be mounted outside of the aircraft. The system sends continuous pulses to the terrain when the aircraft is flying and calculates the exact coordinates for the returned hits. The range measurement systems usually work with the time-of-flight (ToF) or phase shift (PS) techniques (Fig. 2A and 2B). In ToF method the time of flight of the pulse and speed of light are defined and precise timing for the ranging is utilized. PS uses continuous laser illumination why it is also called as continuous wave (CW) method. It compares the phase shift of the emitted and backscattered laser beams for the range measurement [13, 14]. Laser scanning is so called active system and is able to operate without other light sources e.g. sunlight. This can be seen as an advantage compared for example to photogrammetry.



(1, 2)

$$R = \frac{c}{2} * t^f \qquad R = \frac{\lambda}{4\pi} * \phi$$

Figure 2. The principles and formulas of the A) ToF and B) PS laser ranging [13, 15].

where

R is the range between the scanner and the object
c is the speed of light
t^f is the time of flight
λ is the wavelength of the ranging signal
ϕ is the phase resolution

- **GNSS** antenna (e.g. GPS) is usually placed on top of the aircraft to the spot, where the signal to the satellites is most likely undisturbed. System collects the exact position data (x, y, z) of the aircraft, or more precisely of the scanner, using the satellite information. It reconstructs the flight path (trajectory) together with the IMU system, which is recorded to the database [12].
- **IMU** controls the orientation of the system and is located either near the scanner system or is fixed directly to the laser scanner. The component collects the rotation rates (roll, pitch, and yaw) to determine the orientation of the aircraft. It also collects the acceleration data to reconstruct the trajectory paths together with the GNSS system [12].

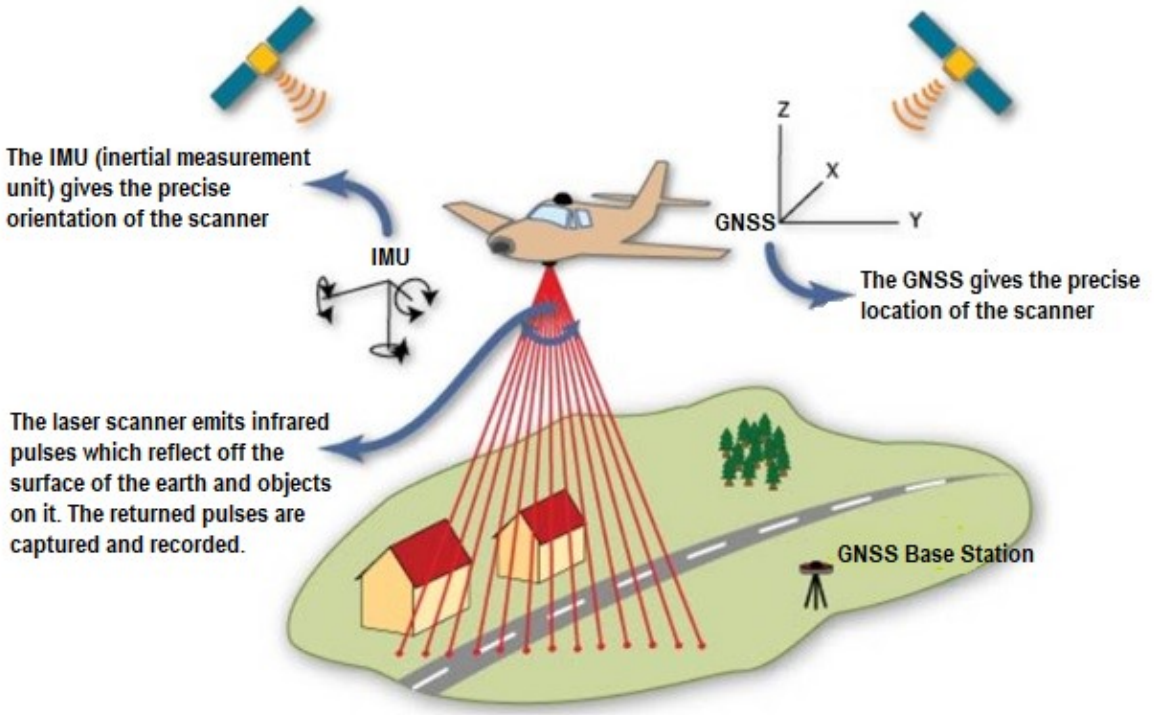


Figure 3. Visualized case of the airborne laser scanning. Location, altitude and the angle of the light ray need to be known, traveling time of the light ray is measured, and according to the observations, the 3D points are calculated (modified from figure of "ALS System Schematic" from the GMW website, 2019) [16].

- **Data recording unit** controls the whole system and stores the data from the scanner and POS for the calculations of the 3D points. Unit is responsible of the time synchronization of the sensors, which have to be precise to allow the good quality data. Computer is part of the communication between the control unit and the user. It also calculates the real time performance of the survey [12].
- Possible **camera system** will allow the correctly colored point cloud, which may help the object recognition. If the camera is located properly, the IMU data can be directly used in the registration of the image data to avoid extra work [12].

The densities of the surveyed point clouds vary between flight paths and equipment. The target area has also an influence to the point density, because the reflectance differ between measured objects. The LiDAR systems allow typically density of 5 – 20 pts/m², but the new scanning methods raise the point density even to 50 – 200 pts/m². The density and quality of the point cloud affect for example to the classification and object recognition of the terrain. Different factors have influence to the point density, for example:

- flying speed
- survey height

- scanner properties (e.g. range measurement method, laser wavelength, swath width, pulse repetition frequency, scanning frequency)
- scan angle and surface angle
- overlap of flying strips
- surface material/reflection
- lighting and weather conditions [12, 16].

Pulse repetition frequency (PRF), also called as Pulse repetition rate (PRR), is an important factor in the LiDAR system. The velocity of the aircraft or a helicopter is usually too fast for the sensor, to wait only one laser beam to hit the target and receive the backscattered beam, to collect the encompassing amount of data. Because of the velocity and altitude, the scanner operates with several laser beams at the same time. PRF is the definition for the leaving number of laser pulses sent per second. The unit of PRF is Hertz (Hz), where for example 100 kHz means ~ 100 000 emitted pulses per second. However, effective measurement rate depends on the scanning mechanism and optics, and signal processing, e.g. multiple echo detection. Generously, it can be said that the higher PRF provides the higher point density [12, 17]. The type and wanted end product of the project affect to the planning of the measurement campaign. The flight plan and device are related to the point density, which is possible to estimate in advance. Rough approximation before the survey and the true point density after the survey can be calculated with the equations below.

To calculate the mean point density must be first calculated the size of the covered area F [12]:

$$F = v * t * sw \tag{3}$$

where

v = flying speed
t = time interval
sw = swath width

And then the mean point density Pm [12]:

$$Pm = \frac{PRF * t}{F} = \frac{PRF}{v * sw} = \frac{PRF}{2vh \tan(\theta/2)} \tag{4}$$

where

PRF = number of laser pulses generated during a time interval
h = height
θ = scan angle

When using the scan mechanism, there is an interrelation between scan frequency and field of view. Different types of systems use different types of deflecting mirrors, which form various types of point patterns on the ground (Fig. 4). Scan frequency (or scan rate) is the

definition for the number of scan lines per second. The unit is Hertz as in the pulse frequency, but they implicate different things, also LPS (lines per second) is sometimes used as a unit. Scan rate relates into line spacing on the ground. If the rate is high the system produces more scan lines per second and the distance between individual lines is smaller. This has influence also to the point density and affects to the mapping of the objects, for example how well and accurately the corners of the roofs are captured. Field of view (FoV) describes the angle that the scanner can measure. The orientation and properties of the scanner, the flight altitude and speed determine the swath width on the ground, and these factors together with the scan speed define the point spacing within a scan line. With the point spacing is meant the distance between two points in the scan line. FoV alongside scan frequency has influence to the accuracy of the data [18].

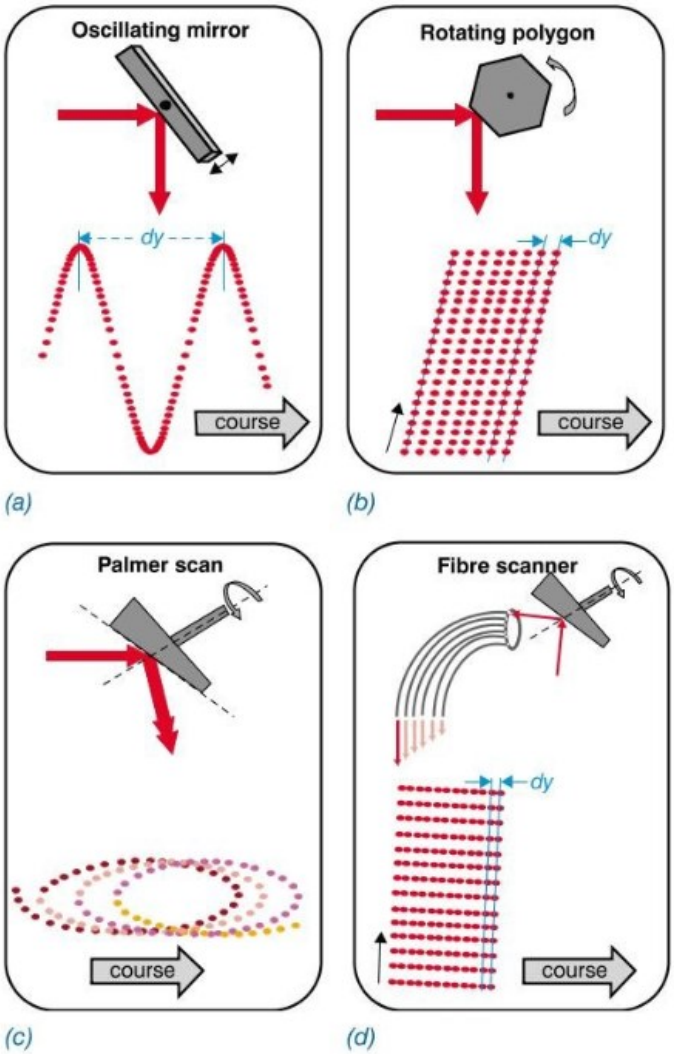


Figure 4. Mirror system has influence to the point pattern [12].

2.1.2 Calibration and registration

The quality of the point cloud depends on the registration and calibration, which are highly related to the observations of the scanner, GNSS and IMU systems. Calibration is done to avoid errors between LiDAR data and the ground truth, and also to improve the reliability of the result. It must be done carefully for each individual sensor, but also between POS system and scanner, to perform with maximum accuracy. The spatial relationship between the sensors is called the mounting, bias or boresight. The boresight misalignment originates from the local reference systems of the components, and is defined by the offset and rotation between the systems. The LiDAR points are obtained to the IMU coordinate system by applying the shift and rotation between the systems. The three rotation angles are around x-, y- and z-axes because of the roll, pitch and yaw changes of the airplane. The common methods in overall system calibration process are block adjustment and a simple trial and error approach, where the angles are changed to fit the LiDAR points to the known surface. These methods are human-based techniques, where the known surface or points are required. Automated methods, which don't need a priori knowledge are also tested. One example uses two or more overlapping strips, and the surface differences are considered as observations. It can be found from the paper by Toth [19].

Gathering LiDAR data can be called as multi-platform, multi-angular, and multi-temporal task. These terms are valid at least in the case, when the different surveying methods (ALS, TLS, MLS) are used together, but stands also when only one method is used. The point cloud registration means merging several point clouds or setting the cloud to the selected coordinate frame. It can be seen as a "same-platform registration" or "registration between different platforms". In the ALS campaign the platform is usually the same aircraft from the beginning to the end of the survey, but the angular and temporal changes are still happening during the work. The scanning height and FoV limit the ground width coverage, which results to the need of several flight lines. This together with the integrated system (more than one component) expose the potential errors. On the other hand, large enough overlapping is a way to recognize and fix the error. Several potential error sources may occur on the flight paths, in ranging, sensor mounting, POS or orientation. Registration methods are either data-driven or sensor system-driven methods, where flight lines are adjusted and errors avoided. The difference adjustment is one of the registration methods, where the differences between point clouds are simply minimized. Sensor system-driven method uses positioning equation as an adjustment model, which is seen more generally than data-driven method. Data-driven method uses geometric features to calculate the rotation matrix and translation vector [20].

In addition it must be mentioned the coarse-to-fine registration strategy which is generally used for the same-platform registrations. To achieve good initial values, first coarse method is used and later the fine method to achieve best possible match between two point clouds. More about the registration and calibration of the laser scanning data can be read from papers [13, 14, 19, 20, 21], and the basic relations and formulas of the airborne laser scanning can be found from paper [22].

2.1.3 Multi-wavelength laser scanning

Laser is the light amplification of stimulated emission of electromagnetic radiation, which consist of wide range of wavelengths and frequencies. Wavelengths and frequencies form the electromagnetic spectrum for the light source, which is divided into regions (Fig. 5). The laser together with the other light sources, e.g. sunlight, sends the pulse, waveform light, where the lengths of the wave peaks can be measured. The amount of energy in a photon and electromagnetic radiation are connected. For example short wavelengths have higher energy than longer waves [23]. Usually LiDAR systems use wavelengths in the range of 0.4 – 1.6 µm (= 400 – 1600 nm). The regions of electromagnetic spectrum are a bit inconstant, but are generally determined as follow (the irrelevant wavelengths of the radiation in ALS are left out: Gamma, X-ray and Radio):

- Ultraviolet 10 – 400 nm
- Visible 400 – 700 nm
- Infrared 700 nm – 1 mm
- Microwave 1 mm – 1 m

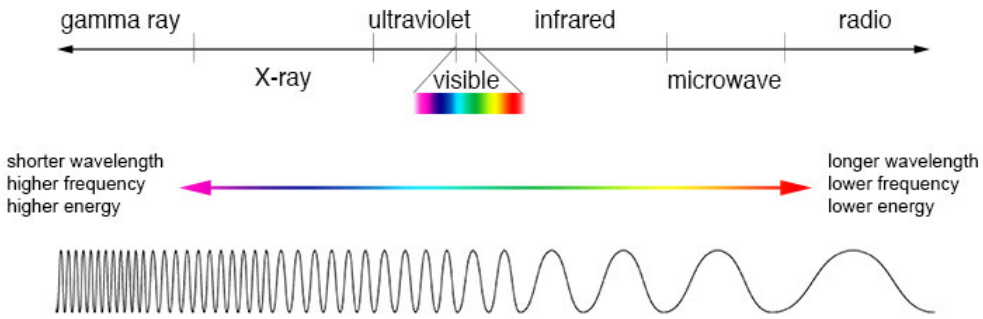


Figure 5. Electromagnetic spectrum [24].

The first multispectral airborne LiDAR, called Titan from Teledyne Optech was introduced in 2014 (Titan will be presented later in this thesis) [8]. As can be inferred, multispectral LiDAR uses two or more channels with different wavelengths. The possibility to acquire a more dense point cloud is not the only benefit of using multiple wavelengths, but also the possibility to interpret the objects from the spectral information of the independent channels (Fig. 6). In many cases the spectral information will simplify the processing of the data because of the spectral characteristics of each point. For example, from the spectrum can be inferred the materials of the roofs, the species of the trees, or if the measured point is reflected from the vegetation or soil. The usage of different wavelengths is emphasized in the water areas and bathymetric surveying, because of the penetration of the water. Near infrared channel (~ 700 – 1400 nm) is mostly absorbed by water, unlike the green channel (~ 495 – 570 nm) in visible region. The total number of wavelengths is big, but only a few can be used in the ALS, which limit the options a little. Wavelengths 355 nm, 532 nm, 905 nm, 1064 nm, 1550 nm and 2050 nm are used in the long-range systems. The reason is that those wavelengths only have lasers, which are at the same time powerful and safe enough to be used in the ALS systems [25].

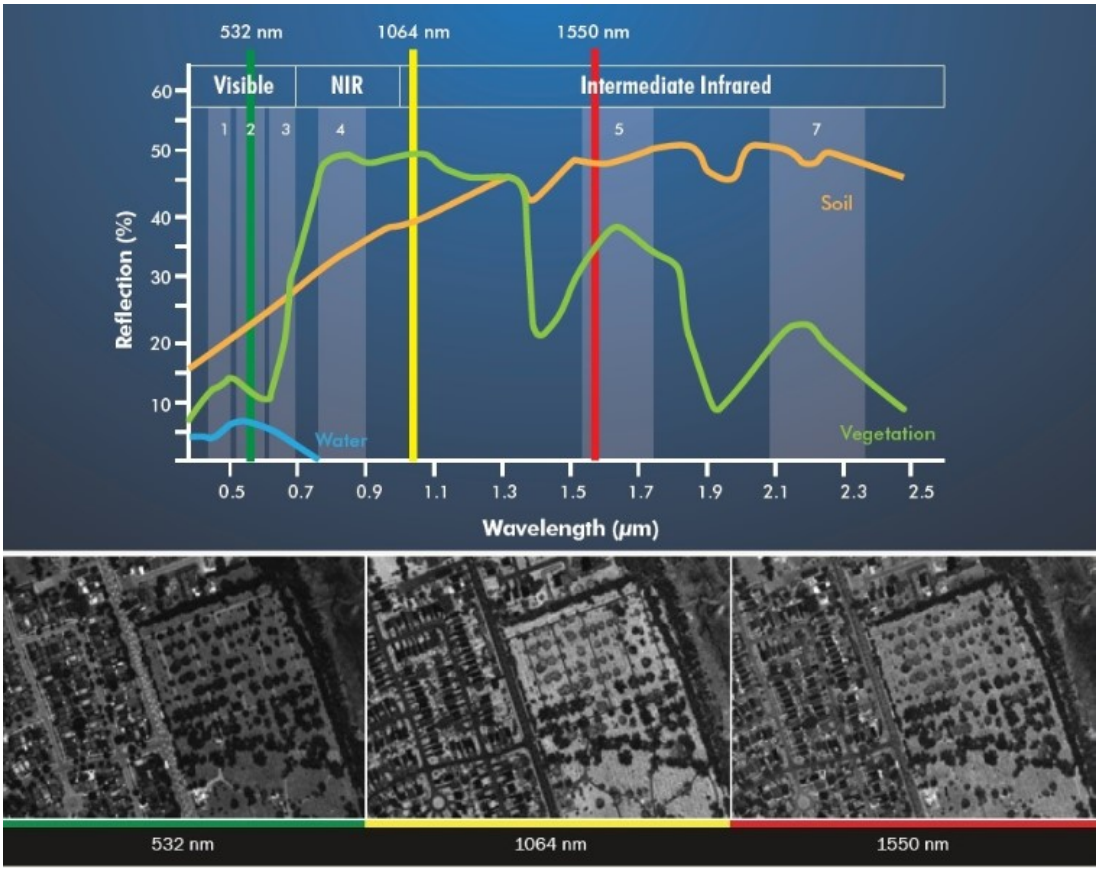


Figure 6. Example of the characteristic changes of soil, vegetation and water with different wavelengths. Titan multispectral LiDAR [8].

2.1.4 Single Photon LiDAR

Conventional airborne LiDAR systems, also called as linear mode LiDAR, can capture up to 1 – 2 million points per second. As was already mentioned, to be able to get the high point density the pulse repetition rate must be increased, which increases also the number of the photons (usually the echo can't be detected with the small number of photons). The higher PRF allows higher flying speed and altitude, and the dense point cloud can be measured at lower costs. The challenge is the limitations of the PRF because of the energy consumption, heating of the device and eye safety. The solution was to change the technology of the system to reduce the required energy of each pulse [17, 26]. Sigma Space Corporation developed the Single Photon-sensitive LiDAR (SPL), which achieves high pulse rates with lower power consumption. The first publication about Sigma's LiDAR data came in 2015 [27] and the first commercial sensor by Leica (will be presented later) in 2017.

SPL system transmits shorter pulses with lower energy because it divides the light beam into hundred beamlets, arranged in a 10 x 10 array. The individual beamlets are imaged onto a pixel in a matching array detector of the telescope. The sensitivity of the SPL system is based on the recording the ToF of each photon of the laser pulse. It requires only one photon to be

detected for the range measurement, and the recovery time of the channel (1.6 ns) is much shorter than in the other systems (Fig. 7). This increases the surface measurement rate notably and the point density is higher. The increased point density per each laser pulse allows the lower PRF, and the lower PRF allows more imaging capability. The higher flying altitude and spatial resolution with a wider swath enable the greater areal coverage. The available systems from Sigma Space Corporation and Leica are efficient in extensive topographic measurements, but map also well the bathymetric areas because of the visible green wavelength (532 nm) [28, 29, 30, 31]. However, the sensitivity of the system increases also the number of the noise points. In the article by Degnan [29], is written that the possible noise of the data originates from the dark counts from the detector, from the solar backscatter or from the laser backscatter from the atmosphere. The atmospheric backscatter has to be taken into account at least when SPL is operated from the space.

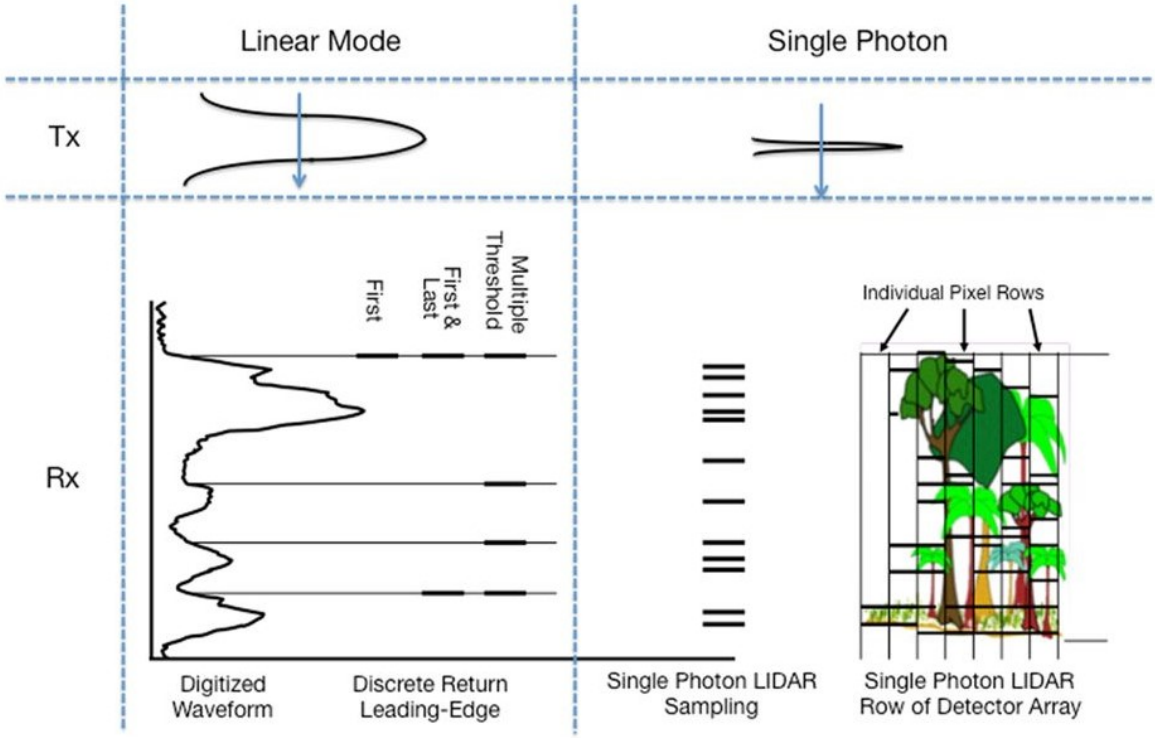


Figure 7. The schematic idea of the common ALS systems and the SPL system. SPL system has a shorter and lower energy pulse (Tx) and it detects the returned energy more efficiently (Rx) [32].

2.2 Computational concepts

Standard deviation is used in statistical analysis to describe the dispersion of the values [10, 28, 31]. The dispersion of the dataset is relative to its mean, where the square root variance between the mean and each data point is calculated. The low standard deviation means that the data points are close the mean and less spread. When only sample from the larger group of observations is used, the sample standard deviation is applied instead of population standard deviation (estimates the deviation of the entire population or observations). The expression of the sample standard deviation is [33]:

$$Std = \sqrt{\frac{\sum(xi - \ddot{x})^2}{n - 1}} \tag{5}$$

where

xi = measured value
ẍ = mean of the data
n = number of data points.

The accuracies of the individual data points can be described by the distances between the measured points and the reference points. The distance between the points can be thought as a line, and to find the line, at least two x and y –coordinates are needed to be known. Z-coordinate can be included if it is wanted to find the 3D distance. In this work are investigated the planimetric differences of the corner points, and the distance formula is a simple geometric calculation [34]:

$$Dist = \sqrt{((x_2 - x_1)^2 + (y_2 - y_1)^2)} \tag{6}$$

where

x_1 = x –coordinate of the point 1
x_2 = x –coordinate of the point 2
y_1 = y –coordinate of the point 1
y_2 = y –coordinate of the point 2.

Root mean squared error (RMSE or RMS) is commonly used for the performance analysis of the point cloud data [13, 32, 35, 36, 37]. RMSE presents the differences between the reference and the measured values. The statistics about the distribution of the data can be calculated by comparing the individual points and the correspondences of the reference. The result of the RMS performs the fitting, in other words the accuracy, of the measured data. The expression of the equation is shown below [13]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(e_1 - e_2)^2}{n}} \quad (7)$$

where

e_1 = measured value
e_2 = reference value
n = number of samples.

2.3 ALS studies for urban areas

2.3.1 Conventional LiDAR

Previous chapters were based on the articles and writings, and introduced the idea and methods about the data capture with the airborne laser scanning. As this thesis concentrates more on the data comparing and analyzing, this chapter discloses some interesting works that have been done with the ALS. While the most of the ALS related works focus on the classification and segmentation of the land cover data, here LiDAR works that include urban areas and building exploration are collected. In many research studies also hybrid systems, where LiDAR data has been combined together with the aerial images, are used. Hybrid method for example color the point cloud and help the object recognition. It is discovered that commonly the hybrid technique improves the accuracy a little, because of the good planimetric accuracy of the digital images, and the good height assessment accuracy of the LiDAR. In 2005, the study by Kaartinen et.al. [35, 36], the participants were investigating the accuracies of the outlines, lengths, heights and inclinations of the 3D building models with the different methods by using laser scanning data, image data or the combination of the datasets. The building outline errors ranged from 20 to 76 cm (mean 44 cm, std. 18.5 cm) with the hybrid methods, and from 20 to 150 cm (mean 66 cm, std. 33.2 cm) with the laser scanning based methods. The building lengths ranged from 19 to 108 cm (mean 59.4 cm, std. 31.2 cm) with the hybrid methods, and from 13 to 292 cm (mean 93 cm, std. 60.9 cm) with the laser scanning methods. The building heights ranged from 9 to 34 cm (mean 18 cm, std. 8.5 cm) with the hybrid methods and from 4 to 153 cm (mean 32 cm, std. 31.5 cm) with the laser scanning methods. The inclinations ranged from 0.6 to 2.3 degrees (mean 1.3°, std. 0.6°) with the hybrid methods and from 0.3 to 9 degrees (mean 2.7°, std. 4.4°) with the laser scanning methods. In 2018 Awrangjeb et.al. [38], also used hybrid data in the 3D roof reconstruction and reached couple of centimeter accuracies in the planimetric and height -coordinates.

Classification, segmentation and 3D modeling are typical stages in the data processing and analyzing. Points belonging to a building can be separated from the other data points with classification, and the redundant data is simpler to detect and cut out. The segmentation of the building structures and the modeling can be done without exhaustive and huge point clouds. The 3D information of the processed data should be analyzed and compared to the other datasets or other processing methods to be able to verify the accuracy. Comparison of the data are done for example with the other aerial based laser scanned or photogrammetric

data, measured field reference points or modifying the degree of automation in the 3D reconstruction [39]. Automated reconstruction processes have progressed significantly and more accurate and robust building models can be produced. Xiao et.al. [40] revealed the paper in 2014 where the segmentation methods provided 0.6 m planimetric and 0.1 m vertical accuracies. In 2016 Cao et.al. [41] provided roof reconstruction technique with 10 cm planimetric and vertical accuracy. In the segmentation of the roof and wall features data-driven (e.g. clustering) or model-driven (e.g. Hough transform or RANSAC) approaches are used. When the building segments are recognized the outlines can be created from the intersection of the planar patches [38, 40, 41, 42, 43, 44, 45, 46]. 3D building models are typically used in the visualization, urban planning and facility management. Model based applications enable even more in environmental studies, for example in the damage detection or solar resource assessment [47, 48]. The combinations of multiple detection and extraction methods in classification and segmentation are seen to improve the accuracies in vertical and horizontal directions, which are typically from few centimeters to one meter [42, 43, 44, 45].

2.3.2 Multispectral LiDAR

Multispectral ALS, and especially Titan from Teledyne Optech, is used in numerous researches related to the map updating, classification and change detection. In 2015 Bakula [25] used the independent channels to create the Digital Terrain Models (DTMs) and Digital Surface Models (DSMs). The vertical accuracies of the channels were investigated by comparing the models together. Between the channels, the mean values of the DTMs were less than 0.03 m with the standard deviations ranging from 0.19 to 0.27 m, and the mean values of the DSMs for the NIR and SWIR channels were -0.13 m and -0.10 m with respect to the visible band.

The three wavelength technology is seen suitable in the land cover classification, and in the building and tree detection. Also in road mapping, the multispectral ALS has achieved good results, which are competitive or even better than the results of aerial image data [49]. Researchers [49, 50] have noticed, that automated classification methods produce better results with multispectral ALS data than with multispectral photogrammetry data, because of the independency of ambient light and shadows. The existence of shadows or occlusions in the data may confuse the algorithm in the automated processes [50]. In the studies [51, 52, 53, 54], the accuracies of the classification are improved for example with the methods of object-based analysis, clustering, optimization, or with the maximum likelihood classifiers. In many papers, the number of land cover classes range from three to nine classes, where the commonly used classes are water, grass, trees, roads and buildings. For comparison, in the data classification for this project six classes are used: ground, hard surface, low vegetation, medium vegetation, high vegetation and buildings. The gained accuracies of the classification methods vary a bit but they show that the results from 90 to 98 % are possible to achieve [51, 52, 53, 54, 55].

In urban area the classification has more uncertainties because the objects are mixed. For example the vegetated roofs and the materials of the roads causes some unsteadiness in the classification. Some articles [37, 55] prove that promising results, even 90 % accuracies, in the urban areas can be achieved with the multispectral systems. One method is to use the spectral and spatial relationships together, to detect the land cover types. The capability of

Titan was extensively tested by Fernandez-Diaz et.al. (2016) [37], where the multispectral capabilities of the system in the classification and qualitative observations, bathymetric and penetration capabilities, range characterization, and what we are interested in, the precision and accuracy assessment, were tested. Fernandez-Diaz et.al., noticed that the height precision and accuracy decrease, when all the channels are combined and analyzed, despite handling one channel at a time (Table 1). The medium altitude data was collected in 2014 and the topographic elevation precision, from the altitude of 800 m with the PRF of 300 kHz, was about 5 cm and the accuracy about 6.5 cm. The performance was also tested from different ranges and with different frequencies, but no significant variation was obtained [37]. In the literature overview no papers, where the multispectral ALS data would have been used in the 3D building reconstruction and accuracy assessment of the models, were found.

PRF	Range	Number of Samples			Height RMSE (m)			
(kHz)	(m)	C1	C2	C3	C1	C2	C3	C123
				Precision				
100	900	658	-	667	0.02	-	0.018	0.030
200	900	1411	93	1542	0.018	0.028	0.018	0.020
300	800	2535	756	4766	0.019	0.026	0.017	0.048
100	500	1038	1061	1060	0.019	0.018	0.021	0.045
200	500	2208	2280	2255	0.020	0.018	0.021	0.022
300	350	7476	7218	7380	0.017	0.016	0.016	0.027
				Accuracy				
100	900	207	-	179	0.051	-	0.082	0.048
200	900	347	27	324	0.055	0.037	0.060	0.044
300	800	420	82	465	0.073	0.044	0.059	0.0649
100	500	219	236	242	0.022	0.037	0.022	0.042
200	500	434	487	500	0.018	0.035	0.020	0.030
300	350	917	895	914	0.019	0.033	0.021	0.026

Table 1. Vertical precision and accuracy results for the different PRFs and flying altitudes. Data collected with the Titan multispectral LiDAR. The results for the separated channels and for the combination are presented [37].

2.3.3 Single Photon LiDAR

SPL has attracted interest in topographic mapping because of the recent features in the laser systems. SPL systems especially allow the higher flying altitude and higher data rate. The performances and differences of the SPL systems are investigated by comparing the systems to the conventional airborne LiDAR systems and/or to field measurements. The SPL is seen to strengthen the mapping of tree canopy structures with the wide coverage, the efficient 3D structures, and the better capability to penetrate dense canopies [57]. The filtering process of the ground points for creating the digital elevation models (DEM) still contains uncertainty. The problem is related to the higher point density captured from the canopy which leads some falsity in the ground point classification. The importance of the calibration of the system and noise filtering are highlighted to avoid the problem [28, 32].

In the large scale mapping the SPL system is seen to be suitable with competitive location and resolution accuracies. As an example, in the study by Degnan (2016) [29], the SPL has met the highest requirements of the USGS with the point density and the elevation accuracy. The USGS is the national program in the United States to collect the high-resolution elevation data for the public use. In the test experiments were used Sigma's HAL system and HRQLS-1 system, whose characteristics are almost equivalent to Leica's SPL100. In the study by Degnan [29], the elevation accuracy received with the HAL system from the altitude of 7.6 km was 10 cm and the vertical accuracy of the HRQLS-1 was 3 cm from the altitude of 2.3 km. In the study in 2016 by Stoker et.al. [56], was received 17 cm vertical accuracy from the altitude of 2.3 km with the Sigma's HRQLS system. Also in bathymetric mapping the SPL has proved centimeter accuracy in certain experiments [59].

A case study using the SPL and conventional LiDAR data was executed in the City of Vienna [31]. The study investigated land cover classification, but also the point cloud deviations for the land cover classes, including buildings (Fig. 8). Based on the research paper, the waveform LiDAR provides a bit better accuracy with sharper and concise data, but the strength of the SPL system is in the coverage performance and vegetation capability. The point cloud dispersion is affected by the roof type, but the SPL100 system seems to be less precise than the waveform LiDAR. The local 3D point cloud deviations were calculated for flat and tilted roofs, and the deviations of 1.0 cm for flat roofs, and 9.5 cm for tilted roofs were achieved with the post processed SPL data. The waveform LiDAR achieves 0.7 cm (flat roofs) and 3.1 cm (tilted roofs) precision. The point cloud deviation is said to be one to five times higher, depending on the surface type, for the SPL data than for the waveform LiDAR.

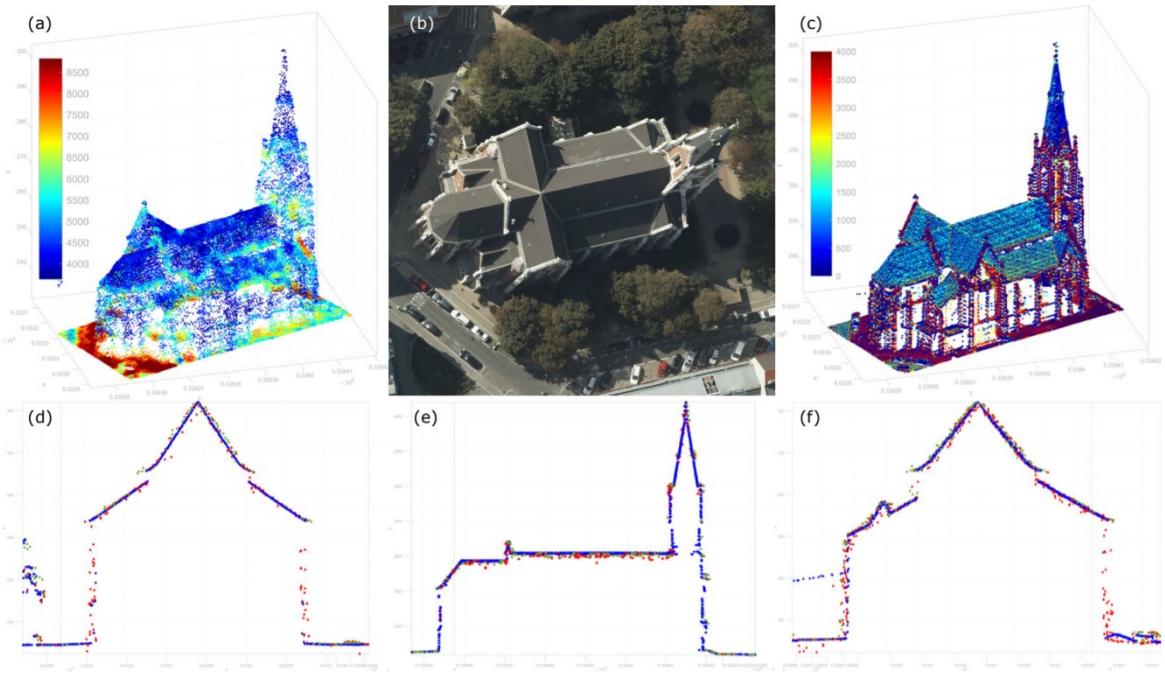


Figure 8. 3D point clouds of a church. a) SPL point cloud colored by intensity, b) aerial image, c) waveform LiDAR point cloud colored by signal amplitude, d) cross section, e) longitudinal section, f) diagonal section. Unfiltered SPL data with red points, postprocessed SPL data with green points and waveform LiDAR data with blue points [31].

3 Materials and methods

3.1 Study area

The area of Espoonlahti is located in Espoo, Finland, about 15 km west of Helsinki. Espoonlahti is used as a regular test site of research of the FGI, where have been made several ALS and MLS studies. The area was selected because the proper density, complexity and type of the buildings for the purpose of the project and the interest toward the urban environment. The size of the study area is about 265 ha and it has quite wide variety of houses with detached houses, row houses, and high-rise buildings (Fig. 9).

Figure 9. The study area of Espoonlahti [60].

3.2 Survey and equipment

3.2.1 ALS

Airborne laser scanning datasets were collected with helicopter and with fixed-wing aircraft. All the datasets included LAS or LAZ –files of the point clouds, the point cloud files of the one minute of flight and the trajectory information (TRJ –files) of the SPL100 and Titan flights. This section presents the laser scanners and the specifications of the data collection. The overview of the laser scanner datasets can be found from Table 2. The laser scanners used in the ALS campaigns:

VUX-1HA from Riegl was introduced in 2015 and is a high accuracy LiDAR sensor for kinematic laser scanning. The scanner uses near infrared wavelength (1550 nm) and is produced to operate in mobile laser scanning with different platforms. Main applications according to the manufacturer are indoor and outdoor mobile mapping, tunnel measurements and railway applications, but it is also suitable for the ALS systems. The scanner is lightweight (3.5 kg) and compact with the high scan speed and measurement rate, where the pulse frequency is up to 1000 kHz. It uses rotating mirror -scanning mechanism (Fig. 4), which provides regular point pattern with parallel scan lines [61]. The VUX-1HA data was collected from the helicopter in May 16, 2019. Flight altitude was about hundred meters above the ground level and the flight speed around 50 km/h. Swath width was between 400 and 500 meters with the 40 to 50 percent overlap. The point density used in the data processing is 71 pts/m² with the total amount of 175 803 560 points. The real number of collected points was over 800 million and the used amount was fifth from the total number of the points. Thinning had to be done to make the data classification possible. The scanner was located on the front side of the helicopter to measure forward to the flight direction with the angle of 15° from the nadir. VUX-1HA data was used as a reference, because the foreknowledge and presumption that the data is accurate, dense and reliable for the purpose of the project. The manufacturer reports the accuracy to be 5 mm at the range of 30 m under RIEGLs test conditions. In this project the accuracy is about 5 cm in x, y and z –coordinates. In the work and in the results the VUX-1HA data is simply called as VUX.

MiniVUX-1UAV from Riegl came in 2016 and is a very lightweight scanner (1.6 kg) designed mainly for unmanned vehicles. It uses near infrared wavelength (905 nm) and the main application is the aerial mapping with the multi-rotor, rotary-wing or fixed-wing UAVs. The device is suitable in many applications, but the purpose, according to the manufacturer, is to measure snowy and icy terrain with the multiple target capability (1-5 echoes per laser shot). Scan speed and measurement rate are lower than in the VUX-1HA and the PRF of the miniVUX is up to 100 kHz. Scanning mechanism is rotating mirror (Fig. 4) [62]. The data was collected from the same helicopter flight with the VUX-1HA, so the flight speed and altitude were equal. The scanner was located on the front side of the helicopter with the angle of 15° from the nadir. Swath width was between 300 and 400 meters with the overlap of 30 to 45 percent. Point density of the dataset is 29 pts/m² with the total amount of 71 081 901 points. In the work and in the results, the simplified miniVUX means the miniVUX-1UAV data.

SPL100 from Leica is the first commercial Single Photon LiDAR system. The system uses laser wavelength of 532 nm and divides the beam into 100 outlet beams, to low the cost per data point, and to make the data collection more efficient. The device was introduced in 2017 and is suited for large area terrain mapping. The scanner weights about 84 kg and is typically used from the airplane. The scan speed and measurement rate are high because of the number of the beamlets. The PRF capability of the laser raises up to 6000 kHz. Scanner uses Palmer scan system (Fig. 4), which provides circular point pattern [63]. SPL100 measurements were done from an airplane in August 8, 2018. This work includes two SPL datasets, SPL100_65 and SPL100_120, which were collected from the different altitudes: 1920 and 3600 meters above the ground level. Swath widths were 1050 and 1950 meters with the overlap of 20 to 25 percent. Point densities were 80 pts/m², with the total amount of 197 986 050 points, for the lower altitude (SPL100_65) and 39 pts/m², with the total amount of 105 959 120 points, for the higher altitude (SPL100_120) flights.

Titan from Teledyne Optech is the first commercial multispectral airborne LiDAR sensor, which was released in 2014. The system operates from the airplane and the weight of the sensor and control unit is more than 100 kg. Titan uses three active beams with independent wavelengths (532 nm, 1064 nm and 1550 nm) to collect the spatial and spectral information of the target. The main applications are topographic mapping and modeling, land cover and vegetation classification, and bathymetric survey. The sampling rate is up to 300 kHz per channel and combined sums up to 900 kHz. Scanner uses oscillating mirror system (Fig. 4) [8]. The data was collected in May 2, 2016 with the flight altitude of 700 meters and speed of 250 km/h. Swath width was over 500 meters with the overlap of 30 percent. The total amount of points was 152 041 797, and the point density 56 pts/m².

	VUX-1HA	miniVUX-1UAV	SPL100_65	SPL100_120	Titan
Flight date	16.5.2019	16.5.2019	17.8.2018	17.8.2018	2.5.2016
Platform	Helicopter	Helicopter	Airplane	Airplane	Airplane
Flight altitude (m)	100	100	1920	3600	700
Flight speed (km/h)	50	50	220	280	250
Wavelength (nm)	1550	905	532	532	532/1064/1550
PRF (kHz)	1017	100	5100	5900	300 (x 3)
Scan frequency (Hz)	200	100	9 (540 RPM)	8 (490 RPM)	40 (x 3)
FoV (°)	360	360	30	30	40
Swath width (m)	400-500	300-400	1050	1950	510
Overlap (%)	40-50	30-45	25	20	30
Point density (pts/m2)	71	29	80	39	56

Table 2. Specifications of the laser scanner datasets.

3.2.2 Field measurements

Beside the ALS campaigns, field measurements were done in June 18 – 20, 2019. The ALS data were examined before the field measurements, so that the field measured data would include as many clearly visible targets as possible, and would cover the area comprehensively. The measurements were done with the GNSS receiver HiPer HR from Topcon [64] with real-time virtual reference system (VRS) service from commercial Trimnet VRS [65], and the total number of the measured points was over 600. The measured data was used mainly as a visual reference in the investigation of the ALS datasets, because most of the points were ground points and weren't needed in the registration of the datasets. Most of the ground points were taken from the paintings on the roads or edges of the sidewalks. Some points were taken from the corners of the roofs of the garages, trash shelters or small storehouses, wherever they were possible to measure with the handheld device. The points from the roofs were used in the coordinate comparison. The VRS RTK accuracy of the equipment is horizontally 5 mm + 0.5 ppm and vertically 10 mm + 0.8 ppm [64].

3.3 Data processing

3.3.1 Software

SPL100 and Titan point cloud datasets used in the thesis were ordered from third party companies (Aerodata AG and TerraTec Oy), who executed the data capture and handled the preprocessing. The trajectory and point cloud processing for the SPL100 datasets were done with the company's own software. For the Titan data the supplier used Optech's and TerraSolid's software in trajectory and point cloud processing. The VUX and miniVUX datasets were already used in an other intention, why the charge for the preprocessing wasn't the author's, even if it is an important part in the data processing and quality ensuring. Cleaning, registration and georeferencing were done with the Terrasolid software.

Most of the work in this thesis is done with the TerraScan software from Terrasolid Ltd. The software is generally used in the laser scanning point cloud processing and managing. The application can be used for example in the city, forest and powerline modeling or road surveying. It provides tools for the point cloud classification, data organizing, trajectory managing and point cloud visualizing, in a 3D view. TerraScan supports several import and export formats and the automatic routines can be combined with manual editing. TerraScan works as an integration with MicroStation, which is a design modeling software from Bentley Systems. MicroStation provides tools for CAD environment [66, 67]. In this thesis the software is used to organize the data, in the point cloud classification and in the 3D modeling of the buildings. The CAD tools are used as a help in the calculation.

The other software used in the point cloud analyzing is the open source and free software CloudCompare, which started in 2003 as a project of PhD of Daniel Girardeau-Montaut. The software is a processing platform for the 3D point clouds. It was originally designed for comparing the point clouds and meshes, but nowadays includes also more processing algorithms, for example for the registration, resampling, color handling and segmentation [68]. In this work the software is used in the investigation of the point deviations by calculating the cloud-to-mesh distances.

3.3.2 Processing

In the thesis the aim of the processing was to separate the roof data out of the other data points to allow the investigation to concern only the roofs. The processing started by dividing the area into four, more or less equally sized, blocks and by classifying all the datasets using TerraScan macros. The datasets were divided into four blocks to ease the calculation and processing. The huge amount of points has impact to the speed of the automatic process by extensive training, which may lead into the problems, when the software can't handle the amount of the data. The speed of the automatic data processing is also related to the power of the computer, what is affecting at least to the number of the points processed at a time. The user can add and edit macro steps to automate the processing which includes classifica-

tion, modifying, deleting, transforming and outputting of points, updating views and executing commands. Macros can be adjusted individually to each dataset to take into account the characteristics of the dataset.

In this work own macros were created for each dataset to find the best possible automatic classification in every case. Each macro set included definitions for ground, vegetation and buildings, and they were run in this order. Ground classification was started by searching low points where the user defines the routines to search the points lower than the surroundings. Low points may also be error points, and therefore filtering must be done before ground point classification. Setting the parameters to the ground option include the building size, iteration angle and distance adjustment. The ground routine uses triangulated surface model to determine the ground data. Model keypoints were also created from the ground points to ensure the possibility to create the surface models. Keypoints are usually a thinned ground data with smaller set of points, which the user has classified by defining the tolerance of the elevation and the point density of the data. Hard surface routine detects surfaces, which are not belonging to vegetation, and are typically roads and buildings. The method is suggested to be used, instead of ground routine, for mobile laser datasets and not for airborne datasets. Typically, in airborne datasets the majority of the data is belonging to the natural terrain. However, in our case the Espoonlahti test site has quantitatively many hard surfaces, why the routine was used to detect points from the roads and roofs. Hard surface routine works almost similarly with the ground routine, but is not sensitive to low points. It creates triangulated surface model iteratively and is interested in the points that form a local plane to add the points to this class. Low points are not necessary to find with this classification method.

Vegetation is defined based on the height from the ground and were categorized in low, medium and high vegetation classes. When defining the higher point classes, the first step was to set all but the hard surface data above 2.5 meters from the ground, to the high vegetation class. Roofs of the buildings can be searched by setting the rules, minimum size and height tolerance of the planar surface. In this case the surfaces were searched from the high vegetation and hard surface classes. The routine starts the building classification from the holes on the ground, and finds possible planar surfaces above the holes, before moving forward [69]. Figure 10 shows the classified roof points from the example building of the kindergarten, which surface area is about 716 m².

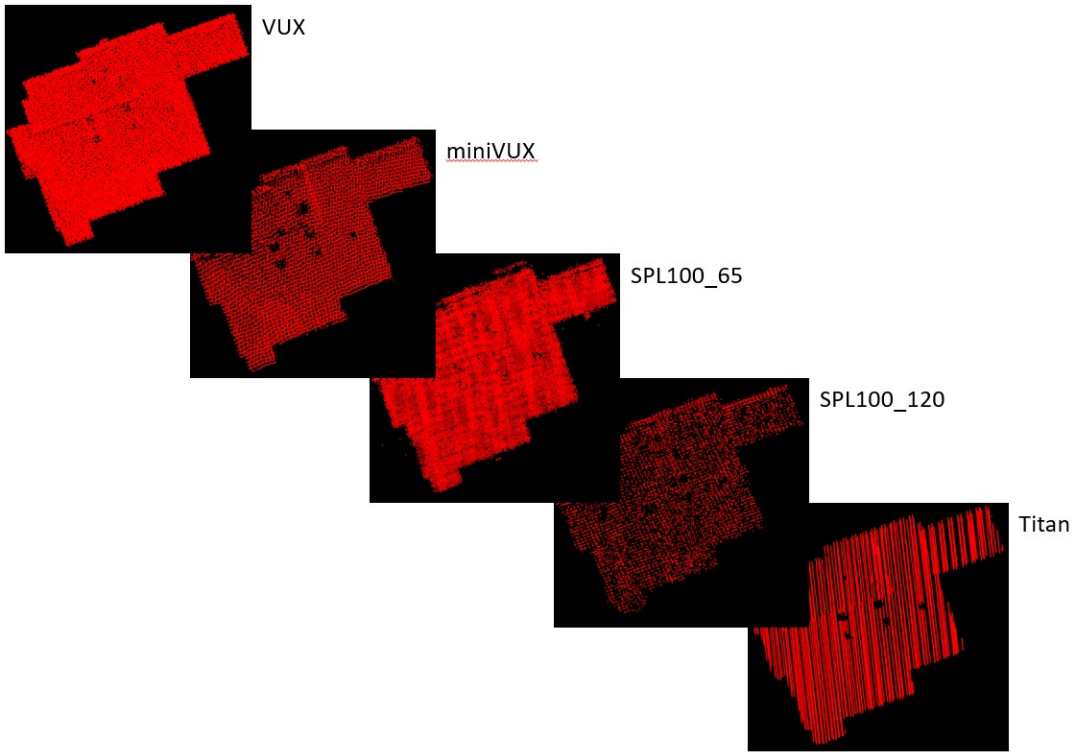


Figure 10. The example results of the roof classification and the visual differences between the point clouds. The points belonging only to building -class are visible. The size of the roof is about 716 m².

For the project different roof types were selected to see if the type has any impact to the results or differences in the data. Four of the roof types were common or basic types and one group included the other and more complex structures. They were selected by searching the building types from the area and by examining the point clouds. Roof types were classified into five groups: flat, oblique, gabled, conical and other types. Selection of the roofs was planned so that every block would include three buildings from each class and from the opposite locations on the block. At the end the number of examined buildings was 57. From the classified point clouds of the blocks, the smaller samples with selected buildings were separated. In addition to automatic classification, some manual work was done to improve especially the roof classification. Without the classification of the ground, building and other higher objects, the vectorization tool of the 3D models is not able to work. The requirements must be fulfilled before the user can define the settings to construct the buildings. All the sample areas were stored individually to fast binary –files, which were used to create the 3D models of the buildings (Fig. 11). The roof points were also stored as a LAS 1.4 –format, which can be loaded and used in the CloudCompare.

TerraScan has a tool for automatic 3D vector modeling which result can be modified by manual editing. The vectorization works when the user is able to show the classified building roof and ground points. Also building footprints can be used in the vectorization to determine the locations of the walls or roof edges, but it was not used in this project. The vectorization is based on the ground and roof points, so the quality of the modeling is determined by the result of the classification and the laser data itself. The density of the laser data usually affects to the quality, and also the TerraScan user's guide tells that more than 10 pts/m² would

allow accurate and detailed models with the roof structures in LoD2 (Level of Detail). This can be also seen as a possible drawback of the vectorization, because it is totally dependable of the quality of the source data, when the missing or bad quality data can lead to the confusion. Examples were, when some of the sights had occlusions due to trees or other buildings and the construction of the buildings were incorrect. These kind of problems were tried to fix by manual editing, but the true location of the building edges could not be determined. One of the notable advantages in TerraScan is that the automatic building vectorization works in a short time for large areas with many buildings and allows also complex roof modeling [69].

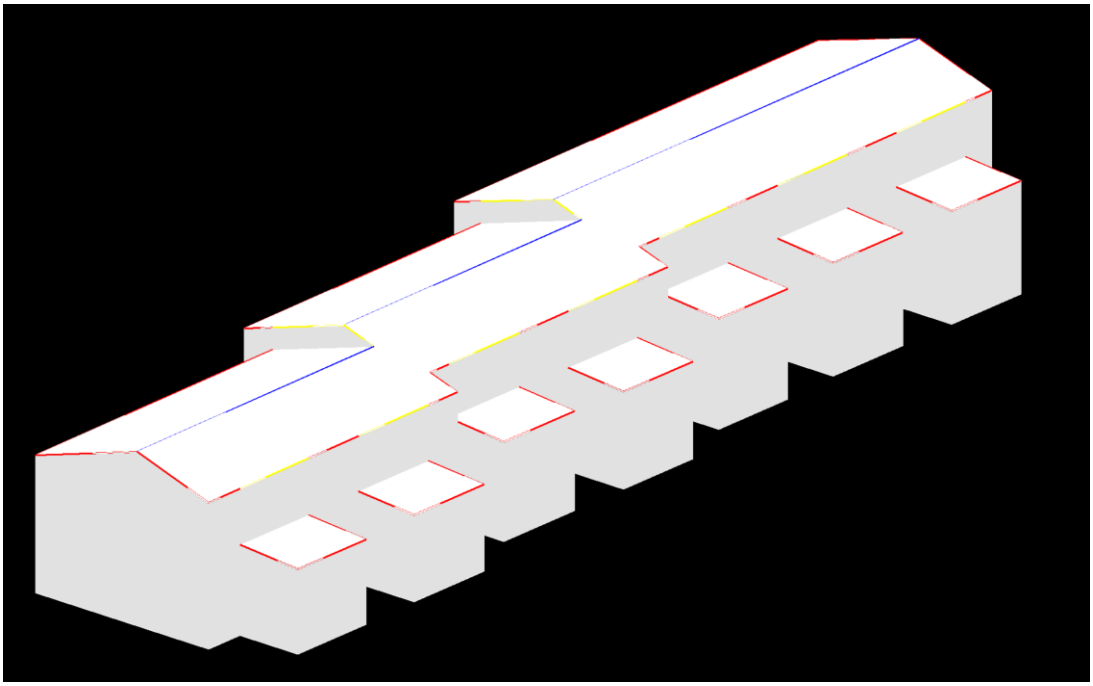


Figure 11. 3D building model with constructed roof lines from the VUX-1HA data. Gabled roof. LoD2.

From the 3D models the roof lines were constructed and used in the data analyzing (Fig. 11). Construction is simple and the roof lines follow the edges of the model. The line colors are determined by the edge types in red (outer edge), blue (intersection of planes) and yellow (elevation jumps) [69]. TerraScan provides an automatic tool to create the line elements from the models, which was used in the project, but also the possibility to construct the roof polygons from the roof lines. The roof lines determine the outlines, from where the polygon tool tries to create the 3D roof polygons, which can be used to reconstruct the 3D building models. Either point or vector data can be used to reconstruct the building models, which shows some flexibility of the software. In this thesis all the data was laser data in the classification and in the modeling, and no vector or image data were used besides, and as was mentioned, the level of automation was quite high which speeds up and eases the processing.

3.3.3 Comparison methods

The topic of the thesis is to compute and compare the different features of the roof data (were listed in the objective 1.2). Notable are that some of the computations are done automatically with the TerraScan and CloudCompare software, and it is thought that the basic calculations of the surface, height, length, width, point density and inclination are not necessary to explain in detail.

In this paper most of the results are expressed as an average and standard deviation values. The idea is to show the amount of variance in the data. The selected roofs used in the data comparison, are not covering all the buildings, why sample standard deviation is used, instead of population standard deviation. Standard deviations are used in the sections 4.2 – 4.7 to describe the distributions of the datasets. The expression of the formula is found from the chapter 2.2 (Equation 5).

The distances between the points are calculated in the section 4.4, when the planimetric corner coordinates are investigated. The equation needs at least two x and y –coordinates, which in this case are the reference coordinates from the VUX data and the corresponding point from the other ALS datasets. The calculations are done separately for the corner points (4.4.1), and for the corners of the roof outlines (4.4.2). The calculations are done for each corner, but the figures (Fig. 15 and 16) show only the average differences of the correspondences of each dataset. The used formula is found from the chapter 2.2 (Equation 6).

The root mean squared error (RMSE or RMS) is used in the section 4.7 to describe the deviation of the points on the roof plane. RMS computes differences between the measured values and the reference plane. The idea is to calculate the differences between individual points and the correspondence on the reference plane, and to show the statistics about the data distribution. The reference is the VUX data. The final RMS –value performs the fitting of the points on the roof plane, thus estimating the accuracy. The expression of the equation is found from the chapter 2.2 (Equation 7).

4 Results

The results of the data comparison which cover previously mentioned point densities, surface areas and outlines, planimetric differences of the corner points, heights, inclinations and point deviations are presented in this chapter. Some results are compared to the measured VRS RTK-points, but the most to the VUX-1HA data. The results are based on the roof data, except the point densities. The point cloud classification is done twice and partly manually to increase the reliability. This chapter presents the results for the roof feature comparison.

4.1 Point densities

The densities of the point clouds, and how they are distributed between the four blocks, are presented in Table 3. Point densities are calculated with the TerranScan's measuring tool, which measures the average number of points per square meter. As was mentioned, VUX data is a fifth from the original total amount of points to make the data processing practicable. The lower point density of miniVUX is notable. The VUX and miniVUX sets don't cover the entire blocks two and four, and the SPL100_65 don't cover the entire block four, from the southern side of the test area. Both VUX and miniVUX data miss about 13 ha on the block two and 16 ha on the block four, and the SPL100_65 about 22 ha. VUX data is missing approximately five million points, miniVUX a bit over two million points and the SPL100_65 a bit more than 17 million points. The western side of the test area is partly water and the water points are missing from the VUX and miniVUX datasets which low the number of points compared to the SPL100 and Titan datasets. The three Titan channels are also separated.

	Block1	Block2	Block3	Block4	Overall pts	pts/m²
VUX	58 348 910	27 912 227	59 120 754	30 421 669	175 803 560	71,85
miniVUX	23 473 609	11 150 728	23 783 685	12 673 879	71 081 901	29,72
SPL-100_65	51 271 396	49 327 887	58 258 192	39 128 575	197 986 050	80,45
SPL-100_120	28 969 490	21 247 210	20 742 587	34 999 833	105 959 120	39,73
Titan	38 418 137	38 137 229	34 642 752	40 843 679	152 041 797	56,80
Titan Ch1	12 689 323	12 661 934	11 233 295	13 533 525	50 118 077	18,82
Titan Ch2	14 154 011	14 332 871	13 007 709	15 399 241	56 893 832	21,31
Titan Ch3	11 574 803	11 142 424	10 401 748	11 910 913	45 029 888	16,85

Table 3. The number of points and densities of the datasets.

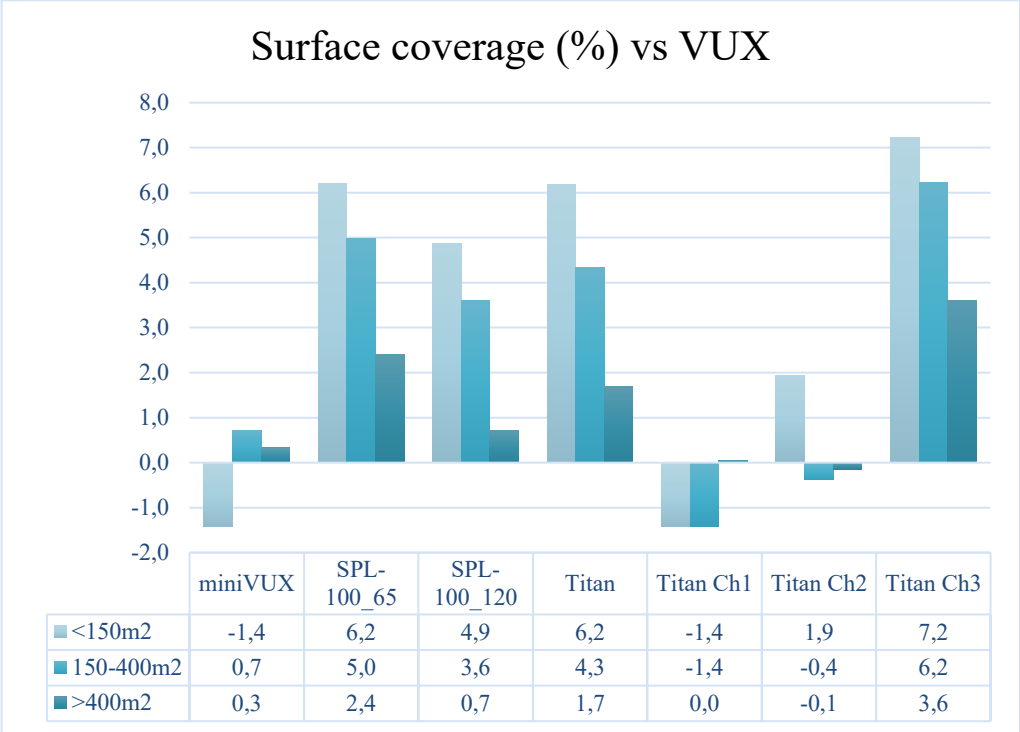
4.2 Surface areas

Surface areas of the roofs are calculated in TerraScan from the 3D building models. Few buildings were left out because of the fuzzy models of complex structures, but the rest were grouped for small (under 150 m²), middle (150 – 400 m²) and large sized buildings (over 400 m²). Small roofs were trash and car shelters. Middle category has detached houses and shelters. Large roofs were row houses, high-rise buildings and markets. The difference for each surface was calculated by subtracting the data value from the VUX value, and for the groups the mean differences and standard deviations were calculated. They were used in the investigation to explain and support the results, but the figures shown here are based on the surface coverage compared to the VUX data. The sizes of the surfaces were calculated, and they were summed to see the total coverage of the buildings, and for the total surfaces the percentage calculation was used to describe the results reasonably. The differences of the roof sizes were large as the smallest roof is 34 m² and some of the largest roofs are almost 4000 m², so the sizes of the roofs had big impact to the mean and standard deviations.

The mean differences are small in the miniVUX data and in the Titan channels 1 and 2, where the surfaces differ 0 - 5 m² despite the sizes of the roofs. Similar behavior can be seen from the Figure 12 where the area difference for every group is under two percentage. The SPL100_65 and the Titan channel 3 have the biggest differences compared to the VUX data. From the calculations it could be seen that all the roof sizes differ occasionally several or even hundreds of square meters, and as a mentioned the mean difference of the large sized roofs for the SPL100_65 was -40 m² and for the Titan channel 3 -60 m². The coverage calculations (Fig. 12) also show clearly the largest differences of the SPL100_65 and the Titan channel 3. The surface differences in the large sized group were partly very huge, but when the percentages are investigated, the results show that the largest roofs have the smallest differences in each of the datasets. The proportional values were used, because otherwise the errors in the larger buildings were huge compared to the smaller buildings, and the results would have been misleading.

One explanation for the larger differences in the SPL100 and Titan datasets, can be related to the impact of the laser beam divergence when hitting the target, but alone it doesn't explain the errors of hundreds of square meters. The datasheets of the scanners show that the beam divergence of the visible wavelength is about 0.2 – 0.4 mrad larger than in the infrared channels [8, 58, 59, 60], which might be tens of centimeters on the ground or at the target plane. Flat high-rise buildings seem to differ quite much in the SPL100 and Titan datasets. Similar behavior can be seen for the lower altitude SPL100 data and for the Titan channel 3, in the case of the gabled roofs. The semi-automatic building model vectorization used in the project might increase the chance for the errors in the models. The variety of the automatic and manual reconstruction may bring some unsteadiness to the data. The accuracies of the corner points of the roofs are revealed later, but they may also lead to few square meter differences in the surface areas. Notable is that the surface areas vary irregular when they are calculated from the miniVUX and first two channels of the Titan. In contrast to that, the SPL100 and Titan channel 3 construct the buildings constantly larger than the VUX data, which are seen from the results on the Figure 12. In the all building groups (small, middle and large) the largest differences can be found from the Titan channel 3, where the maximum values are 20.0 m² in the small sized buildings, 53.1 m² in the middle sized buildings and 199.2 m² in the large sized buildings. Minimum and maximum surface differences for each dataset can be found from the Appendix 1.

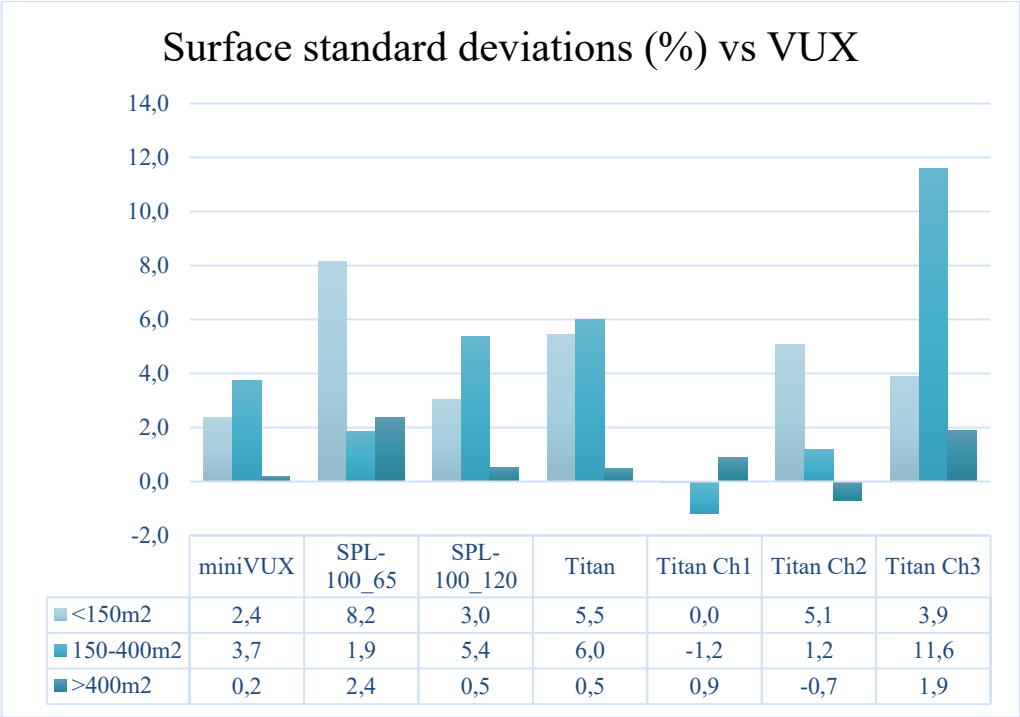
In each dataset, when the percentages are investigated, the smaller roofs seem to have a bit more variance than the larger buildings. The smaller size range of the small and middle sized roofs (than in the large sized roofs) may explain the effect. Some point clouds of the smaller buildings are also a bit sparser than the larger buildings and that's why they are more sensitive for errors. The smaller roofs are mostly shallow trash and car shelters and the sparsity is caused by the trees next to the buildings. Shelters are also more exposed to occlusions from the higher buildings and trees.



	miniVUX	SPL-100_65	SPL-100_120	Titan	Titan Ch1	Titan Ch2	Titan Ch3
<150m2	-1,4	6,2	4,9	6,2	-1,4	1,9	7,2
150-400m2	0,7	5,0	3,6	4,3	-1,4	-0,4	6,2
>400m2	0,3	2,4	0,7	1,7	0,0	-0,1	3,6

Figure 12. Percentage differences of the surfaces compared to the VUX data. Colors present the groups based on the sizes of the roofs.

Sample standard deviations were calculated for each dataset and for each building group (Equation 5). The Appendix 1 shows clearly the effect of the large sized roofs, where the deviations are remarkable. The used grouping for the roof sizes also highlights the differences of the large surfaces versus the small or middle roofs, where the areas are sometimes even ten times bigger than in the small and middle sized buildings. That's the reason, why the percentage calculation was also used for the standard deviations of the datasets. In overall seems that the miniVUX and Titan channels 1 and 2 have a bit smaller deviations than the other datasets (Fig. 13). The SPL100_65 and the Titan channel 3 have larger deviations in the small and middle sized roofs. The deviations of the large sized roofs seem to be sufficient in all the datasets. Based on the surface coverage and standard deviation calculations, the conclusion is that in the surface area investigation would be worth to use the miniVUX instead of the SPL100. From the Titan data the channel 3 should be separated to get more accurate information.



	miniVUX	SPL-100_65	SPL-100_120	Titan	Titan Ch1	Titan Ch2	Titan Ch3
<150m2	2,4	8,2	3,0	5,5	0,0	5,1	3,9
150-400m2	3,7	1,9	5,4	6,0	-1,2	1,2	11,6
>400m2	0,2	2,4	0,5	0,5	0,9	-0,7	1,9

Figure 13. Percentage differences of the standard deviations compared to the VUX data. Colors present the groups based on the sizes of the roofs.

4.3 Roof outlines

The lengths and widths of the buildings were investigated separately, even if the surface comparison is quite much related to the outlines. However, in this section it is practicable to describe the mean differences of the outlines (Fig. 14) in meters, instead of percentages, because the relations of the sizes of the roofs don't affect similarly to the results as previously. The lengths and widths were investigated separately and combined, and were based on the created roof lines of the 3D building models. Buildings were divided into two groups based on the total lengths of the outlines (length + width): outlines under 40 m (Mean_S) and outlines more than 40 m (Mean_L). The VUX data stood as a reference. The smallest differences were in the miniVUX data, 0.4 cm (Mean_S) and 1.5 cm (Mean_L), and the largest differences were in the Titan channel 3, -49.3 cm (Mean_S) and -53.9 cm (Mean_L). Notable detail was that the miniVUX seem to map the building lengths a bit shorter and the widths a bit wider than the VUX data, which could originate from the scanning directions. The behavior can be seen from the outline figures in the Appendix 2. Otherwise, the separated length and width values (from the Appendix 2) follow the shapes of the combination of the outlines on the Figure 14. Similarly to the surfaces, the Titan channel 3 seems to be the error source and causes the largest differences to the lengths and widths of the roofs. The datasets of the SPL100 and Titan channel 3 map the outlines larger, which are in this case identified from the negative sided columns. The averages of the standard deviations are 33 cm for the building outlines under 40 m, and 46 cm for the building outlines more than 40 m.

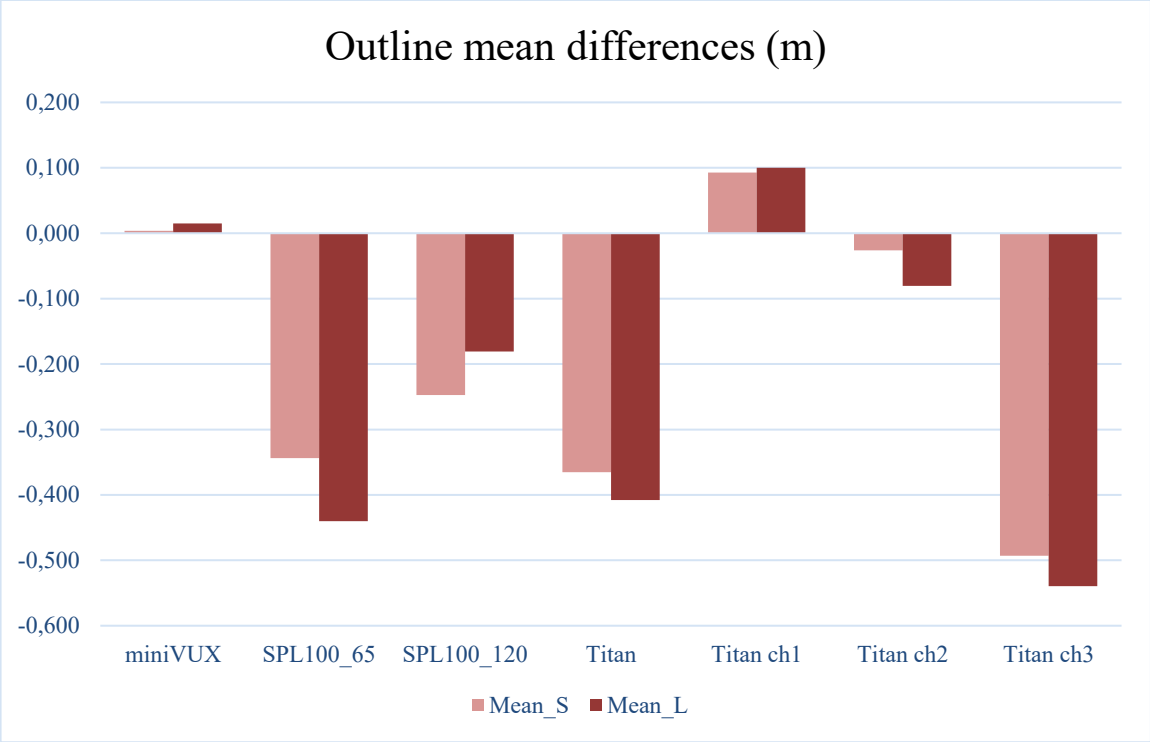


Figure 14. Outline (length + width) differences of the roofs compared to the VUX data. Means_S = outlines under 40 m, Mean_L = outlines more than 40 m.

Based on the results, the data used in the area computation and in the outline investigation are suggested to be the miniVUX data instead of the SPL100, but also the Titan channels 1 and 2 provides competitive accuracy. The behavior of the visible wavelength in the SPL100 and the performance of the third channel in Titan are remarkable compared to the miniVUX and the Titan channels 1 and 2, which should be noticed. To be able to confirm the typical error, caused by the visible wavelength for the hard surfaces, more investigation is needed, but according to this research the laser of 532 nm is more imprecise to map the roof surfaces. The results show that the visible wavelength 532 nm in Titan causes the error and grows the difference to the ground truth. The usage of the visible wavelength constantly widens the roof planes than the infrared wavelengths, and the assumption is that the reason is the beam divergence of the laser. It was unexpected that the higher altitude SPL100 data provides better accuracy than the lower altitude SPL100, at least in the case of the surfaces. In average, the larger roofs don't affect to the outline differences similarly than in the lower SPL100 dataset.

4.4 Planimetric coordinate differences

4.4.1 Corner points

The corner coordinates of the roofs were searched for each building from the five roof type groups. The coordinates of the points were compared to the corresponding point from the VUX data by subtracting the values. For each group of roof types (flat, oblique, gabled, conical and others) the mean differences and standard deviations were calculated. The planimetric distances were computed from the mean values of the x and y –coordinates, which showed the average differences between the true corner coordinates, with the assumption that the VUX data was the true value, and the measured corner coordinates (Equation 6). From the point cloud data and from the coordinate comparison, the systematic error in both of the SPL100 datasets was noticed. To remove the systematic error from the coordinates, the field measured roof points were used as a reference, 67 points for the SPL100_65 and 65 points for the SPL100_120. In the SPL100_65 data the error was: x = 0.527 m, y = 0.591 m. In the SPL100_120 data the error was: x = 0.465 m, y = 0.717 m. Figure 15 shows the planimetric accuracies, technically the distances between measured points and the reference points, of each dataset and each group of roof type, compared with the VUX data. Black line shows the average differences of all roof types combined. The smallest and largest planimetric differences are both in the Titan channel 2: flat roofs 0.003 m and other roofs 0.132 m.

For the miniVUX, the gabled roof data (green) seem to reach the accuracy of 1.5 cm. All the roof types differ 1.4 – 11.9 cm from the VUX data and the overall distance difference was 4.3 cm (x = 4.1 cm and y = -1.4 cm). The miniVUX data is also most sparse point cloud and achieves similar accuracy than the denser clouds, even if the usual assumption is, that the sparser point clouds provide lower accuracy. Standard deviation was 26 cm (x = 29 cm and y = 23 cm).

The averages of the SPL100 coordinates are in a range of 1.4 – 8.1 cm after the correction of the systematic error. The Leica's datasheet reports the horizontal accuracy to be under 15 cm at the altitude of 4000 meters with the flight speed of 360 km/h. Based on the calculations, the results of the both SPL100 datasets seem to be in the reported accuracy. The distance difference of the SPL100_65 was 3.8 cm (x = -3.4 cm and y = -1.7 cm) and the SPL100_120 was 4.3 cm (x = -2.9 cm and y = -3.2 cm). The lower altitude and lower flight speed (than in Leica's test experiment) are affecting positively to the data. Also the dense point clouds have an impact to the accuracy of the corner coordinates. Standard deviations were 33 cm (x = 33 cm and y = 32 cm) for the SPL100_65, and 25 cm (x = 24 cm and y = 26 cm) for the SPL100_120.

Teledyne Optech reports the horizontal accuracy to be calculated from the equation: 1 / 7500 * altitude. In this flight campaign the altitude was 700 m, so the accuracy should be around 9.5 cm. The average distance difference was 3.8 cm of all the roof types combined and the reported accuracy seem to be reached reliably. The mean difference of the x-coordinate is 0.6 cm and the y-coordinate 3.7 cm. The peak in the group "others", which lead to the channel 2, can be explained by the low number of the points collected from these roofs. There were only two buildings involved in the calculation and can be assumed that using the larger sample size it would compensate the distribution of the coordinates. Standard deviation of the Titan data was 31 cm (x = 32 cm and y = 30 cm).

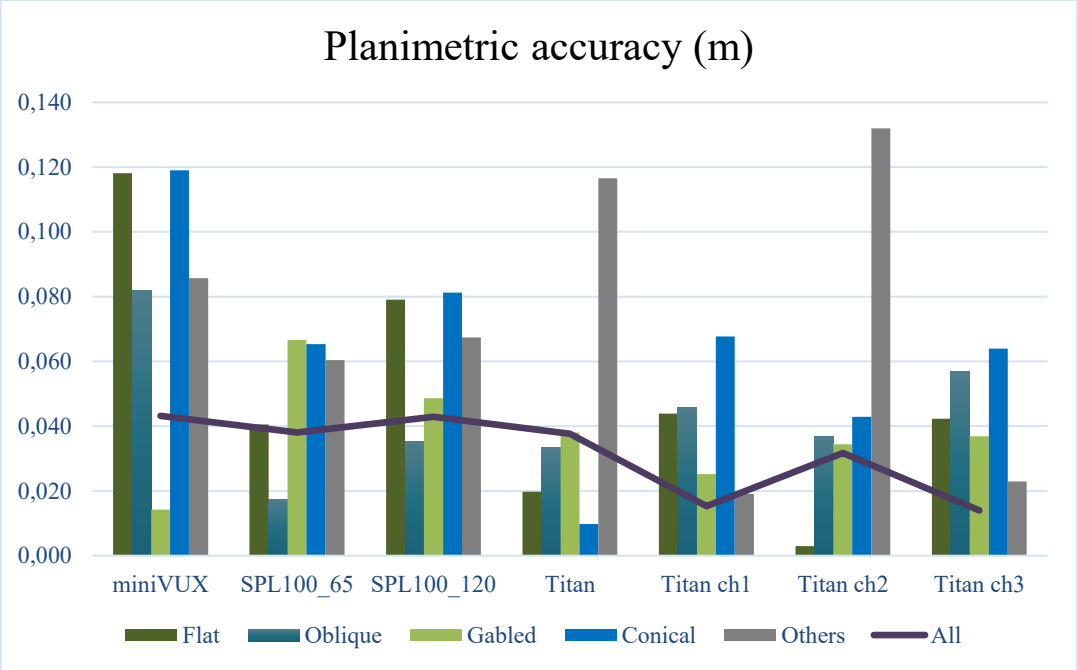


Figure 15. Planimetric (x, y) accuracies of the corner points of the roofs, compared to the VUX data. Bars present the roof types: flat, oblique, gabled, conical and others. Black line is the average of all roofs.

Titan's three wavelengths were examined separately. Almost all the average coordinate differences of the roof types are under 6 cm. The peak in the channel 2 may be the result of the low number of the points as previously mentioned, where the complexity of the roofs also affects. Surprising was that the mean difference of the channel 3 is that small, when the influence to the previously investigated surfaces and outlines were notable. The average distance differences of the points were, 1.5 cm (x = -1.2 and y = 0.9 cm) for the channel 1, 3.2 cm (x = -2.7 cm and y = 1.7 cm) for the channel 2, and 1.4 cm (x = -1.3 and y = 0.5 cm) for the channel 3. The standard deviations were 18 cm (x = 18 cm and y = 19 cm), 24 cm (x = 22 cm and y = 26 cm) and 35 cm (x = 35 cm and y = 36 cm) in the same order.

4.4.2 Outline corners

The coordinate differences were also calculated for the roof outlines of the 3D building models. Similarly the mean, standard deviation and planimetric distance were calculated for the all roof types and compared to the VUX data. The systematic error in the SPL datasets were controlled similarly with the same correction values. Again the smallest and largest planimetric differences were found from the Titan channel 2, where the smallest was 0.013 m for the gabled roofs and the largest was 0.224 m for the other roofs. Figure 16 shows the planimetric accuracies based on the corner coordinates of the building outlines.

With miniVUX, the estimated 5 cm accuracy of the survey is achieved by the gabled roofs (green). The roof types differ in a range of 3.7 – 12.4 cm and the distance difference of all roof types combined with the mean of the x and y –coordinates was 3.4 cm (x = 3.4 cm and

y = -0.5 cm). The standard deviation was 26 cm (x = 30 cm and y = 23 cm). Compared to the coordinates of the corner points of the roofs, the outline corners provide a bit better mean accuracy. The changes are small, but the oblique roofs seem to be a bit more accurate, which affect also to the total average.

All the roof types of the SPL100 datasets differ in a range of 1.6 – 14.9 cm and distance differences are 5.8 cm (x = -4.6 cm and y = -3.5 cm) for the SPL100_65 and 3.5 cm (x = -2.1 cm and y = -2.8 cm) for the SPL100_120. Standard deviation of the SPL100_65 was 32 cm (x = 33 cm and y = 31 cm), and for the SPL100_120 the deviation was 26 cm (x = 26 cm and y = 25 cm). Previously the lower altitude SPL100 dataset provided a bit better accuracy, even if the deviation was larger, but when the outline corners were investigated it is seen that the higher altitude SPL100 provides better accuracy and smaller deviation. The flat roofs increase the error in both of the datasets and they are notably larger than previously. The reconstructed models of the flat roofs seem to have a couple of centimeters inaccuracy, but still the datasets are in the reported 15 cm accuracy.

Titan's roof types differ in a range of 2.0 – 17.6 cm and the distance difference is 6.3 cm (x = 0.9 cm and y = 6.2 cm). The standard deviation was 34 cm (x = 34 cm and y = 34 cm). The corner points based on the outlines of the building are more inaccurate than the manually collected corner points. The conical and other roof types seem to grow the error notably, which can be assumed to come from the 3D model vectorization. The tilted and more complex roofs seem to affect quite much to the building models and to the accuracies of the outline corners of the roofs. The effect and accuracy of the channel 3 is again a bit questionable, when remembering the impact to the surface areas and outlines.

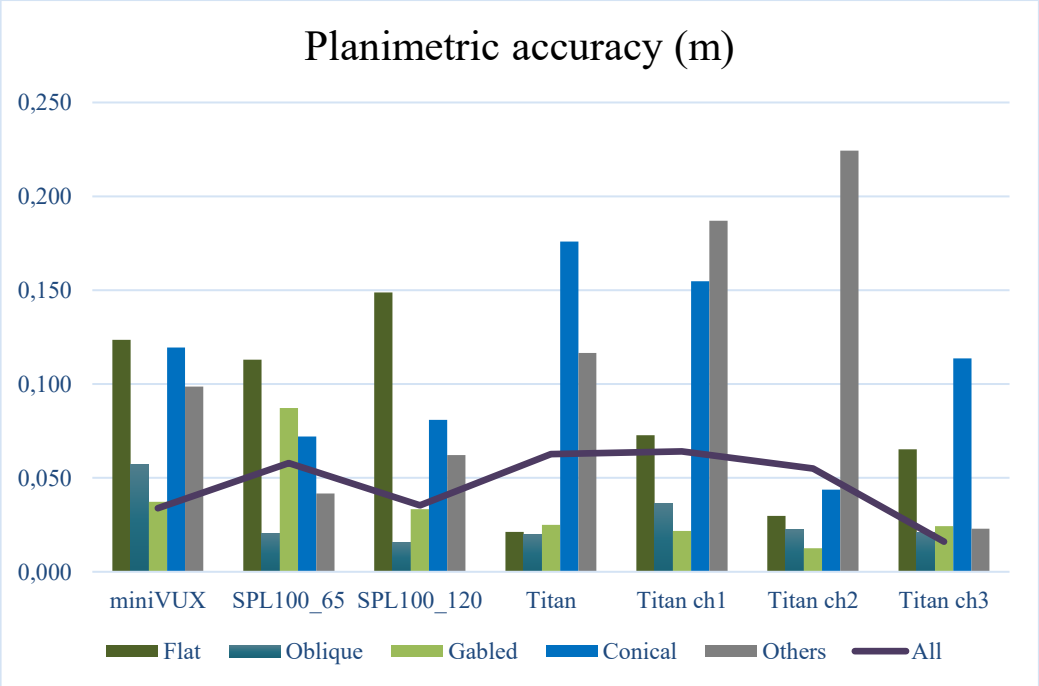


Figure 16. Planimetric (x, y) accuracies of the corner outlines of the roofs, compared to the VUX data. Bars present the roof types: flat, oblique, gabled, conical and others. Black line is the average of all roofs.

The separated Titan channels 1 and 2 have inaccuracies with the buildings “others”, where the largest mean is more than 20 cm. Somehow in the channel 3 the outlines are accurately constructed. Channel 1 has also notable errors with the conical surfaces. The distance difference of the channel 1 was 6.4 cm (x = -3.6 cm and y = 5.3 cm), channel 2, 5.5 cm (x = -3.4 cm and y = 4.3 cm) and channel 3, 1.6 cm (x = -1.4 cm and y = -0.8 cm). The standard deviations were 30 cm (x = 31 cm and y = 29 cm), 25 cm (x = 23 cm and y = 28 cm) and 35 cm (x = 35 cm y = 34 cm), respectively.

4.5 Building heights

Heights of the buildings are examined from the point clouds in TerraScan by subtracting the height of the closest ground point from the corner point of the roof. The height differences are calculated similarly than previously by subtracting the value from the reference data (Fig. 17). Some reference heights are measured with the Topcon GNSS receiver, but the most are from the VUX-1HA ALS data. The height measurements were done for 35 buildings with two opposite corner points of each building. There were found two outliers from the data, which were removed from the calculations, therefore the overall number of points was 68 in each dataset.

Usually the assumption is that the higher point density provides not only better planimetric accuracy but also height accuracy. In advance it was interesting to find out, if similar behavior is involved and differences discovered in the case, where even the lowest density is more than 16 pts/m². From the line on the graph (Fig. 17) can be assumed that the individual Titan channels can almost provide as accurate height information than the higher point density miniVUX and SPL100 datasets. Couple of centimeter increase can be seen in the individual Titan channels. The higher altitude SPL100 data has the best height accuracy based on the results by providing 4.2 cm (std. = 14.4 cm) average difference to the reference data. The largest difference 8.4 cm (std. = 14.4 cm) is in the Titan channel 2. The height determination of all the datasets ranged from one millimeter to 55 cm. Even if the datasets were dense, the most inaccuracies were caused by the occlusions from the trees or from the other buildings. Complexity of some of the roof structures were also affecting to the precision. The ground shape in the area was partly rugged and hilly which augmented some small changes on the ground and especially near the buildings. Some occlusions in the data was also caused by the blocking shape of the ground or the middle height vegetation. These issues caused the situation, where there weren’t points right below the corner of the roof, and the small distance outside from the building caused the difference and inaccuracy in the height determination. Although there were just a couple of these situations, the effect of the issue was approximated to be 1 - 10 cm in height.

The red bars in the Figure 17 show the standard deviations of the datasets which range quite similarly. The smallest deviations of 14.0 cm are in the miniVUX and in the Titan channel 1. The largest deviation of 16.8 cm is in the Titan channel 3.

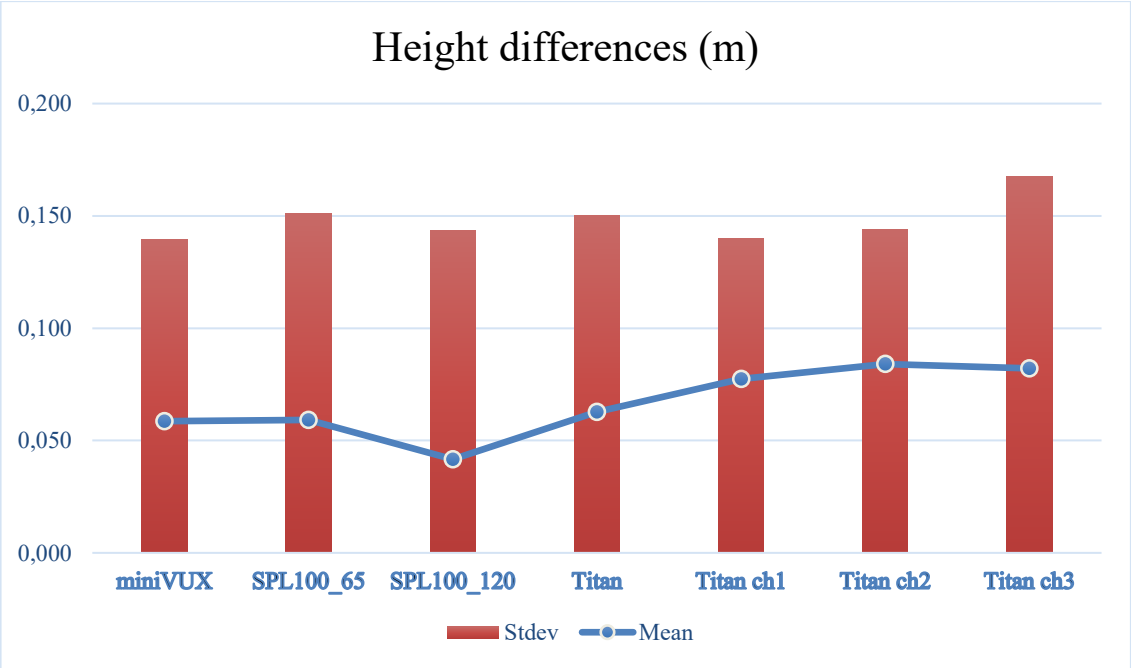


Figure 17. Standard deviations and average differences of the roof heights compared to the VUX data.

4.6 Roof inclinations

Inclinations for the roof types of oblique, gabled, conical and others (just two roofs) were calculated in TerraScan. Flat roofs were ignored after the first calculation of a small sample, because of the assumption that the rest of the differences are also so small and close to 90 degrees, that the investigation is pointless. Finally, the total number of samples for inclination comparison was 32. From the gabled roofs the angle between the roof planes was calculated. From the oblique, conical and other roofs the tilts of the roofs compared to the vertical walls were calculated. The differences to the VUX data were calculated for all the angels of the roofs. The mean and standard deviations were calculated for each dataset. The maximum differences of the samples were picked and removed from the calculus to see the effect of the outliers. Below on the Table 4 are shown the mean values and standard deviations with and without the maximum differences in the data.

	Inclination differences °			Inclination differences ° (max values removed)	
	Mean	Std	Max	Mean	Std
miniVUX	-0,159	0,638	2,4	-0,079	0,479
SPL-100_65	-0,097	0,638	1,4	-0,052	0,553
SPL-100_120	-0,094	0,711	2,6	-0,010	0,547
Titan	-0,203	1,122	4,8	-0,045	0,723
Titan Ch1	0,013	0,247	0,8	-0,014	0,201
Titan Ch2	-0,086	0,759	2,7	0,007	0,579
Titan Ch3	-0,485	1,794	7,8	-0,204	1,060

Table 4. Mean, standard deviation and max differences of each dataset compared to the VUX data. Values are in degrees.

When so called outliers, if for example the maximum value of the Titan channel 1 can be called as an outlier, are removed, the most of the mean differences decrease quite much, and also the most of the standard deviations change in a positive way. This shows that in most of the datasets the maximum differences were individual angles with larger error. The graph showing visually the effect, when the maximum values are removed from the datasets, is shown in Figure 18. The standard deviations ranged from 0.25° to 1.79° when the maximum values were from 0.8 to 7.8 degrees. After the removal, the standard deviations ranged from 0.20° to 1.06° and the maximum values were from 0.5 to 2.1 degrees. In overall the inclination differences were quite small between the datasets and most of the differences came from the small sized roofs including the trash and car shelters. It seems that the smaller surfaces are more error sensitive than the larger surfaces, but more buildings should be included to the research to confirm the behavior. Also, the lower amount of points hitting the surface, the occlusion caused by the trees or from the larger buildings might affect to the inclinations and angles. Also in this investigation it was seen the worst accuracy in the Titan channel 3, where the mean and maximum differences together with the standard deviation were the largest before and after the outlier removal.

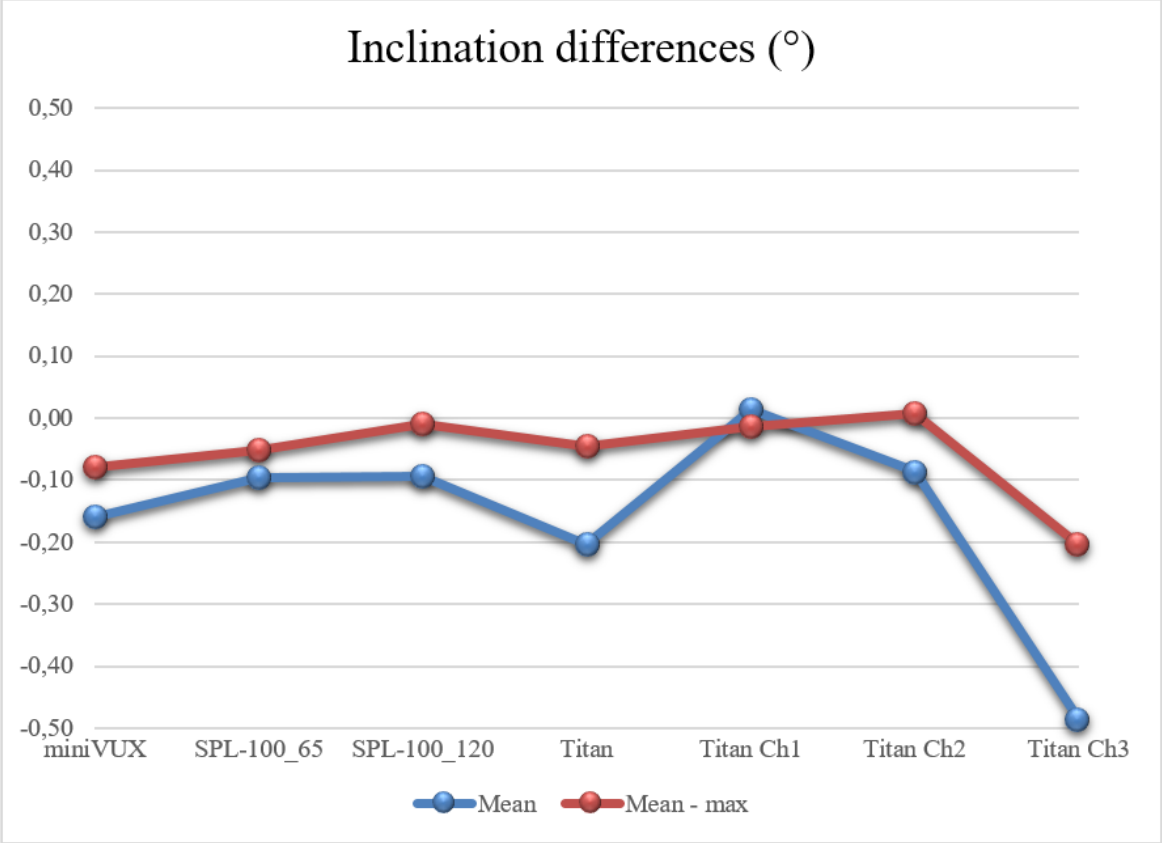


Figure 18. Inclination differences of the roofs compared to the VUX data. Blue line includes all the data and the max differences are removed from the data of the red line.

4.7 Plane deviations

The point deviations on the roof planes were investigated from the datasets. The roof point data was exported from TerraScan as LAS1.4 –files, which were used in CloudCompare. Flat and oblique roofs were used to investigate the dispersions on the whole surfaces of the roofs. Later the smaller samples were segmented, because the large deviations were noticed on the surfaces, which weren't caused by the scanner. For the conical roofs the roof planes were investigated separately. They were used to see if the flight direction, the location of the trajectory or the orientation and tilt of the roof plane cause any differences to the point dispersions. With the CloudCompare's fitting tool the planes can be created based on the point cloud. The tool works automatically when launched and creates a triangular mesh plane. In this work the dispersions of the roof planes are based on the cloud to mesh distances between the created roof planes and the point clouds. CloudCompare provides tools to examine visually the deviation to the plane (Fig. 19). When the plane is fitted to the point data, the software computes the root mean square –values to see some statistics of the disparity. Other information are normal vectors of the plane, dip and its direction, and the transformation matrix of 4x4 to make the plane horizontal, but these information weren't that important for the analysis. Because the software provides the RMS –values automatically they are used in this chapter to show the statistics together with the standard deviations. From the roofs including two or more sides or levels, as many planes as needed were created. For example,

for the conical roofs four planes were created based on the orientations of the roof surfaces. In this section the VUX data is not working as a reference, but is dealt individually together with the other datasets. From the VUX data the points of one minute of flight was separated beside the whole dataset, to see the possible differences of one scan (one flight line) and the combination of several scans (multiple flight lines). However, the one minute of flight data was used only for the sample investigation.

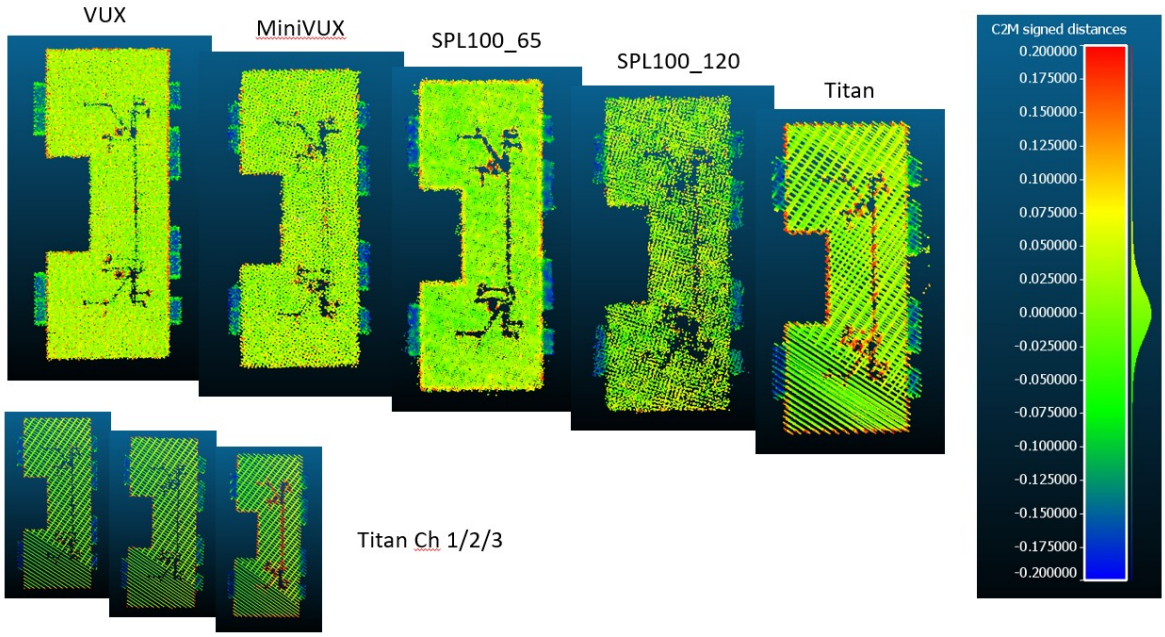


Figure 19. Computed and colored cloud-to-mesh distances. The planes concern the whole surface of the roof.

CloudCompare's cloud-to-mesh distance tool needs a point cloud and one or two meshes to get launched. In this case one mesh per point cloud is used as a reference. First the software calculates the approximate distances between the point cloud and the reference cloud. The plane mesh contains the hidden reference cloud. The approximate distances are displayed if the user wants to operate with the parameters, which is not usually necessary. The main parameters for the computation are octree level and maximum distance. The signed distance, flip normals and so called multi-threaded are the check box selections for the computation. The computation is performed in the subdivisions of octrees, where inside each cell the distance for each point is computed. The idea is that the octree level specifies the size of the cells to do as little computation as possible, and due to that the computation time will be lower. CloudCompare sets the number of the cells and cell size parameters automatically, which are recommended to use. The maximum distance determines how far the points are searched. With larger datasets than single roof points the distances between entities can be large and the search of their nearest neighbors can be slow. Signed distances determines, whether the distance should be signed with the triangle normal or not, and if it is signed, the flip normal inverts the sign. The multi-threaded selection decides if all available CPU cores of the computer are used for the processing [70].

4.7.1 Flat and oblique roofs

The mean RMS –values of each dataset for the total surfaces of the roofs were calculated from the individual planes. Notable were the size of the deviations and the much larger mean values of the flat roofs. With the attributes and accuracies of the equipment used in the data collection, the deviations were too large. The maximum disparities of the datasets varied between 24 and 76 cm for the flat roofs and between 14 and 20 cm for the oblique roofs. The largest objects on the roofs were cut out but the point clouds still had some smaller details, for example air conditioning units etc., which increased a bit the rate of the point deviation. The biggest problem with the flat roof category was, that at least the largest roofs, which were about 75 % of the investigated buildings, weren't totally flat. As is known also the flat roofs usually have some kind of inclination to guide and remove water, which were clearly seen, when the buildings were examined more precisely. This lead to the fact that a single planes didn't follow the exact shapes of the roofing materials. The largest roofs might also have some smaller varying on the face of the roof that affect to the point disparity. From the Figure 19 can be seen for example, that the edges of the roofs are a bit higher than the surfaces, the roofs of the balconies are a bit lower than the rest of the roof plane, and some details on the roof. From this building the pipes from the middle of the point cloud are removed.

Because the total surfaces had unsteadiness and didn't really explain the performance of the scanner, the smaller pieces were cut out from the roofs. The visual examples of the roof samples are in the Figure 20, where the samples are collected from the oblique roof of the car shelter. The size of the sample is about 40 m². Even if the range of the colored distances is the same (from -0.20 to 0.20 m) as previously in the Figure 19, the difference is that the average disparities are much smaller than with the larger surfaces. The mean values of the datasets are mostly under one centimeter and the RMS –values a couple of centimeters. From the point clouds it can be seen that the roof has elevated edges (at least the side on the top), where the distances to the plane are more than 15 cm. Besides the thinned VUX data in Figure 20, a sample with all the VUX points on the roof is shown (small image). This shows that the VUX data of one minute of flight is not as dense as the complete VUX dataset. Mentionable is that the shelter has wavy profile to remove the water and the VUX_1min data was the only dataset, where the profile of the roof was visible and the user was able to detect the single gutters of the roof. If the image (VUX_1min in the Fig. 20) is viewed carefully, the stripes in the point cloud can be seen. Some datasets have more points than the VUX_1min, but the detailed profile of the roof couldn't be recognized.

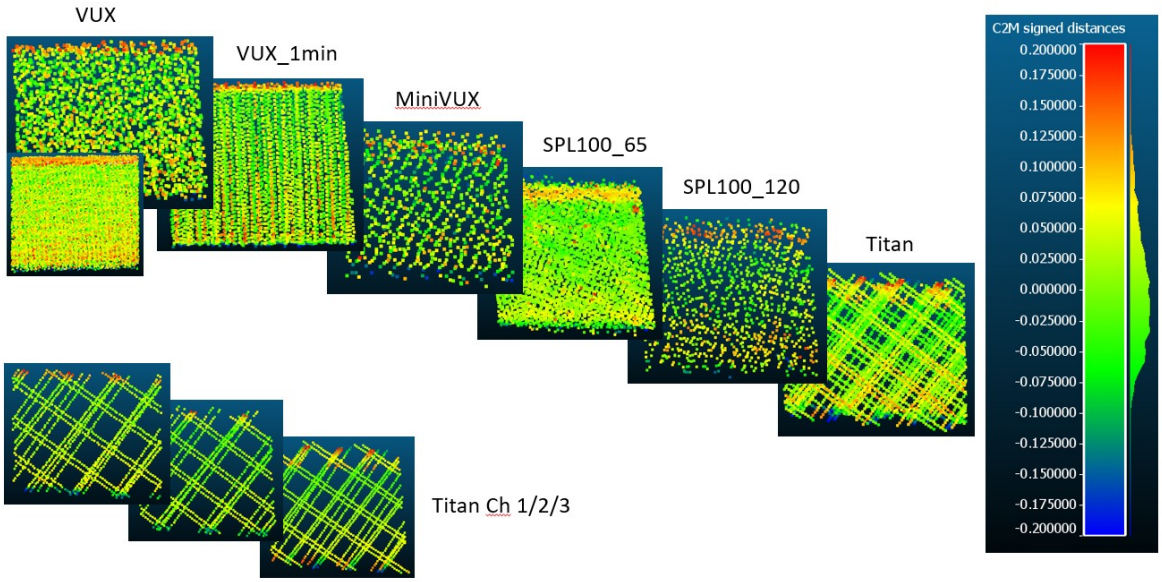


Figure 20. Computed and colored cloud-to-mesh distances. The planes concern the samples of the roof.

Figure 21 shows the mean RMS-values of the samples taken from the roof data. In overall the mean RMS-values have decreased a lot and the differences between the flat and oblique roofs are leveled. The earlier problem with the roof details and falls are not anymore involved, and can be assumed that now the results are more truthful concerning only the point deviations on a constant planar roofs. From the results it can be seen that in average the point deviation of the miniVUX data is the largest, which includes both of the roof types. The data of the lower altitude SPL100 seem to be a bit better than the higher altitude data. From the Titan data are seen that the channel 1 is the most precise and the channel 3 again the most imprecise. The VUX data of one minute of flight seems to be more accurate than the VUX data including several scans. This behavior was beforehand expected. However the differences between datasets are not dramatic. The maximum deviations varied between 6 and 10 cm for the flat roofs and between 5 and 15 cm for the oblique roofs. The Titan channel 3 had the max deviation of the flat roofs, 9.8 cm. For the oblique roofs the biggest deviation was in the miniVUX data, 14.5 cm.

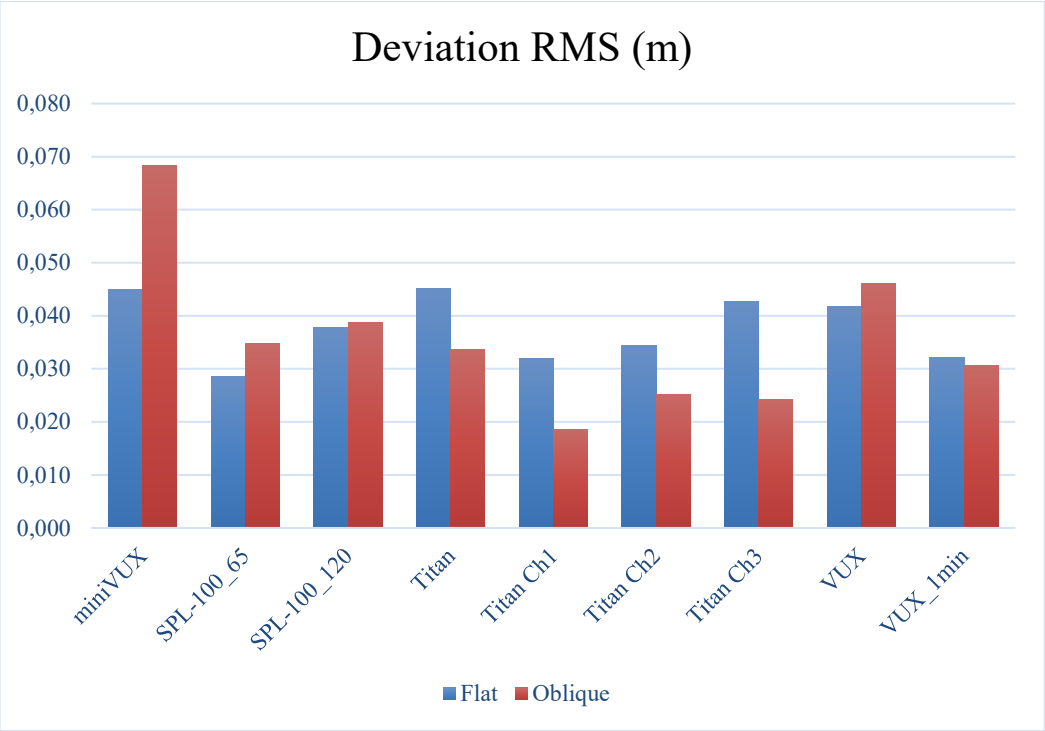


Figure 21. Mean RMS -values of the cloud-to-mesh distances. Planes concerned the samples from the roofs. Bars present the flat (blue) and oblique (red) roofs.

4.7.2 Conical roofs

Last thing of the point deviations was to investigate the conical roofs (Fig. 22). The idea was to see do the flight direction, the location of the trajectory, the orientation and tilt of the roof plane or the mechanism of the scanner cause any differences to the point deviations. To find out the effects, point cloud data of the different flight lines which consist the conical roofs were extracted: from the VUX data the point cloud data of one minute of flight, from the SPL100_120 data flight line 4, and from the Titan data the data of the flight lines 11 and 12. All the conical buildings had four differently oriented roof planes, which were handled individually, and whom were created the reference planes. From the outcome it could be seen that the miniVUX was the most diffuse data, where almost all of the RMS –values were close to 10 cm. Otherwise the deviations of the datasets were alternately between 2 and 10 cm. From the datasets, 15 outliers were found, and four of them were in the miniVUX data. The miniVUX and the SPL100_65 data had some roofs with noise below or on top of the plane. For the other datasets the outliers were in the individual roof planes and mostly with the low amount of points. From the Appendix 3 can be found an example data of the detached house, which were used in the investigation. The number of the points per each plane of the building and the distribution of the points are demonstrated. The planes are named with the compass directions. The vertical axis shows the number of the points and the horizontal axis the distance to the plane. The fitting curves of the normal (Gaussian) distribution are figured. The mean values and standard deviations of the datasets are also discovered from the figures.

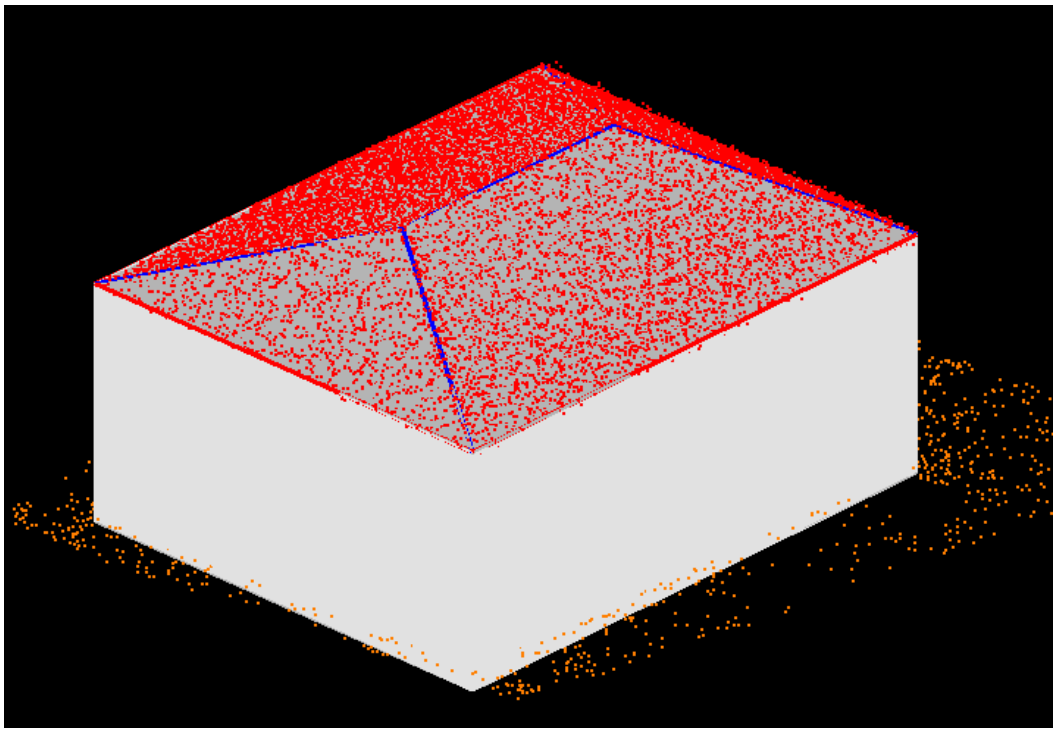


Figure 22. Building model with the conical roof. Outlines, ground and roof points are visible.

In overall the mean values, the standard deviations and the RMS –values of the roof planes are small in the VUX data, where most of the mean values are from -2 mm to 2 mm. In the example (Appendix 3), there are more points in the northern and southern roof planes, which come from the obvious reason that the surfaces are larger than the eastern and western planes. In the investigation was noticed that the number of the points decrease quite quickly together with the distance from the trajectory line. This was clearly seen at least from the data of one minute of flight. After 100 meters away from the scanning line the point density is much lower and the orientation of the roof plane starts to affect to the number of the hits. But again, it should be noted that the used VUX data was only a fifth from the whole dataset. From the point deviations it could be noticed that the sides, where the laser beam hits first and whose normals are towards the scanner, have the smaller dispersion than the further side, where the plane normals are pointing to the opposite direction. The RMS and standard deviations are couple of centimeters (mean RMS 6.8 cm). When the trajectory goes right above the building, the orientations of the planes don't affect to the point dispersion. Otherwise the location of the flight line has influence to the point deviations on the roof planes. It was expected that the VUX data of one minute of flight has smaller RMS (mean 3.5 cm) and standard deviation than the full VUX data, and from the example values it can be seen that the assumption was correct. Mentionable is that the VUX_1min data have more points than the thinned VUX data. When the target is further from the trajectory, the orientation of the plane (is it towards or against the scanner) has bigger influence to the number of the points, and in the investigated roofs the differences were even thousands or more points. However, the point deviations seem to be similar with high and low amount of points. The point deviations in respect to the tilt of the planes were investigated with a couple of buildings. The conical roofs had the inclinations between 102 and 110 degrees. No impact of the inclination change to the deviation values was found.

The number of the points in the miniVUX data was the lowest and it has the largest standard deviations and RMS –values from all the datasets. From the example figures in the Appendix 3, and from the other buildings too, two peaks formed to the graphs can be seen. Because the miniVUX data wasn't divided to the one minute of flight datasets, the presumption is that the shapes are coming from the different flight lines, which were done from the different directions (north-south and east-west directions). From the shapes and the values of the figures, it can be assumed that the north-south directed scans form the plane either a bit lower or higher, than the east-west directed scans. The second thing noticed from the data was that the orientations of the planes with respect to the flight direction affect more than in the VUX data. The planes pointing away from the scanner have less points than the planes toward the scanner. The RMS –values range between 3 and 12 (mean 9.8 cm) centimeters and the number of the points differ from 2 to 8 points/m². It is unknown how much the scanning direction against the flight direction have influence to the effect of the point layout. The last thing noticed was that the more inclined roof planes are more sensitive to the errors and changes than in the VUX data. When the tilt of the target grows, the laser footprint grows and the point cloud will be sparser, which seem to affect to the deviation.

The figures of the SPL100_65 data show that the point distributions are quite sharp and there are many points with the same distance from the roof plane. In the example the southern plane has more than thousand points with the same distance, which is much more than in the planes of the other datasets. In average the deviations are a bit smaller than in the VUX data, but there are single points, which have a longer distance to the roof plane (noise). The mean of the RMS is 5.4 cm. Notable was, that when the tilt of the plane is perpendicularly away from the trajectory there are systematically lower number of the points and larger disparity, as in the plane toward the scanner. The number of the points differ about 100 – 200 points per roof plane which is about 4 – 6 pts/m². The deviation differences of these planes were a couple of centimeters more than in the other planes. The roofs parallel to the flight direction don't have this kind of behavior either in the number of the points or in the point deviation. The flight line and the formed point pattern of the scanner seem to cause the differences to the roof planes. The SPL100_65 data was handled as one, because the buildings were captured reasonably enough only from the one flight line. The data from the other flight lines were too sparse or included holes.

The SPL100_120 data differs from the lower altitude SPL100 and Titan flights in a way, that it has one flight line crossing the study area from south to north. The effect of the crossed line is clearly seen in the number of the points, where the planes inside the swath and toward the cross section of the flight lines have much more points than the other roof planes of the same building. The difference between the roof towards the cross section and the opposite side can be even more than 500 points which is from 10 to 15 points/m². The cross section affects also to the other two roofs (on the sides along the trajectory), but the difference is not that obvious, and like in the example, the number of points can be quite nicely even. In average, the planes with the north-south directed normals have smaller point deviation, and at least on the plane, which is closer to the flight line. This behavior seems to be systematic, but the differences are not significant. In the east-west direction similar systematics can't be found. The mean of the RMS is 4.9 cm. From the SPL100_120 data the flight line four, which included almost all the investigated conical buildings, was separated. The amount of the point data is about half or one third from the whole SPL100_120 data. The distribution graphs are quite similar as with the total dataset, where the orientation of the roof planes

affects to the number of the points and the planes toward the flight line have more points. The mean of the RMS is 5.2 cm, and it seems that there is no systematic behavior in the point deviation.

Titan data was handled as one and the channels weren't separated. In overall it has small mean values, standard deviations and RMS –values. Almost all the data is from the flight line 11, which is flown from west to east. Seems that the flight direction and the location of the trajectory don't affect to the point deviation, but to the number of the points. The mean of the RMS is 4.1 cm, for the flight line 11 the RMS is the same 4.1 cm, and for the flight line 12, 3.6 cm. For the roof planes parallel or closer to the trajectory the number of the points is bigger. In the example, the difference is a couple of hundred points and similar behavior are seen from the other buildings and flight lines too. The differences are from 100 to 200 points on a plane which is about 1 – 5 points/m². One observed effect of the tilt to the point deviation, was with the inclination angle of 110 degrees, where the two planes, with normals pointing to the opposite direction from the scanner, have larger point deviations of a couple of centimeters. More buildings and wider coverage of inclinations should be investigated to be able to confirm the effect of the tilt to the point deviation.

5 Discussion

5.1 ALS data comparisons

This chapter summarized the results from the previous studies and the result of this project. In overall, the accuracies in this project are similar or a bit better than accuracies mentioned in the other studies. The horizontal accuracies of the conventional LiDAR works are mostly varying from 5 cm to one meter. In 2005 Kaartinen et.al. [35, 36] had the mean of the building outlines of 66 cm and the mean of the length accuracy of 93 cm for the laser scanner data. In 2015 Xiao et.al. [40] had the planimetric accuracy of 60 cm with the segmentation method, and in 2017 Cao et.al. [41], 10 cm with the roof reconstruction method. In this project were got better accuracies than the previous studies. The accuracies of the outlines for the conventional LiDAR (miniVUX) were clearly better than 10 cm and the mean difference of the corner points was 3.4 cm. Fernandez-Diaz et.al. [37] got the 5 cm accuracy with the multispectral Titan, and made the observation that the accuracy decreases, when all the channels are combined and the channels are recommended to be used separately. Similar issues were recognized in this project. The corner coordinates of the Titan reached the 5 cm accuracy and the outline corners were almost the same. When the roof outlines were investigated, it was found that the channel 3 (532 nm) causes the most of the error. Channels 1 and 2 reach the 10 cm accuracy, but the channel 3 extends the outlines. Based on this study the horizontal accuracy of 5 cm can be reached, but the behavior of the channel 3 is unreliable. Degnan [29] reported the accuracy of 10 cm with the SPL systems, and in this project the investigation of the corner points and corner outlines reached that easily. However, the differences of the surfaces of the SPL100 are quite large, so there seem to be some unreliability with the green wavelength.

Kaartinen et.al. [35, 36] were investigating the building heights for the conventional LiDAR data and the mean of 32 cm was reported. Xiao et.al. [40] and Cao et.al. [41] reached the 10 cm vertical accuracy for the building roofs. In this work, the miniVUX reached the height difference of 5.9 cm to the reference, which is better than in the other researches. With the Titan was reached 6.5 cm accuracy by Fernandez-Diaz et.al. [37]. In this project, the mean heights of the Titan were between 6.3 and 8.4 cm, which are close the previous studies. Degnan [29] got 3 cm vertical accuracy with the SPL data and Stoker et.al. [58] 17 cm accuracy. The vertical accuracy of 3 cm wasn't reached with the SPL100 data, but quite close, when the mean differences were 5.9 and 4.2 cm to the reference. Kaartinen et.al. [35, 36] also investigated the inclinations and the mean of 2.7° was reported. In this project the differences were small and the mean values of all the datasets were less than 0.5°. Plane deviations of 0.7 – 3.1 cm for the conventional LiDAR data and 1.0 – 9.5 cm for the SPL data were reported by Mandlburger et.al. [31]. In overall the deviations of the VUX (4.4 cm) and miniVUX (5.7 cm) datasets were a bit more than the conventional LiDAR results in the paper by Mandlburger, and the point deviations of the SPL100 datasets were close (SPL100_65: 3.2 cm, SPL100_120: 3.8 cm).

Table 5 is the overview of the outcome, which includes the most and least accurate results for each comparison. The datasets and outcome are implicated. "Surface areas" shows the mean differences in percentages compared to the VUX data, and they are divided by the size of the buildings. Roof outlines had two classes: the outlines less than 40 m and more than

40 m. Planimetric coordinate differences were investigated from the corner points and corners of the outlines as were presented in the paper. Roof inclinations had also datasets when the outliers (max differences) were removed. Investigation for plane deviations were done for the flat and oblique roofs separately.

	Best result	**Worst result**
Point densities (pts/m²)		
	SPL100_65 80,5	Titan Ch3 16,9
Surface areas (%)		
Buildings < 150 m²	miniVUX/Titan Ch1 -1,4	Titan Ch3 7,2
Buildings 150 - 400 m²	Titan Ch2 -0,4	Titan Ch3 6,2
Buildings > 400 m²	Titan Ch1 0,0	Titan Ch3 3,6
Roof outlines (cm)		
Outlines < 40 m	miniVUX 0,4	Titan Ch3 -49,3
Outlines > 40 m	miniVUX 1,5	Titan Ch3 -53,9
Planimetric coordinate differences (cm)		
Corner points	Titan Ch3 1,4	miniVUX/SPL100_120 4,3
Outline corners	Titan Ch3 1,6	Titan Ch1 6,4
Building heights (cm)		
	SPL100_120 4,2	Titan Ch2 8,4
Roof inclinations (°)		
All data	Titan Ch1 0,01	Titan Ch3 -0,49
Max values removed	Titan Ch1/Titan Ch2/SPL100_120 0,01	Titan Ch3 -0,20
Plane deviations (cm)		
Flat roofs	SPL100_65 2,8	miniVUX/Titan 4,5
Oblique roofs	Titan Ch1 1,9	miniVUX 6,8

Table 5. The overview of the most and the least accurate results of each comparison.

5.2 Limitations and recommendations

As was seen, the features of the laser scanners differ, which have influence to the quality and attributes of the data. The capabilities and limits of the scanners are somehow known, but this type of accuracy investigation and data comparison including the new mechanisms of multispectral or single photon LiDARs, were found just few or none at all. Different algorithms and papers for point cloud classifications were found, but the number of the accuracy investigation of the point clouds or 3D models was small. This can be seen kind of a limit also for the project, that the knowledge of the capabilities of the data, including example situations and analysis, were so small. On the other hand, or actually, that was the idea of the thesis to investigate the new technologies and compare those to the more traditional and high accurate ALS laser scanners. One quite big limitation in the work was concerning the VUX-1HA data, technically the number of used points. For the more fluent processing the amount of the data had to be reduced. Finally, the VUX data was just fifth from the total amount of the dataset. The amount of the data could have been a bit more to make the processing still practicable, but however the processing took time, and it was decided to reduce the data quite much to make the processing adaptive. It can be assumed that this may affect to the accuracy and precision of the data, but the average difference between the measured points and VUX data was 2 cm in x-coordinate and 6 cm in y-coordinate. ALS campaigns and RTK –measurements also have some uncertainties and errors in the navigation and positioning, but the impact was thought to be max 5 cm. Preprocessing of the data and point cloud adjustment may also include some error. In this thesis, they cannot be analyzed, as they were done by the third party companies, who were responsible for the ALS campaigns and the selection of the preprocessing methods. The accuracies of the scanners and the datasets were presented in the text, and here are disclosed some possible error sources and other observations, which were found or might be involved in the data processing and calculations of the project. The systematic errors in the SPL100 datasets were taken into account only, when the coordinates of the corner points of the roofs were calculated, because it shouldn't affect to the other investigated problems. The question is that were the number of the points enough to remove the error reliably and was it smart to use only the roof points to remove the error. If more points including ground points would have been used and the removal of the systematic errors had been done for the complete datasets, how much the accuracy and reliability would have improved. One thought was to investigate the ground points before the roofs, but it was left out because the time set the limit for the project and the object detection based on the intensity, without the aerial images, was challenging.

The classifications of the point clouds were one way to increase the reliability of the results in the project. To make the point cloud classification trustworthy, the TerraScan macros were tested step by step for each dataset, and the automatic classification was done twice. The roof points of the sample areas were inspected, and manually edited when needed to ensure that all roof points were in the sample datasets. The reliability of the 3D models was a bit unsure, because many of them needed editing after the vectorization. The simple flat, oblique and gabled buildings were reconstructed correctly, but for the large sized buildings and complex structures it was difficult to find correct parameters for the vectorization. Also the smaller details on the roofs confused the construction. For these buildings manual editing was executed, following the point clouds as good as possible, but still there might be small differences between the models. However, it must be remembered that the point clouds also differ from each other, which might bring some variance to the reconstruction. Some of the buildings needed to be left out because the TerraScan's building reconstruction wasn't able

to create the complex building models. The number of the details seem to be one reason for the problems in the vectorization. For the more detailed models, another software could be tested to create the higher LoD models, where the smaller details can be created. As was seen the surfaces, outlines and roof inclinations were based on the 3D models, and the accuracies of the reconstructions were crucial. If the work and comparison of the datasets would continue, the next step could be to use the modeling method for the higher level of details.

It was already mentioned that some of the research projects used the hybrid method, where the laser scanner data is used together with the images [35, 36, 38]. Most of them reported that the hybrid methods provide higher accuracy and affect positively to the 3D reconstruction, than using only the laser scanner data. The hybrid of singe photon or multispectral LiDAR data together with the images, is one method which no-one has published yet. The hybrid method might help in the data processing and it is assumed, that the hybrid with the single photon or multispectral data, works similarly, as with the conventional LiDAR data, and increases the accuracy. Again, if the work would continue this can be step to try and it probably will affect to the results. The data used in this project included also features or information, which weren't taken into account, but might be useful at least in the point cloud classification. The waveform and echo information of each point can be handled and used also in TerraScan [69]. The waveform data can be linked to the laser points and trajectories, and used in the classification. The echoes can be extracted (first, last, all echoes) from the data and the properties of the length, normality and position, can be compared between the points to ease and clear the data processing. The waveform profile shows for each point the number of the photons returned to the scanner and the location of the selected point. When comparing the waveform data of the point to the typical graphs of the land cover classes, the target detection might be more reliable. With the echo extraction, points to the missing returns can be created, where the generated points are not optimal. The method is used for example, when there are missing returns from the objects or to create the ground points in a densely vegetated areas by using the order of the echoes. Titan's multispectral data provides the possibility to detect the ground classes and objects from the spectrum of all the three wavelengths. In this case, only the roofs, where the number of the samples were quite small were investigated, and the manual work was practicable. This work didn't need the spectral information to find the points belonging to the roof features. With the larger amount of buildings, or if more than roof data is studied, the automatic classification and object recognition might get better results, if the echo or spectral data is involved in the processing.

When building investigation is done, the next thing is to include also the wall information to the comparison. Depending on the scanning geometry, the high density ALS point clouds may have quite many points returned from the walls as well, and the research could be extended to cover the entire buildings. With these datasets, the wall investigation would be possible, even if the higher altitude flights SPL100 and Titan have some missing surfaces and a bit lower point density on the walls. The lower altitude and wide field-of-view VUX and miniVUX datasets have scanned the walls excellently. The similar research as for the roof data can be done. It can be hypothesized that the differences are even larger between the datasets when the vertical walls are investigated. As an assumption, when the footprints of the laser beams are larger, the point densities of the datasets and the distributions between the walls are more varied. If more roof data is wanted to be studied, the details can be added to the investigation. It was thought to include the detail detection to this work, but eventually it was left out. The roof objects can be used to investigate the effect of the dataset to the

dimensions and to the 3D reconstruction. It would be helpful for the future research to examine the accuracies, and to get some knowledge about the processing of the smaller details from the ALS data. Also the amount of the buildings could be increased, and, what was obtained less emphasis in this work, to search the different shapes of the roofs for the surface and dispersion investigation. Round shapes, for example domes, were tried to find from the roofs of the area, but none was found.

6 Conclusion

This thesis presented the airborne laser scanning data comparison based on building roof features in an urban area. In the data collection the state-of-the-art technologies were used, multispectral and single photon LiDARs, together with the conventional LiDARs. The lack of accuracy investigations and data comparisons were the main reasons for initiation of the project. Another motivation was the notably low number of publications on urban area features. The work presented the background and the principles of laser scanning, technologies and new methods for the data capturing. Also comparison of the equipment can be found from the study. The preprocessing wasn't part of the project, but otherwise the relevant data processing methods and software were explained.

The high density point clouds allow more detailed data and as such ease the object recognition. The datasets used in the project were dense, the most sparse point density was 16 pts/m². However, all the benefits of the very dense point clouds can't always be used. In this case the total dataset of the very high density VUX-1HA couldn't be used in the processing because the lack of computing power. The VUX-1HA data was used as a reference, and the point cloud was reduced to one fifth from the original dataset. The assumption was that the reduction affects the accuracy about a couple of centimeters. In overall the scanners investigated in this project reach a few centimeters accuracy.

In the surface area investigation the smaller surfaces have more uncertainties than larger ones because the sparser point clouds and possible occluded areas in the roof data. The miniVUX and Titan channels 1 and 2 seem to be the most accurate, where the differences were less than two percent from the reference. The SPL100 and Titan channel 3 use the visible (green) wavelength, which seem to widen the roof planes because of the larger laser beam divergence compared to the others. A bit surprising was that the higher altitude SPL data (3600 m) provided better accuracy than the lower altitude SPL data (1920 m). The seemingly greater amount of noise points in the lower altitude scan would explain the phenomenon.

For the roof outline investigation the results were more or less identical with the surface calculations. The miniVUX and the two infrared channels of Titan are the most accurate (< 10 cm), and the larger laser beam divergence of the green wavelength affects clearly to the lengths and widths of the buildings, where the mean differences were from 30 to 60 cm to the reference.

The planimetric corner coordinate differences were mostly between 2 and 10 cm. The differences between the datasets are not dramatic, and as expected the small variance is involved when either coordinate points or the coordinates of the outlines of the 3D models was used. Titan channel 3 seems to be one of the most accurate data when the coordinates of the corner points were investigated. This is a bit contradictory in reflection to the results for roof outlines and surfaces. The building heights differed between 4 and 8 cm, and the higher altitude SPL100 data was the most accurate based on the calculations. On average the roof inclination angles didn't show large variance. The most differences to the inclinations were observed in the Titan channel 3, but the mean differences were still less than one degree.

When the plane deviations are investigated the smaller samples of roof planes should be used instead of the total surfaces. Then the problems with the roof details and non-planar shapes

of the roofing material are largely eliminated. The deviations were mostly between 2 and 5 cm. The miniVUX-1UAV data had the largest disparity on a plane (4 – 7 cm), where the point density might have an effect. The VUX-1HA data of one minute of flight provides the possibility to detect the shapes of the roof material. The wavy profiles of the roof plates and tiles were detectable from the VUX-1HA data, but not with the other datasets. The flight trajectory as well as tilt and position of the roof affect the resulting point cloud properties from different scanners.

The results obtained in this study show that a few centimeter accuracy of the large area ALS is achievable. More accuracy and comparison research is needed for the state-of-the-art scanning systems and this study was one of the opening investigations for the urban areas. Based on the results the low altitude miniVUX-1UAV provides competitive accuracy compared to the denser SPL100 and Titan datasets. Overall, the higher altitude SPL100 seems to provide better results than the lower flight, but more investigation is needed to better understand the behavior of the instrument. With Titan, the three channels are recommended to be handled not only together, but also separately. Channel 1 of Titan (1550 nm) seems to provide the best results from the three channels. The visible (green) wavelength of the SPL100 and Titan's channel 3 have larger laser beam divergences, which might affect to the accuracies. Finally, different scenes, targets and volumes should be investigated to acquire more information about the performance of the new technology scanners in the perspective of topographic mapping, and to provide more knowledge for the future sensor and methodological development.

References

[1] Jurovich Surveying (2019), Short History of Land Surveying, website. Available online [https://www.jurovichsurveying.com.au/faq/history-land-surveying]. Accessed on September 13, 2019.

[2] Harley, J.B. (1987), The History of Cartography, Volume 1. Cartography in Prehistoric, Ancient, and Medieval Europe and the Mediterranean. Chapter 1, The Map and the Development of the History of Cartography. p. 1-5.

[3] Ropponen, J-M. (2013), Kartografia Suomessa. Opinnäytetyö. Rovaniemi University of Applied Sciences. Theseus. Available online [https://www.theseus.fi/bitstream/handle/10024/60971/Ropponen_Jouni-Matti.pdf?sequence=1&isAllowed=y]. Accessed on September 13, 2019.

[4] Environmental Science (2014), Cartography: More Than a View From Above, website. Available online [https://www.environmentalscience.org/cartography]. Accessed on September 16, 2019.

[5] Petrie, G. and Toth, C.K., Introduction to Laser Ranging, Profiling, and Scanning. In the book: Shan J. and Toth C.K., Topographic Laser Ranging and Scanning: Principles and Processing. Boca Raton, USA. CRC Press, Taylor & Francis Group. p. 1-27. ISBN-13: 978-1-4200-5142-1.

[6] Flood, M. (2001), Laser Altimetry: From Science to Commercial Lidar Mapping. Photogrammetric Engineering and Remote Sensing. p. 1209-1217.

[7] Jaakkola, A. et.al. (2010), A Low-Cost Multi-Sensoral Mobile Mapping System and Its Feasibility for Tree Measurements. ISPRS Journal of Photogrammetry and Remote Sensing. Volume 65(6). p. 514-522. DOI: [10.1016/j.isprsjprs.2010.08.002].

[8] Titan (2019), Teledyne Optech, website. Available online [https://www.teledyneoptech.com/en/products/airborne-survey/titan/]. Accessed on September 24, 2019.

[9] The evolution of LiDAR (2019), Hexagon Geosystems, website. Available online [https://rail.hexagongeosystems.com/news/customer-magazine/reporter-80/the-evolution-of-lidar]. Accessed on September 24, 2019.

[10] Alho, P. et.al. (2008), Application of boat-based laser scanning for river survey. Earth Surface Processes and Landforms 34. p. 1831-1838 (2009). DOI: [10.1002/esp.1879].

[11] European Global Navigation Satellite Systems Agency (2017), What is GNSS?, website. Available online [https://www.gsa.europa.eu/european-gnss/what-gnss]. Accessed on December 10, 2019.

[12] Vosselman, G. and Maas, H-G. (2010), Airborne and terrestrial laser scanning, E-book. Whittles Publishing, Scotland. ProQuest Ebook Central, [https://ebookcentral.proquest.com/lib/aalto-ebooks/detail.action?docID=3417283].

[13] Kukko, A. et.al. (2013), Mobile Laser Scanning – System development, performance and applications. Finnish Geodetic Institute. Aalto University, Doctoral dissertation, Real Estate, Planning and Geoinformatics. ISBN (pdf): 978-951-711-307-6.

[14] Kouva, V. (2017), Ilmalaserkeilainten kehitys. Opinnäytetyö, Lapland University of Applied Sciences. Theseus. Available online [https://www.theseus.fi/bitstream/handle/10024/138649/Ville_Kouva.pdf?sequence=1&isAllowed=y]. Accessed on September 17, 2019.

[15] Wehr, A., LiDAR Systems and Calibration. In the book: Shan J. and Toth C.K., Topographic Laser Ranging and Scanning: Principles and Processing. Boca Raton, USA. CRC Press, Taylor & Francis Group. p. 129-172. ISBN-13: 978-1-4200-5142-1.

[16] Geospatial Modeling and Visualization (2019), 3D Scanning, Scanning Overview, Airborne Laser Scanning, website. Available online [http://gmv.cast.uark.edu/]. Accessed on September 18, 2019.

[17] Mandlburger G. (2019), Recent developments in airborne lidar, GIM International, website. Article, Available online [https://www.gim-international.com/content/article/recent-developments-in-airborne-lidar-2]. Accessed on September 20, 2019.

[18] Ullrich, A. et.al. (2013), How to read your LIDAR spec – a comparison of single-laser-output and multi-laser-output LIDAR instruments, Riegl Laser Measurement Systems GmbH. Available online [http://www.geocenter.fi/NGC/wp-content/uploads/2015/03/Airborne_Laser_Scanning-what_to_expect_from_your_sensor_2013-03-08.pdf]. Accessed on October 4, 2019.

[19] Toth, C.K. (2002), CALIBRATING AIRBORNE LIDAR SYSTEMS, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume 34. p. 475-480.

[20] Cheng, L. et.al. (2018), Registration of Laser Scanning Point Clouds: A Review. Basel, Switzerland. Sensors, Volume 18(5), 1641. DOI: 10.3390/s18051641.

[21] Shan J. and Toth C.K. (2009), Topographic Laser Ranging and Scanning: Principles and Processing. Boca Raton, USA. CRC Press, Taylor & Francis Group. ISBN-13: 978-1-4200-5142-1.

[22] Baltsavias, E.P. (1999), Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry & Remote Sensing, Volume 54. p. 199-214. DOI: [https://doi.org/10.1016/S0924-2716(99)00015-5].

[23] Juola, J. (2019), Multi-angular measurement of woody tree structures with mobile hyperspectral camera. Master's thesis. Aalto University. p. 12. Available online in Aaltodoc [http://urn.fi/URN:NBN:fi:aalto-201908254904]. Accessed on September 20, 2019.

[24] NASA (2013), Imagine the Universe, website. Available online [https://imagine.gsfc.nasa.gov/science/toolbox/emspectrum1.html]. Accessed on September 20, 2019.

[25] Bakula, K. (2015), Multispectral airborne laser scanning- a new trend in the development of LiDAR technology. Archiwum Fotogrametrii, Kartografii i Teledetekcji, vol. 27. p. 25-44. ISSN 2083-2214, eISSN 2391-9477. DOI: [10.14681/afkit.2015.002]

[26] Sirota, M. and Roth, R. (2017), The Evolution of Lidar, GIM International, website. Article, Available online [https://www.gim-international.com/content/article/the-evolution-of-lidar]. Accessed on September 24, 2019.

[27] Degnan, J. et.al. (2015), Scanning, Hundred Beam, 3D Imaging Lidars Operating at Megapixels per Second. LM3D.4. Imaging and Applied Optics. DOI: [10.1364/LSC.2015.LM3D.4].

[28] Li, Q. et.al. (2016), First Evaluation on Single Photon-Sensitive Lidar Data. Photogrammetric Engineering & Remote Sensing, Volume 82. p. 455-463. DOI: [10.14358/PERS.82.7.455].

[29] Degnan, J. (2016), Scanning, Multibeam, Single Photon Lidars for Rapid, Large Scale, High Resolution, Topographic and Bathymetric Mapping. Remote Sensing, Volume 8, 958. DOI: [10.3390/rs8110958].

[30] Jutzi, B. (2017), Less Photons for More LiDAR? A Review from Multi-Photon Detection to Single Photon Detection. 56th Photogrammetric Week (PhoWo) 2017.

[31] Mandlburger, G. et.al. (2019), A Comparison of Single Photon and Full Waveform LiDAR. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume IV-2/W5. p. 397-404. DOI: [https://doi.org/10.5194/isprs-annals-IV-2-W5-397-2019].

[32] Swatantran, A. et.al. (2016), Rapid, High-Resolution Forest Structure and Terrain Mapping over Large Areas using Single Photon Lidar. Scientific Reports, Volume 6. Article 28277. DOI: [10.1038/srep28277].

[33] Khan Academy (2019), Population and sample standard deviation review, website. Available online [https://www.khanacademy.org/math/statistics-probability/summarizing-quantitative-data/variance-standard-deviation-sample/a/population-and-sample-standard-deviation-review]. Accessed on October 30, 2019.

[34] Ajmera, A. (2000), Mathematical Formulas, Coordinate Geometry, website. Available online [https://orion.math.iastate.edu/dept/links/mathforms.html]. Accessed on November 4, 2019.

[35] Kaartinen, H. et.al., (2005), EuroSDR building extraction comparison. ISPRS Hannover Workshop High-Resolution Earth Imaging for Geospatial Infromation. Available online [https://www.researchgate.net/publication/228804651_EuroSDR_building_extraction_comparison].

[36] Kaartinen, H. et.al., (2005), Accuracy of 3D city models: EuroSDR comparison. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences.

Volume 36(3/W19). Available online [https://www.researchgate.net/publication/228820122_Accuracy_of_3D_city_models_EuroSDR_comparison].

[37] Fernandez-Diaz, J.C. et.al. (2016), Capability Assessment and Performance Metrics for the Titan Multispectral Mapping Lidar. Remote Sensing, Volume 8. Article 936. DOI: [10.3390/rs8110936].

[38] Awrangjeb, M. et.al. (2018), An Effective Data-Driven Method for 3-D Building Roof Reconstruction and Robust Change Detection. Remote Sensing. Volume 10. Article 1512. DOI: [https://doi.org/10.3390/rs10101512].

[39] Gülch, E. et.al., Quality of Buildings Extracted from Airborne Laser Scanning Data: Results of an Empirical Investigation on 3D Building Reconstruction. In the book: Shan J. and Toth C.K., Topographic Laser Ranging and Scanning: Principles and Processing. Boca Raton, USA. CRC Press, Taylor & Francis Group. p. 535-573. ISBN-13: 978-1-4200-5142-1.

[40] Xiao, Y. et.al. (2015), Building segmentation and modeling from airborne LiDAR data. International Journal of Digital Earth, 8:9. p. 694-709. DOI: [10.1080/17538947.2014.914252].

[41] Cao, R. et.al. (2017), 3D building roof reconstruction from airborne LiDAR point clouds: a framework based on a spatial database. International Journal of Geographical Information Science. p. 1-22. DOI: [10.1080/13658816.2017.1301456].

[42] Jung, J. and Sohn, G. (2019), A line-based progressive refinement of 3D rooftop models using airborne LiDAR data with single view imagery. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 149. p. 157-175. DOI: [10.1016/j.isprsjprs.2019.01.003].

[43] Park, S-Y. et.al. (2019), Segmentation of LiDAR Data Using Multilevel Cube Code. Journal of Sensors, Volume 2019. Article 4098413. DOI: [https://doi.org/10.1155/2019/4098413].

[44] Biljecki, F. et.al. (2017), Generating 3D city models without elevation data. Computers Environment and Urban Systems, Volume 64. p. 1-18. DOI: [10.1016/j.compenvurbsys.2017.01.001].

[45] Demir, N. and Baltsavias, E. (2012), Automated modeling of 3d building roofs using image and LIDAR data. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume I-4. p. 35-40. DOI: [10.5194/isprsannals-I-4-35-2012].

[46] Zheng, Y. et.al. (2017), A Hybrid Approach for Three-Dimensional Building Reconstruction in Indianapolis from LiDAR Data. Remote Sensing, Volume 9. Article 310. DOI: [10.3390/rs9040310].

[47] He, M. et.al. (2016), A 3D Shape Descriptor Based on Contour Clusters for Damaged Roof Detection Using Airborne LiDAR Point Clouds. Remote Sensing, Volume 8(3). Article 189. DOI: [https://doi.org/10.3390/rs8030189].

[48] Lingfors, D. et.al. (2017), Comparing the capability of low-and high-resolution LiDAR data with application to solar resource assessment, roof type classification and shading analysis. Applied Energy, Volume 205. p. 1216-1230. DOI: [10.1016/j.apenergy.2017.08.045].

[49] Karila, K. et.al. (2017), Feasibility of Multispectral Airborne Laser Scanning Data for Road Mapping. IEEE Geoscience and Remote Sensing Letters, Volume 14, No 3. p. 294-298. DOI: [10.1109/LGRS.2016.2631261].

[50] Matikainen, L. et.al. (2017), Object-based analysis of multispectral airborne laser scanner data for land cover classification and map updating. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 128. p. 298-313. DOI: [https://doi.org/10.1016/j.isprsjprs.2017.04.005].

[51] Zou, X. et.al. (2016), 3D LAND COVER CLASSIFICATION BASED ON MULTISPECTRAL LIDAR POINT CLOUDS. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.. Volume XLI-B1, 741-747. DOI: [https://doi.org/10.5194/isprs-archives-XLI-B1-741-2016].

[52] Morsy, S. et.al. (2017), Multispectral LiDAR Data for Land Cover Classification of Urban Areas. Sensors, Volume 17(5). Article 958. DOI: [https://doi.org/10.3390/s17050958].

[53] Teo, T-A. and Wu, H-M. (2017), Analysis of Land Cover Classification Using Multi-Wavelength LiDAR System. Applied Sciences, Volume 7. Article 663. DOI: [10.3390/app7070663].

[54] Morsy, S. et.al. (2016), Airborne multispectral lidar data for land-cover classification and land/water mapping using different spectral indexes. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-3. p. 217-224. DOI: [10.5194/isprsannals-III-3-217-2016].

[55] Bakula, K. et.al. (2016), TESTING OF LAND COVER CLASSIFICATION FROM MULTISPECTRAL AIRBORNE LASER SCANNING DATA. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., Volume XLI-B7. p. 161-169. DOI: [https://doi.org/10.5194/isprs-archives-XLI-B7-161-2016].

[56] Huo, L-Z. et.al. (2018), Supervised spatial classification of multispectral LiDAR data in urban areas. PLoS ONE 13(10): e0206185. DOI: [https://doi.org/10.1371/journal.pone.0206185].

[57] Wästlund, A. et.al. (2018), Forest Variable Estimation Using a High Altitude Single Photon Lidar System. Remote Sensing, Volume 10(9). Article 1422. DOI: [https://doi.org/10.3390/rs10091422].

[58] Stoker, J.M. et.al. (2016), Evaluation of Single Photon and Geiger Mode Lidar for the 3D Elevation Program. Remote Sensing, Volume 8. Article 767. DOI: [10.3390/rs8090767].

[59] Mandlburger, G. and Jutzi, B. (2019), On the Feasibility of Water Surface Mapping with Single Photon LiDAR. ISPRS International Journal of Geo-Information, Volume 8(4). Article 188. DOI: [https://doi.org/10.3390/ijgi8040188].

[60] Google Maps (2019), Google. Available online [https://www.google.fi/maps]. Accessed on October 1, 2019.

[61] Riegl VUX-1HA (2015), Riegl Laser Measurement Systems, website. Product info. Available online [http://www.riegl.com/products/newriegl-vux-1-series/newriegl-vux-1ha/]. Accessed on October 2, 2019.

[62] Riegl miniVUX-1UAV (2016), Riegl Laser Measurement Systems, website. Product info. Available online [http://www.riegl.com/products/unmanned-scanning/riegl-minivux-1uav/]. Accessed on October 2, 2019.

[63] Leica SPL100 (2017), Leica Geosystems, Hexagon Geosystems, website. Product info. Available online [https://leica-geosystems.com/fi-fi/products/airborne-systems/topographic-lidar-sensors/leica-spl100]. Accessed on October 2, 2019.

[64] HiPer HR, GNSS receiver, Topcon, website. Product info. Available online [https://www.topconpositioning.com/gnss-and-network-solutions/integrated-gnss-receivers/hiper-hr#panel-product-specifications]. Accessed on October 22, 2019.

[65] Trimnet VRS, Geotrim Service, website. Available online [https://geotrim.fi/palvelut/trimnet-vrs/]. Accessed on December 16, 2019.

[66] TerraScan, Terrasolid, website. Product info. Available online [http://www.terrasolid.com/products/terrascanpage.php]. Accessed on October 22, 2019.

[67] MicroStation, Bentley Systems, website. Product info. Available online [https://www.bentley.com/en/products/product-line/modeling-and-visualization-software/microstation]. Accessed on October 22, 2019.

[68] CloudCompare, Open Source Project, website. Available online [https://www.danielgm.net/cc/]. Accessed on August 29, 2019.

[69] Soininen, A. (2000 - 2016), TerraScan User's Guide, Terrasolid. 32-bit version of TerraScan. Available online [https://www.terrasolid.com/download/tscan.pdf]. Accessed on October 23, 2019.

[70] CloudCompare User Manual (2015), English Version 2.6.1. Available online [https://www.danielgm.net/cc/doc/qCC/CloudCompare%20v2.6.1%20-%20User%20manual.pdf]. Accessed on November 8, 2019.

Appendix

Appendix 1. Mean, minimum and maximum differences, and standard deviations in surface calculation. 2 pages.
Appendix 2. Outline figures of the lengths and widths of the buildings. 2 pages.
Appendix 3. Data deviations of the conical roof of the detached house from the block 4. 4 pages.

Appendix 1. Mean, minimum and maximum differences, and standard deviations in surface calculation

Buildings under 150 m²

MiniVUX	Mean	1,56	Min	0,25
	Std	3,23	Max	6,36
SPL-100_65	Mean	-7,36	Min	1,50
	Std	5,44	Max	19,15
SPL-100_120	Mean	-5,72	Min	0,70
	Std	4,40	Max	12,25
Titan	Mean	-7,34	Min	0,99
	Std	6,30	Max	15,93
Titan Ch1	Mean	1,57	Min	0,20
	Std	2,98	Max	5,57
Titan Ch2	Mean	-2,21	Min	0,10
	Std	3,83	Max	7,11
Titan Ch3	Mean	-9,09	Min	2,81
	Std	6,01	Max	19,98

Buildings between 150 - 400 m2

MiniVUX	Mean	-0,76	Min	1,23
	Std	9,40	Max	20,30
SPL-100_65	Mean	-12,29	Min	1,07
	Std	9,78	Max	25,98
SPL-100_120	Mean	-8,76	Min	0,41
	Std	10,53	Max	35,60
Titan	Mean	-11,00	Min	1,59
	Std	10,35	Max	32,24
Titan Ch1	Mean	3,27	Min	0,55
	Std	6,44	Max	10,12
Titan Ch2	Mean	0,29	Min	0,48
	Std	6,66	Max	11,03
Titan Ch3	Mean	-17,70	Min	0,75
	Std	15,31	Max	53,12

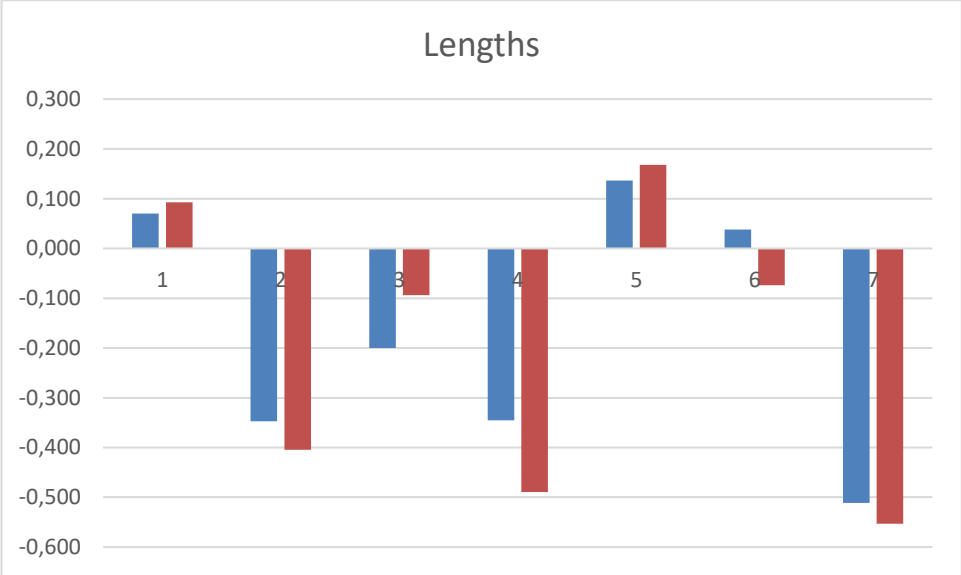
Buildings over 400 m²

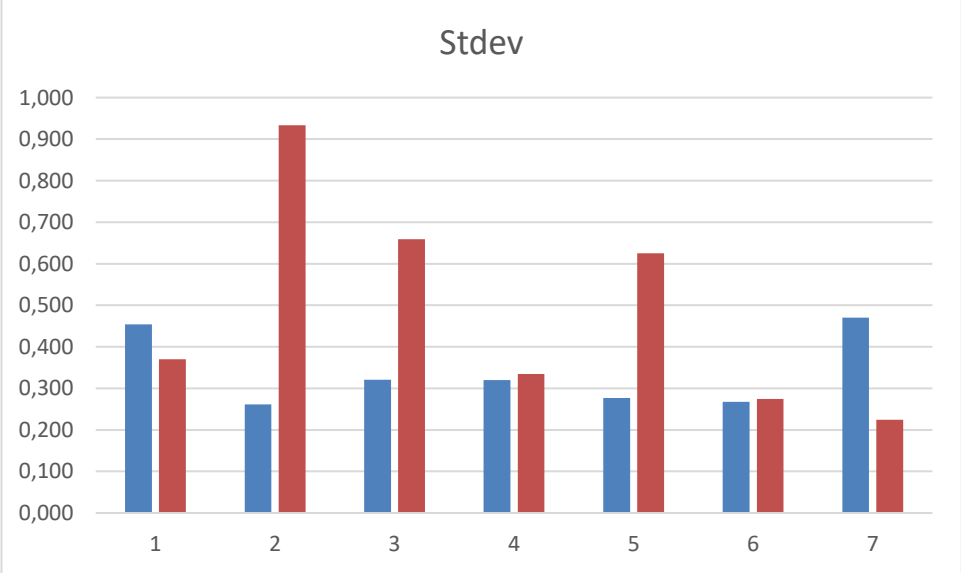
MiniVUX	Mean	-5,25	Min	0,83
	Std	42,02	Max	135,23

SPL-100_65	Mean	-40,42	Min	8,91
	Std	67,43	Max	170,56
SPL-100_120	Mean	-11,37	Min	1,15
	Std	55,53	Max	175,08
Titan	Mean	-27,44	Min	9,05
	Std	53,56	Max	168,00
Titan Ch1	Mean	-0,78	Min	3,09
	Std	55,88	Max	180,66
Titan Ch2	Mean	2,29	Min	0,16
	Std	54,33	Max	149,37
Titan Ch3	Mean	-59,84	Min	1,41
	Std	64,58	Max	199,15

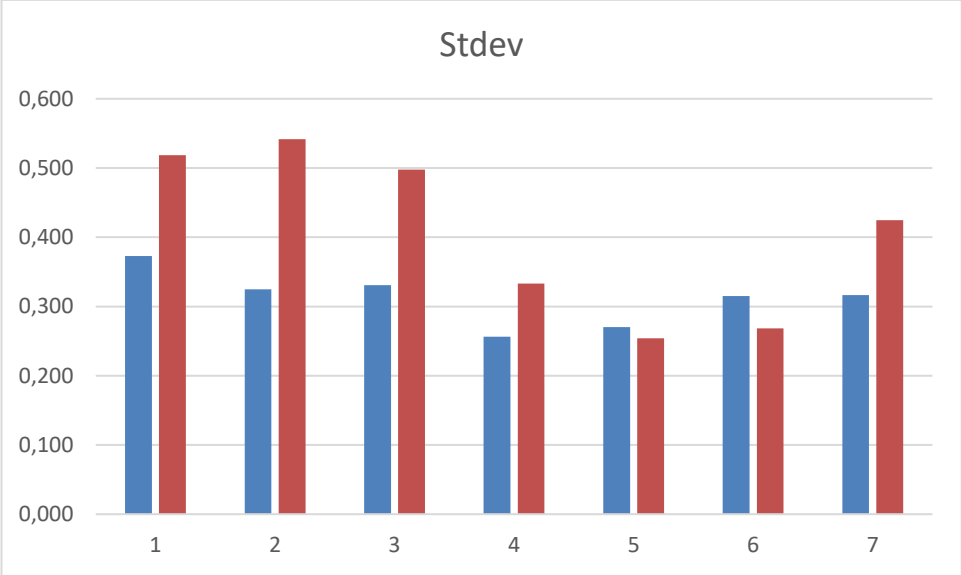
Appendix 2. Outline figures of the lengths and widths of the buildings

Length averages and standard deviations of the buildings. Blue color = building outlines under 40 m. Red color = building outlines more than 40 m.



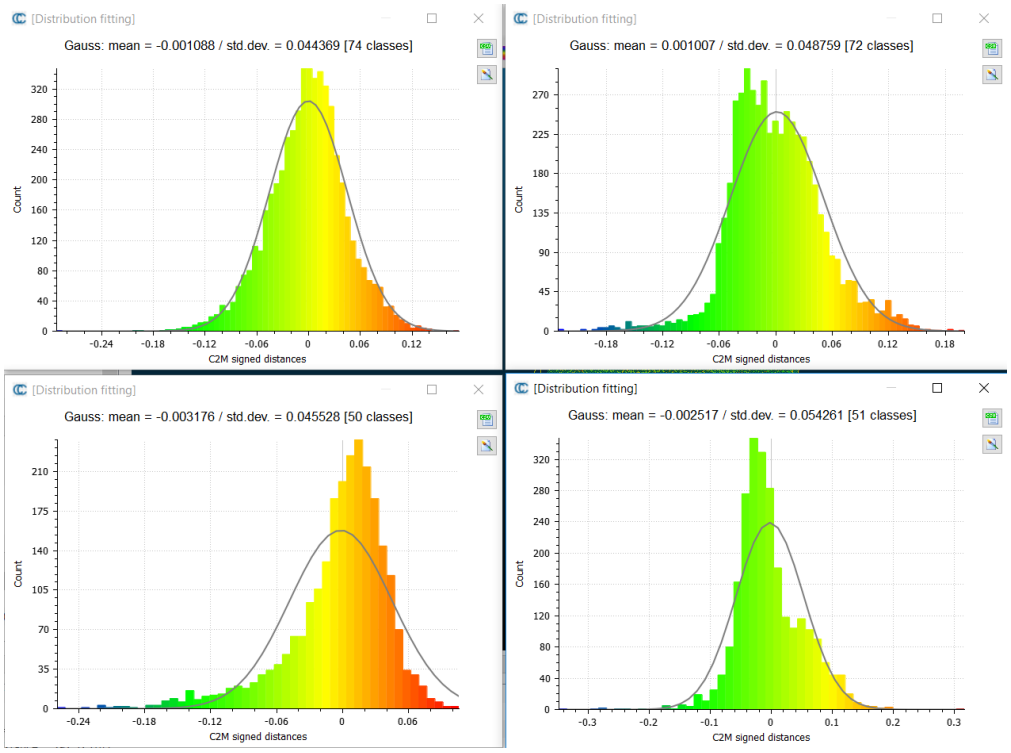


Width averages and standard deviations of the buildings. Blue color = building outlines under 40 m. Red color = building outlines more than 40 m.

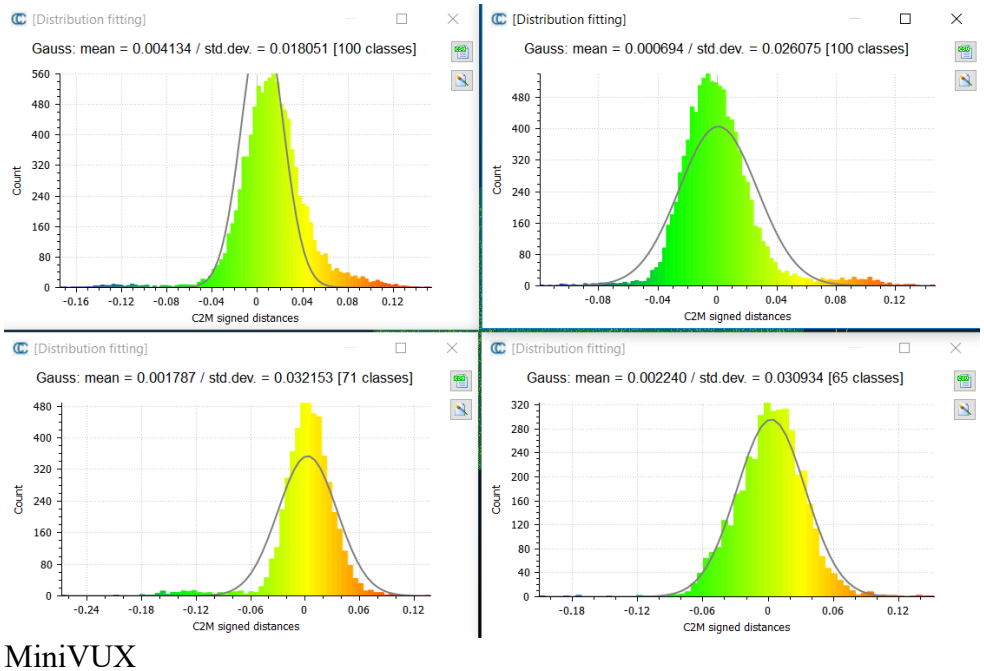


Appendix 3. Data deviations of the conical roof of the detached house from the block 4

VUX
Number of points of the planes (North, South, West, East): 5369, 5119, 2471, 2504

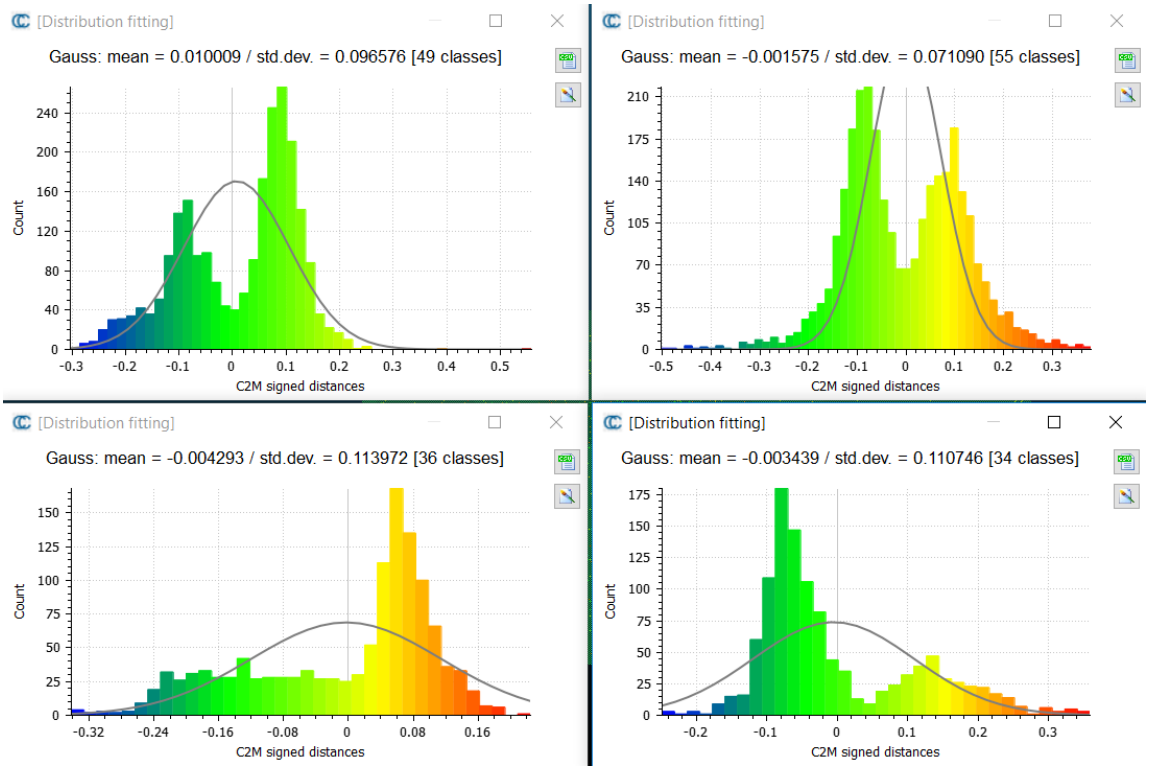


VUX_1min
Number of points (N, S, W, E): 9851, 9979, 4991, 4115

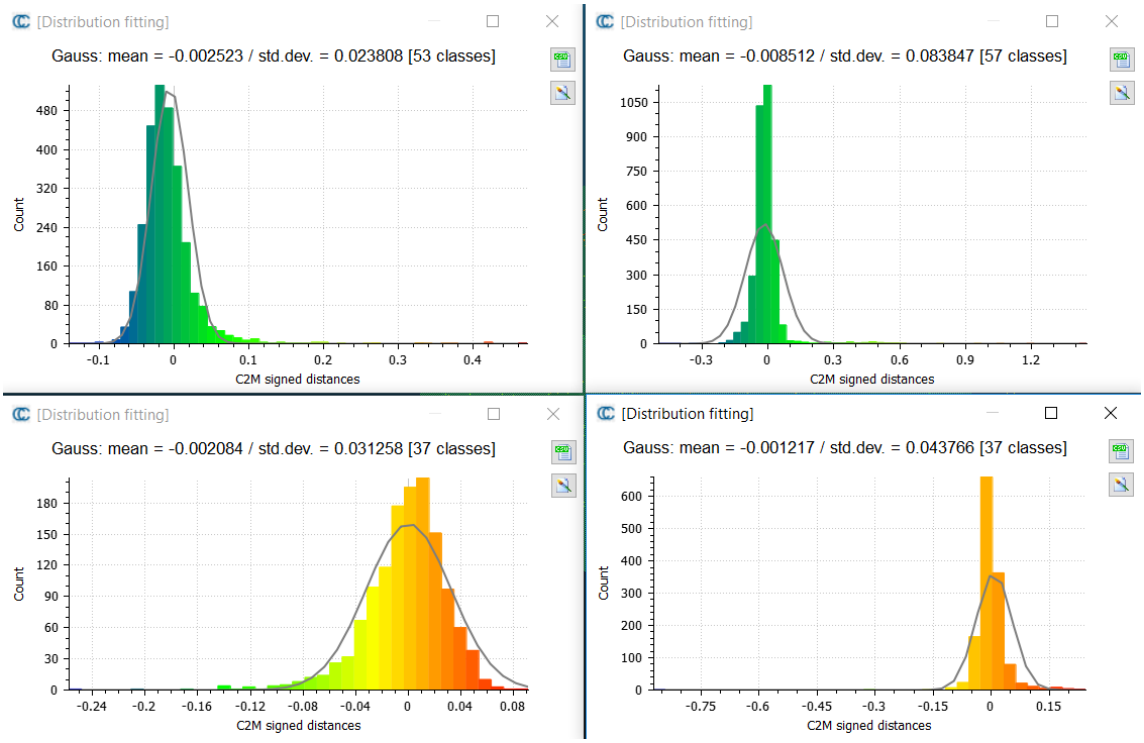


MiniVUX

Number of points (N, S, W, E): 2352, 2944, 1250, 1150

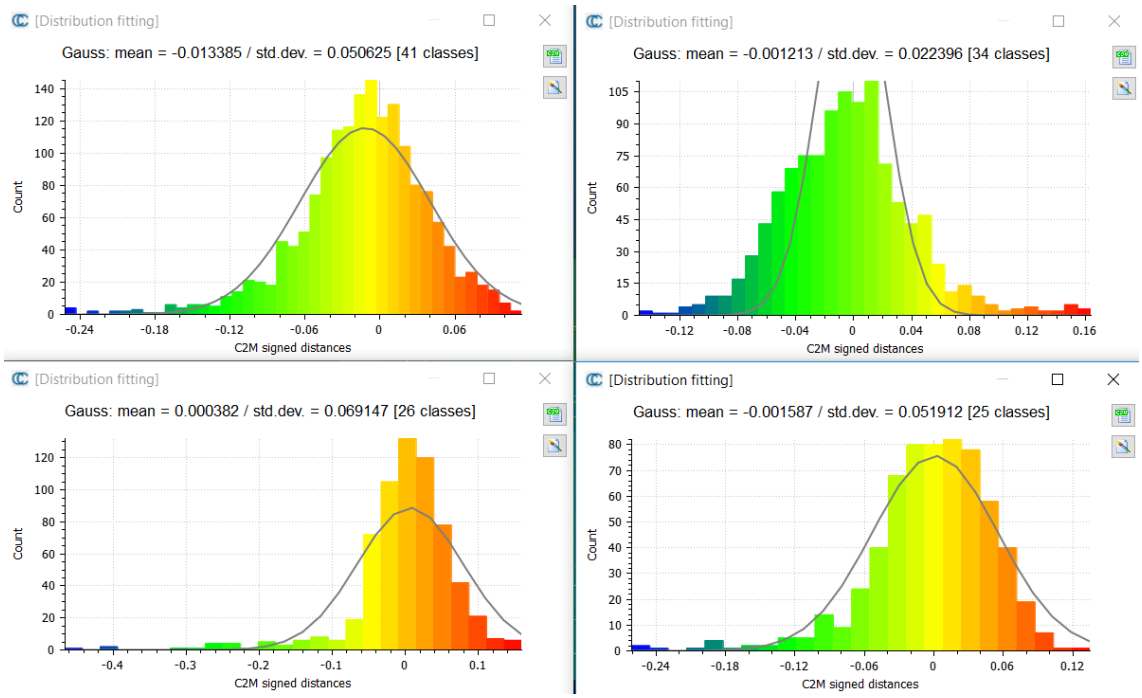


SPL100_65
Number of points (N, S, W, E): 2759, 3232, 1334, 1352

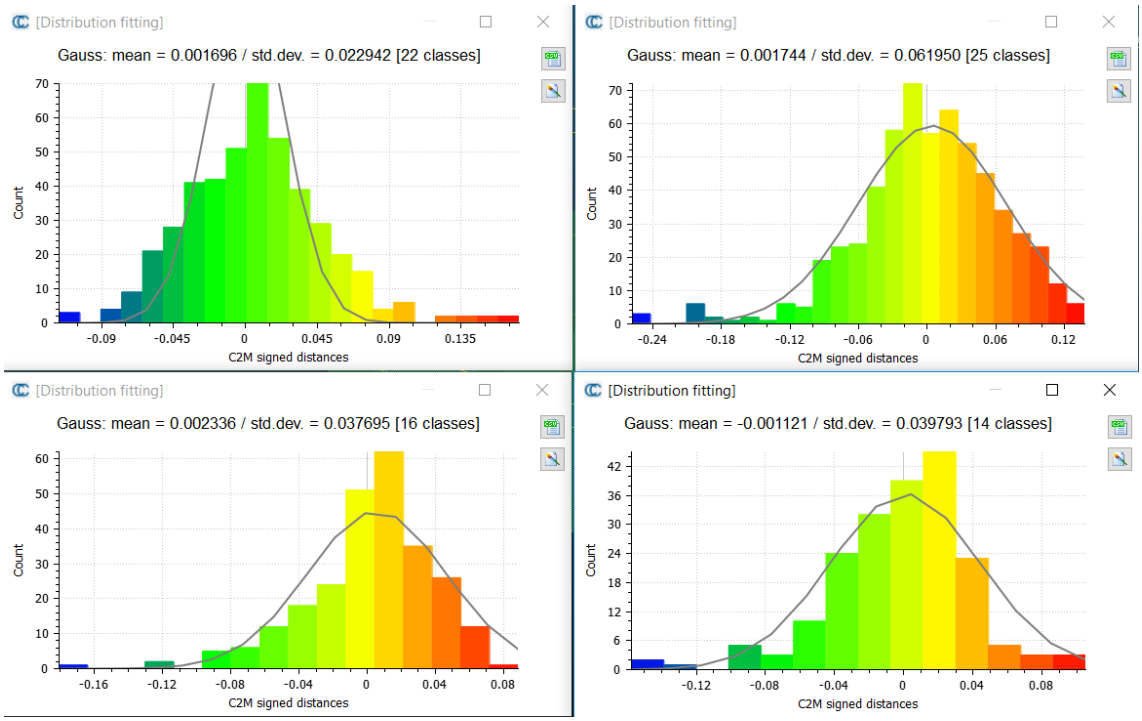


SPL100_120 (all flight lines)

Number of points (N, S, W, E): 1648, 1105, 644, 624

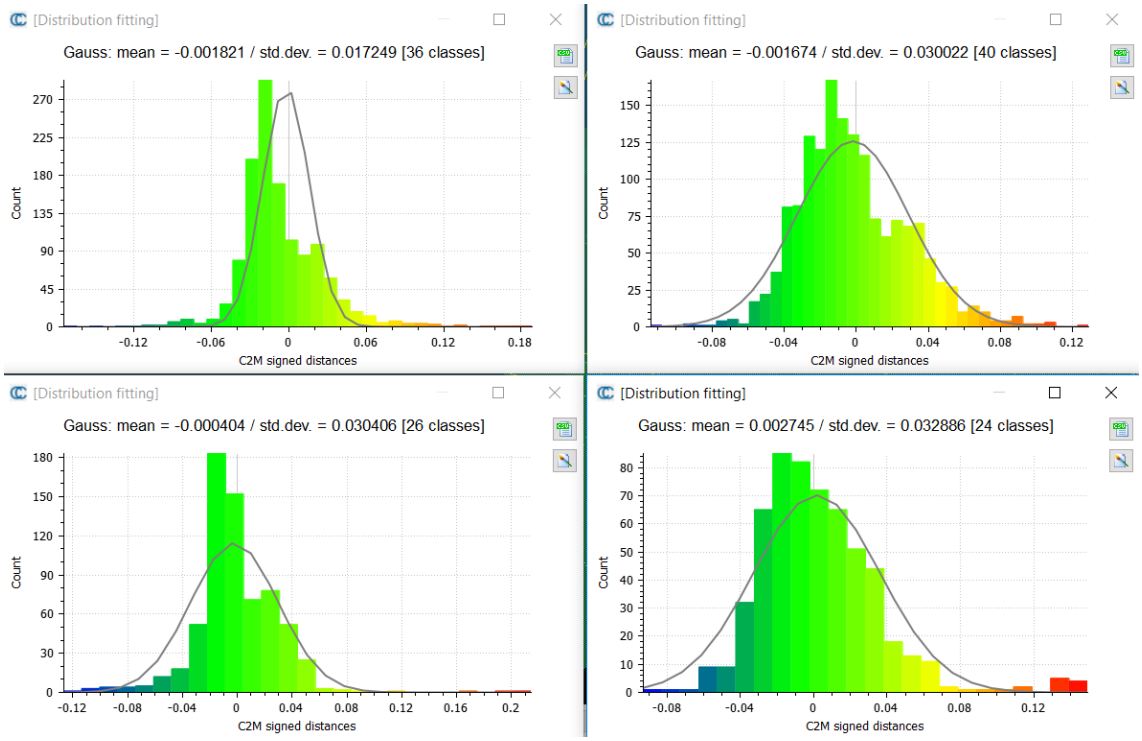


SPL100_120 (flight line 4)
Number of points (N, S, W, E): 444, 585, 255, 195



Titan (all flight lines)

Number of points (N, S, W, E): 1241, 1562, 670, 575



Titan (flight line 12)
Number of points (N, S, W, E): 1241, 1553, 684, 567