Comparing AHN point clouds for their performance in representing 3D buildings in Zuid-Holland

A quantitative and qualitative performance review between AHN3 and AHN4

Master Thesis

Author:	**Niek Manders**
Student Number:	**4237307**
Email:	**n.manders@students.uu.nl**
Supervisor:	**Edward Verbree**
Responsible Professor:	**Peter van Oosterom**

Abstract

This thesis investigates the performance enhancement of the AHN4 point cloud in comparison to the AHN3 point cloud for representing 3D buildings. The study focuses on the key performance aspects: Point Density and completeness, which both impact the perceptive quality of an object for both the Human Vision System and for Computer Algorithms.

The analysis of an object by computer algorithms pertains to the ability of a computer or software system to interpret and comprehend the characteristics, shape, and properties of a 3D object. Point Density offers insights into the quantity of points representing an object. However, it does not provide information about the spatial distribution of these points or the presence of regions without any points. In order to compare point clouds it is necessary to calculate the completeness of the point clouds in 3D by finding gaps in the point clouds. A gap is a piece of surface in a point cloud where there are no points. A good indication for gaps in a 3D point cloud is Point Spacing. Point Spacing can be seen as the Point Density in 3D. Point Spacing is the average distance from a point in a point cloud to all adjacent points. In this research a Delauney triangulation is used on the 3D points to generate a mesh. The triangles in this mesh are an indication of Point Spacing. Every triangle in the generated mesh is the result of three distances to adjacent points. A triangle large 3D area of such a triangle could therefore indicate a gap.

The research reveals that AHN4 exhibits a significantly higher Point Density than AHN3, with the majority of buildings demonstrating improved Point Density in AHN4. This higher Point Density suggests a general improvement in the perceptive quality of the point cloud for the Human Vision System. While AHN4 tends to have fewer gaps on average all planes of a building considered, both datasets show a deficiency in representing wall elements however. Also it is researched that in AHN4 there is less uniform in point density compared to AHN3 and that there are bigger regional differences.

Despite fewer data capture flights for AHN4, the results indicate that this reduction does not necessarily lead to decreased Point Density or increased gaps. Weak correlations suggest that increased distance to flight lines may slightly improve Point Density while introducing slight increases in gaps. A case study confirms that AHN4 generally offers higher Point Density and improved completeness compared to AHN3, even in areas where AHN3 had extra flights conducted and therefore more overlapping scans. However, certain localized regions slightly favor AHN3 in terms of completeness due to these overlapping scans.

In conclusion, the main research question *"To what degree is the AHN4 point cloud an improvement in performance compared to the AHN3 point cloud for the representation of 3D buildings?"* will be answered. The AHN4 point cloud is a improvement for both the Human Vision System and for Computer Algorithms, primarily due to its higher Point Density. Notably, there are areas where AHN3 excels in terms of completeness. These are however limited but underscore the need for future scanning requirements to consider occlusion in acquiring point clouds. The study introduces an alternative approach to evaluating gaps and occlusion by proposing the use of Point Spacing. The methods used in this research can contribute in calculating the performance of future point clouds.

Table of contents

1. Introduction

In this research two point clouds will be compared on their performance as a 3D representation of buildings. The two point clouds that are being compared are AHN3 and AHN4. The main objective is to determine whether AHN4 is an improvement compared to AHN3 in representing 3D buildings. This introduction part of this research will present the problem that this research is trying to address in the problem statement. This part will also include contextual information highlighting the necessity for conducting this study. Then the research objectives of this research will be given. The scope of this research will be mentioned at the end of this chapter.

1.1 Problem statement and context

Digital Twins, originally developed for the manufacturing industry, are becoming increasingly popular in the field of Geographic Information Systems (GIS). A Digital Twin is a digital replica or representation of a physical asset, allowing simulations of changes to be made without affecting the real-world object. While it is relatively straightforward to create Digital Twins for manufactured products with a fixed design, the application of Digital Twins in GIS presents a new set of challenges and opportunities.

Creating a Digital Twin in the GIS context is a complex task. Despite there being many different definitions of a Digital Twin in GIS, it is generally accepted that it includes a 3D model of a physical area (Schrotter & Hürzeler, 2020). The physical world is constantly changing and highly complex, making it difficult to create an exact replica. It is nearly impossible to spatially model the world 1:1 in content, scale, detail, and time, but it is important to strive for an accurate representation of reality as much as possible. With the latest versions of the AHN datasets, one could argue that it is achieving success, as it often provides a remarkably accurate depiction of reality.

Due to limitations in computing power, loading times, and storage capacity, it may not be possible to use a high level of detail in Digital Twin models for all applications. Therefore, any geospatial representation of reality is a model at a specific level of detail at a specific time (Stoter et al, 2021). Nonetheless, computers are improving over the years resulting in cheaper storage and shorter translation times. Therefore in most cases a higher level of detail in Digital Twins can be chosen. However, the added value of this higher level of detail depends on the specific application.

Digital Twins can be used for various types of analyses, such as shadow and sun analysis. These analyses require a relatively high level of detail to be accurate, for example, to check for potential shading issues for neighbors during a permit application process or to estimate the amount of solar panels which could be placed on a roof and to estimate the power generated from these panels. It is essential the shape is correctly registered, since such an estimation would be useless otherwise.

The creation of this Digital Twin starts with a point cloud in most cases. In many aspects, it can be said that point clouds are a Digital Twin in themselves. It is a good Digital Twin due to the fact that it is a very close copy of the reality, with every point representing a location where there is a surface. Point clouds do in most cases however miss information about the object it represents, in most instances just a classification. Point clouds refer to unstructured datasets consisting of enormous collections of individual points with their associated X, Y, and Z coordinates (Wang et al., 2020). Point clouds are used to represent objects, environments, or spatial phenomena from the real world in a discrete and approximate way (Richter, 2018). While one point only provides a small sample of a surface, thousands, millions, or even billions of these together can form something greater than the sum of their parts: a three-dimensional ‘Digital Twin’ (Batty, 2018).

Point cloud data can be acquired through various means. A very important method is using a LiDAR scanner. A LiDAR scanner is an active remote sensing instrument; that is, it transmits electromagnetic radiation and measures the radiation that is scattered back to a receiver after interacting with the Earth its atmosphere or objects on the surface (Longley et al, 2015).

Point clouds can be used to generate 3D objects, but the process often results in a loss of detail due to simplification. Humans have a superior ability to recognize and interpret object shapes and fill in gaps. Computers however often struggle with this task, resulting in inaccuracies or strange artifacts when generating 3D objects from point clouds, especially when dealing with complex shapes, vegetation, or limited data points. This is due to the fact that computers need algorithms in order to interpret object shapes. There are no algorithms that are able to do this better than humans. But potentially a combination of Artificial intelligence and machine learning could be helping developers to generate algorithms that can match the human brain.

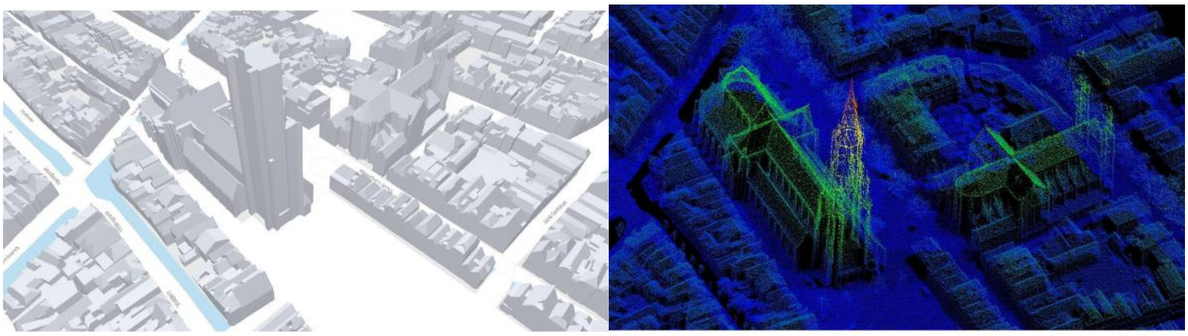


Figure 1 The Nieuwe Kerk (Delft) in the 3D BAG (left), and the corresponding AHN4 point cloud (right).

In the research by Baauw (2021), it is mentioned that for some buildings, this loss of detail can be explained by the way the modeling algorithm works. The algorithm models the object from top to bottom, assuming that the contour of the roof corresponds with the contour of the building's 2D footprint. After this, the facades are modeled downwards.

Generating a 3D model from a point cloud results in a loss of detail, but it still has its benefits. Points in a point cloud represent the precise location of a surface. When a 3D object model is generated there most of the time generalisation in place which means that the precise location of a surface is lost, but also other details like intensity and color of a point is lost. One advantage of a 3D object model is the ease of data management, as 3D objects only require storage of attribute data of the object once, instead of for each point in a point cloud.

Additionally, some 3D analyses can be more straightforward when using a solid surface model. For instance, calculating an object’s volume is often easier in this format. As a result, converting point cloud data into solid surfaces through post-processing can be useful for these types of analyses.

The most essential type of object for a Digital Twin of a city or town are buildings. Buildings form the foundation of a 3D city model (Ketzler et al., 2020). This research will therefore focus on the 3D representation of buildings.

There are three common measurement methods for capturing LiDAR data. The first is Terrestrial Laser Scanning (TLS). TLS uses a fixed location from which the surrounding area is captured by a laser scanner. This method is mostly used for small areas. The second method is Mobile Laser Scanning (MLS), which captures data from a moving object like a car or a bike. The last measurement method is Airborne Laser Scanning (ALS), where the laser scanner is attached to an aircraft. This allows it to capture LiDAR data for large areas.

ALS is used in the Netherlands to generate LiDAR data for the entire country. This nationwide point cloud is included in the Actueel Hoogtebestand Nederland (AHN, 2022). The program started in 1997 with AHN1, followed by AHN2, created between 2007 and 2012, and its successor, AHN3, spanning 2014 to 2019 (Leusink, 2019). For AHN4, the next iteration, data has already been captured in 2020 and 2021. Since January 1, 2023, the entire AHN4 dataset is available, and scanning for AHN5 will commence in 2023 (AHN, 2023).

The Dutch registry of buildings (containing all buildings in the Netherlands), known as the Basisregistratie Adressen en Gebouwen (BAG), is made into 3D with extensive use of the AHN dataset (Peters et al., 2021). Figure 1 illustrates how this dataset appears. This dataset is updated with the new AHN4 dataset.

The documentation of January the 3D BAG of TU Delft (2023) states that the current 3D BAG dataset utilizes AHN3 data, which was collected between 2014 and 2019. Further along in the same documentation, the following statement is made:

> *"At the moment of writing the new AHN4 will soon become available for a part of the country. The new AHN will have a different, improved quality, compared to the AHN3, and we expect that this will have a visible impact on the 3D BAG as well"*
>
> - *Documentation 3D BAG (TUDelft3D, Januari 2023)*

This statement suggests that the new AHN4 dataset will improve the generation of 3D buildings compared to the previous version (AHN3). However, there are concerns that it could represent a step backwards in certain aspects. Generally, it is expected that AHN4 will offer a higher resolution, with more points per square meter on average. Nonetheless, there are reports from colleges in the GIS industry indicating that fewer flights were conducted to capture data for the entire Netherlands. The impact of fewer flights is unknown since it is not known how high the airplanes flew and whether it impacts the quality of the point clouds. However, fewer flights on the same height could lead to less overlaps, resulting in more blind spots in the point cloud and consequently, gaps in the data. This aspect will be further explored in this thesis.

At the moment of writing the 3DBAG did update their dataset using the AHN4 point cloud. However they found that AHN4 for many buildings lacked completeness due to gaps. This is why they analysed whether there are notable gaps and for these buildings they looked whether the building has changed since the acquisition of AHN3. For these buildings they kept using AHN3 instead of the new AHN4 point cloud (TUDelft3D, 2023)

1.2 Research Objectives

The primary objective of this research is to enable a comparison between two point clouds in terms of their ability to represent a 3D building. This will be achieved by determining whether AHN4 is an improvement over AHN3 for the representation of 3D buildings. A performance evaluation will be carried out between these two datasets, leading to the formulation of the following main research question:

“To what degree is the AHN4 point cloud an improvement in performance compared to the AHN3 point cloud for the representation of 3D buildings?"

A literature review will be conducted first to determine the important aspects of point clouds for representing 3D buildings. Furthermore, related works will be examined to compare the significance of these aspects for both the Human Vision System and for Computer algorithms in their perceived quality of 3D buildings. Computer Algorithms are used to interpret and understand the characteristics, shape, and properties of a 3D object. The corresponding sub question is *"What are the key performance aspects of a point cloud that affect the perceived quality for the Human Vision System, and how does this affect compare to the perceived quality for Computer Algorithms?”.*

In Tomljenovic & Rousell (2014) is stated that for the automated building extraction from ALS data the completeness and density is important. The density of a point cloud refers to how many points there are in a point cloud in each area. This is a quantitative aspect that will be researched in the sub question: *"What is the average Point Density of buildings in AHN3 and AHN4, and how do these densities compare between the two datasets?".*

Completeness of a point cloud refers to the fact whether a point cloud has an even distribution of points around an object (in this case a building) without any gaps. This qualitative aspect is harder to research than the density of a point cloud due to the fact that more is not always better. Therefore the Point Spacing is calculated. Point Spacing is the average distance from a point in a point cloud to its adjacent points. By calculating this distance in 3D it is possible to use Point Spacing to find gaps in a point cloud. How the Point Spacing is calculated is stated in the methodology chapter. The corresponding sub question is: *"What is the Point Spacing on building elements in AHN3 and AHN4, and how do the patterns in Point Spacing compare in terms of completeness and gaps in the point clouds?"*

The problem statement indicates that there may be fewer flight lines used in AHN4 compared to its predecessor, which could negatively impact its performance. This research aims to investigate whether this is indeed the case and whether the reduction in flight lines has a significant impact on the performance of AHN4 compared to AHN3. This effect will be researched in the sub question: *" How do the quantitative and qualitative differences between AHN4 and AHN3 relate to the distance from a building to the nearest flight line?".*

Also a case study will be conducted in which will be looked in an area where the scanning areas per flight can be retrieved. In this area will be looked more subjectively with the Human Vision System what the impact is on the amount of overlaps between the scanning areas and whether less flight lines also resulted in less overlaps.

By answering the sub-questions, it will be possible to determine if AHN4 performs better in representing 3D buildings than AHN3. Additionally, the main question can be further answered by distinguishing between the direct use of point clouds or when 3D objects need to be generated from the point clouds. When point clouds are used directly, the perceived quality for the Human Vision System is the primary concern, while when a computer generates a 3D object, the perceived quality for Computer Algorithms is more important. The literature review of the first sub-question will provide insight into the important aspects for both types of usage.

1.3 Scope of the research

This research primarily aims to evaluate the performance of AHN3 and AHN4 in representing 3D buildings, and as such, will not extensively cover or have a limited scope in regard to:

- The accuracy of the points in the point cloud will not be researched in this research. This research will focus mainly on the completeness and the density of point clouds.
- Building Information Modelling (BIM) can be a very detailed 3D model of a building, because every material and piece of a building can be included. All the way to which bolts are used in the construction. Due to the fact that they are not available for all the buildings in the entire Netherlands it will not be included in this research.
- The modelling of the interior of buildings cannot be deducted from a ALS point cloud like AHN so will not be dealt with in this research.
- Photogrammetry is a method of using overlap in images to generate a 3D point cloud. In this research however will be focussed on LiDAR point clouds.
- The texture and colour of buildings will not be addressed in this research. While it may enhance the visual representation of buildings, it has little impact on the usability of most 3D GIS analyses. The most crucial aspect for these analyses is the accuracy of the building's shape. Furthermore, incorporating texture on an inaccurately modelled 3D building can result in a jarring visual appearance. When points of a point cloud have a color attribute it does however makes a big difference on the human perception of a surface. However this attribute is not included in the AHN point clouds and therefore will not be included in the research.

2. Literature Review

In this Literature Review, all publications and sources related to the research topic will be examined. First, the background of LiDAR and its role in generating and applying Nationwide Point Clouds will be provided. Next, the perception of objects in point clouds will be explored. This section will focus on the Key Performance Aspects of point clouds and how these aspects influence the perceived quality of point clouds. The research question, *"What are the key performance aspects of a point cloud that affect the perceived quality for the Human Vision System, and how does this affect compare to the perceived quality for Computer Algorithms?"* will be answered based on the insights derived from this literature review. Finally, the limitations of using Point Density will be discussed, and the argument will be made for considering Point Spacing as a potentially better indicator as a Key Performance Aspect for establishing requirements for a point cloud.

2.1 Background LiDAR

2.1.1 What is LiDAR?

LiDAR, which stands for Light Detection and Ranging, is an advanced surveying technique that utilizes laser beams to precisely measure the surrounding environment. By analysing the time it takes for the laser beams to return and the strength of their reflections, LiDAR calculates the distances to the surfaces they encounter. This process, combined with the scan angle, enables the determination of precise positions for each reflection point, resulting in accurate x, y, and z coordinates. The culmination of these calculations produces a comprehensive 3D point cloud, where each point represents the location of a reflecting surface (Bochove, 2019).

A LiDAR scanner can be placed on three different kinds of ways. The most used way is Airborn Laser Scanning (ALS), Terrestrial Laser Scanning (TLS) and Mobile Laser Scanning (MLS). In figure 2 are all the (dis)advantages of each way of laser scanning (Pradhan & Sameen, 2020).

ALS	MLS	TLS
Direct view of pavement and building tops	Good view of pavement, unable to capture building tops	Good view of pavement with details
Oblique view of vertical faces	Direct view of vertical faces	Direct view of vertical faces with more flexibility
Fast coverage	Slow coverage	Slower coverage
Large footprint	Small footprint	Small footprint
Far-range travel	Short-range travel	Short-to-moderate range travel
Not limited to the area visible from the roadway	Limited to objects close to and visible from the roadway	Limited to objects close to the roadway
Low point density (1–60 point/m^2)	High point density (100 points/m^2)	Very high point density (500 points/m^2)
Limited options for setup locations	Good options for setup locations	Better options for setup locations
Difficult to operate and requires adequate training	Difficult to operate but easier than ALS as a pilot is not needed	Easy to operate with less training
Provides a low level of details	Provides a high level of details	Provides the highest level of details
Low accuracy and resolution of road features	High accuracy and resolution of road features	Higher accuracy and resolution of road features
Highest cost-effectiveness	High cost-effectiveness	Low cost-effectiveness

Figure 2 Comparison ALS, MLS and TLS source: Pradhan & Sameen (2020).

This research focuses on Airborne Laser Scanning (ALS), a LiDAR method used for acquiring AHN datasets. ALS offers high accuracy, providing precise measurements for detailed mapping of terrain, objects, and vegetation. Its accuracy makes it valuable for applications like urban planning, forestry management, and flood modelling. Additionally, ALS is non-invasive, enabling data collection from inaccessible or hazardous locations such as dense forests, mountainous regions, or disaster-affected areas (Okyay et al., 2019).

Another advantage of ALS is its rapid data acquisition capability. LiDAR systems mounted on aircrafts can cover large areas quickly, acquiring vast amounts of data within a short period. This capability is utilized in many countries to generate a nationwide point cloud, which also serves as a elevation model (Ahokas et al., 2008). AHN, as mentioned earlier, is an example of such a nationwide point cloud with a corresponding elevation model.

One drawback is the lower Point Density compared to other surveying methods. ALS may not capture as many data points per unit area, resulting in a less detailed representation of the terrain or objects being scanned. This reduced Point Density can affect the accuracy and precision of the generated 3D models (Pradhan & Sameen, 2020).

Another challenge is dealing with the large amounts of data generated by ALS. The scanning process collects vast volumes of point cloud data, which can be overwhelming to manage and process. It requires specialized software and hardware capabilities to efficiently handle, store, and analyze this data, adding complexity and potentially increasing costs (Van Oosterom et al., 2015).

In the context of the Netherlands, the high air traffic density poses a specific challenge for ALS operations. The country has a densely populated airspace, especially near urban areas and major airports. Coordinating flight paths and ensuring the safety of aircraft becomes crucial when conducting ALS surveys. This can lead to increased regulatory requirements, coordination efforts, and potential limitations on when and where ALS can be conducted (Bochove, 2019).

Weather conditions also play a significant role in ALS operations. Adverse weather, such as heavy rain, fog, or low cloud cover, can impede the effectiveness of ALS systems. These conditions can cause reduced data quality or even make data collection impossible. It becomes essential to consider weather patterns and select suitable timeframes for ALS surveys to ensure optimal data acquisition and accuracy (Bochove, 2019).

There are multiple different type of scan patterns for ALS. The spatial pattern of points is largely dependent on the type of scanner. Rotating mirror scanners and oscillating mirror scanners produce parallel scan lines (assuming constant flight direction) or a zigzag line (Glennie et al, 2013). While parallel scan lines result in equal Point Density (Situation A and B in figure 3), a zigzag line always results in variable Point Density (Situation C in figure 3).

Scanners with conical scan patterns (Situation D in figure 3), measure with a constant scan angle in all directions (Fernandez-Diaz et al, 2014). Consequently, each location is scanned twice from different angles, and the conical pattern causes very high density on the edges of each swath, unless the scanner is changing the speed of scanning (Petras et al, 2023).

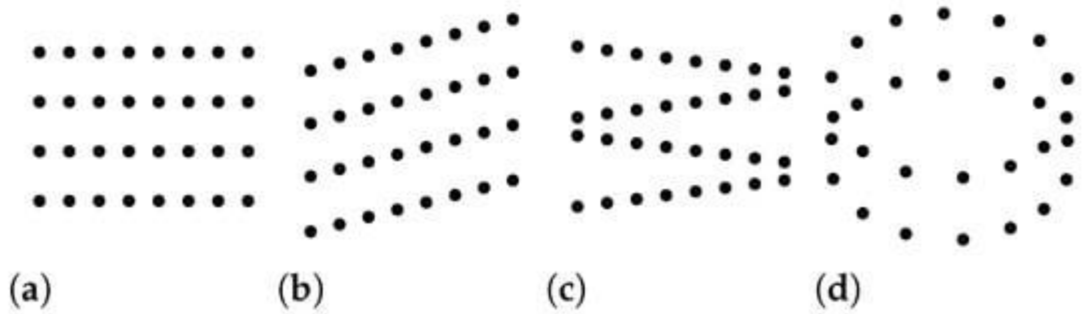


Figure 3 Different scanning patterns for ALS (Petras et al, 2023)

2.1.2 Applications of Nationwide Point Clouds

Nationwide Point Clouds have multiple applications. In 2012, the most common applications of the AHN dataset were water management, archaeological research, and scientific research (Donker and van Loenen, 2013). For example, the water board uses the AHN to check if the dikes are not sinking. Rijkswaterstaat, on the other hand, extensively uses the elevation data to generate sound maps and determine the placement of sound barriers along highways.

The main funders of the AHN are the Provinces, Water Boards, and Rijkswaterstaat. Since AHN2, these funders have covered the entire costs and offered the AHN data as open data. Since the data became open, many other users have also emerged. The results of the survey conducted by Bregt et al. (2016) suggest a shift in focus towards Construction & Infrastructure, Environment Soil & Nature, and Spatial Planning. Additionally, the research indicates that the majority of AHN users are now commercial users, accounting for around 80%.

An example of a commercial application of the AHN is the Zonatlas. Using the AHN point cloud, Zonatlas creates a 3D model and performs a 3D analysis to determine the number of solar panels that can be installed on a house's roof and calculate their potential yield and payback period as shown in figure 4 (Zonatlas, 2023).

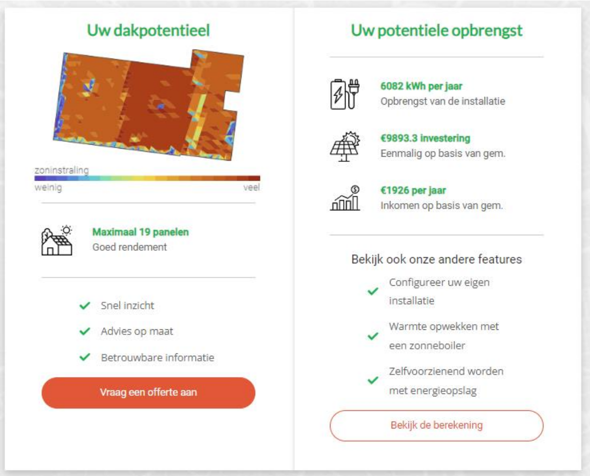


Figure 4 The potential yield of a house analyzed by Zonatlas (Zonatlas, 2023)

One of the primary applications of point cloud data lies in Digital Twin cities. Digital Twin cities serve as virtual extensions or replicas of real-world cities (Schrotter & Hürzeler, 2020). These Digital Twin cities ideally are connected to lots of realtime sensor data from the actual cities (Papyshev & Yarime, 2021).

Digital Twin cities can take various forms, with no one-size-fits-all approach in their applications. However, the most common form is a 3D city model (Peters et al., 2022). These models accurately capture the physical aspects of cities, making them valuable tools in urban planning and informed decision-making for governments (De Kruijf & Steenbakkers, 2021). Typically, a 3D city model is created by integrating point clouds with other administrative data. In the Netherlands, the TU Delft has undertaken the ambitious challenge of generating a 3D building model of every building in the country, called the 3D BAG. This 3D building dataset can be seen as the foundation for a Digital Twin city. (Peters et al., 2022). See Figure 5 for an example of 3D BAG. The current iteration of the 3D BAG is based on AHN3, and ongoing research is evaluating the quality of AHN4 to determine whether it should be incorporated into the dataset.

Figure 5 De Dom (Utrecht) in 3D BAG

2.2 Key Performance Aspects of Point Clouds

This chapter aims to answer the research question: *"What are the key performance aspects of a point cloud that affect the perceived quality for the Human Vision System, and how does this affect compare to the perceived quality for Computer Algorithms?"*. The chapter will begin by examining the general key performance aspects. It will then explore the components of perceived quality and assess the significance of these key performance aspects for both the Human Vision System and Computer Algorithms.

Key performance aspects are indicators on how well a point cloud is suited for an application. The importance of these indicators will vary on the type of application. First will be looked at what kind of key performance aspects there are. After that the importance for both the Human Vision System as for Computer Algorithms will be discussed.

2.2.1 Point Density

Point Density is the most common used indicator for assessing the quality of airborne LiDAR data. In the United States Geological Survey (USGS) standard (Heidemann 2014), the Point Density is commonly referred to as the Nominal Pulse Density (NPD). It is a metric used to quantify the density of laser pulses emitted per defined area. Since every laser pulse that is emitted results in a point, the NPD is typically expressed as the average number of points per unit area, denoted as points per square meter (pts/m2) (Rupnik et al. 2015)

There are two common methods of calculating Point Density: local density and mean density. The most commonly used is the mean density as it gives a representation of the density across the entire project area. The mean Point Density is determined by dividing the total number of point by the overall project area. However the Point Density can vary within the project area which can be a problem for applications like generating 3D objects. For these kind of applications will often local Point Density be used. The local data density is calculated by point number dividing area in a small area (mostly a square meter). The most common way of calculating this is by making a grid (with a cell size of for example 1 square meter) and executing a point on area overlay analysis (Wu et al, 2011).

Point Density as a requirement

With every iteration of AHN new requirements are being developed. These requirements are used in the Tender which is used to find the best company to acquire the point cloud data. This company needs to deliver the point cloud in a given time and this point cloud must meet the minimum requirements that is stated in the Tender.

The requirements for AHN3 and AHN4 were primarily focused on vertical accuracy, with a maximum allowable offset of 5cm. AHN3 did not have a specific minimum Point Density requirement, but it typically ranged between 6 and 10 points per square meter. In contrast, AHN4 had a defined minimum Point Density specified in its tender, mandating at least 10 points per square meter overall and a minimum of 5 points per square meter under vegetation (Leusink, 2019). Generally, AHN4 datasets exhibited a Point Density ranging from 10 to 14 points per square meter (Actueel Hoogtebestand Nederland, 2023).

However, it is important to note that the requirements did not include any provisions for Point Spacing. Consequently, the company responsible for acquiring the AHN4 dataset could potentially have gaps in the data, as in the example of figure 6. The motivation for the company to mitigate these occlusion gaps is therefore significantly reduced. While larger overlaps during data acquisition could prevent these gaps, such an approach would incur higher costs. As a result, the incentive for ensuring the completeness of the dataset rests solely on accuracy and Point Density, disregarding the potential presence of data gaps.

The 3DBAG is a dataset of 3D buildings created by researchers from the TU Delft. During a presentation at the AHN – BM congress in Amersfoort on May 9th, the researchers discussed their findings regarding AHN4 and their plans to incorporate it into the next iteration of the 3DBAG dataset. In this presentation, it was stated that in most cases, the AHN4 dataset has a higher Point Density compared to AHN3, typically twice as high. This higher density yields more detailed information on the derived 3D buildings. However, it was observed that AHN4 exhibits more occlusion, leading to errors in building generation. Figure 5 illustrates the gaps (white areas) present in the AHN4 dataset but absent in the AHN3 dataset. Consequently, the conclusion drawn was that for buildings with such occlusion, AHN3 would be preferred (assuming the building has not changed in the meantime).

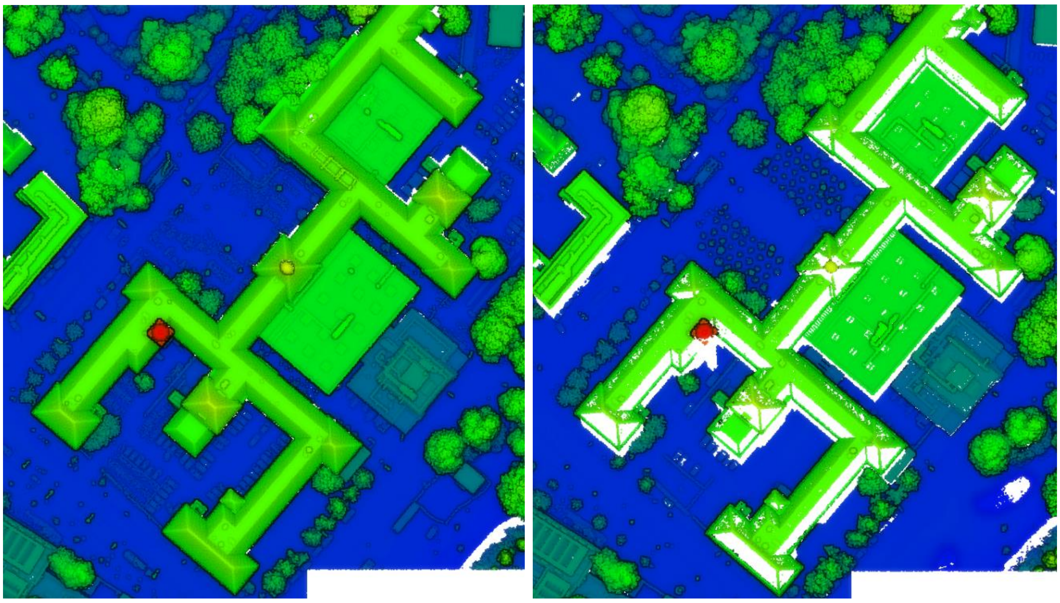


Figure 5 Comparison gaps AHN3 (left) and AHN4 (right) (Peters et al., 2023)

2.2.2 Completeness

A important key performance aspect is the completeness of a point cloud. Completeness refers to the degree to which all the objects and surfaces in a scene are represented in the point cloud data. It measures how well the point cloud captures the entirety of the environment being scanned or observed. If a point cloud is incomplete, important details may be missing, leading to inaccurate measurements, incomplete reconstructions, or flawed interpretations of the scene. For applications such as 3D modelling or object recognition, a complete point cloud is essential to ensure a faithful representation of the real-world environment.

A common phenomenon with LiDAR that affects the completeness of point clouds is occlusion. Occlusion is the phenomenon where objects or surfaces in a scene obstruct the line of sight to other objects or surfaces, causing them to be partially or completely hidden from view. In the context of point clouds, occlusion means that certain areas or features of the scene may not be adequately captured by the LiDAR scanner because they are obscured by other objects. It can result in missing or sparse data in regions that are occluded. For example, if a tree is blocking the view of a building during a LiDAR scan, the point cloud may lack accurate information about the obscured parts of the building.

In figure 6 is a schematic drawing of LiDAR scanning and the area that is occluded. The area that is captured by more than a single plain is overlap. These overlaps result in more acquisition costs because more flight hours are needed to capture the same region. Excessive overlaps are economically unjustified because scanning the same fragment several times does not increase the amount of information about the acquired object (Warchoł, 2019)

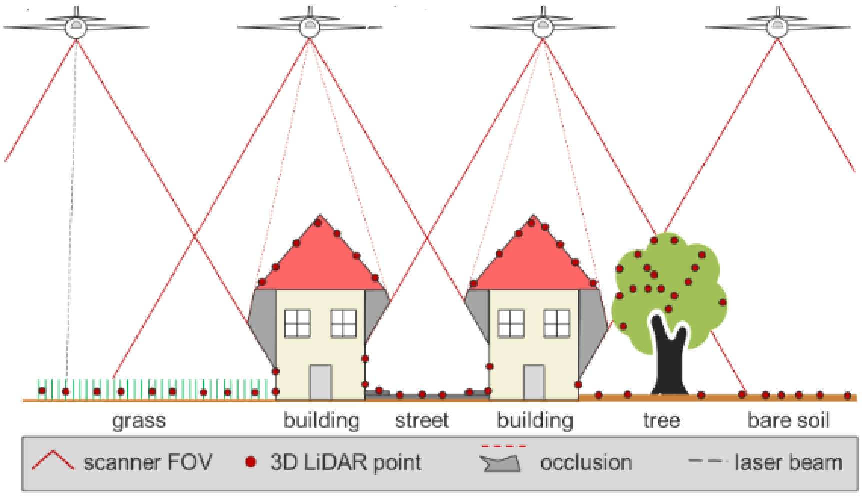


Figure 6 schematic example of LiDAR scanning and the occlusion (Mandlburger et al, 2017)

However in most cases there are to an extend overlaps, because it decreases the amount of occlusion and therefore improves the completeness of the point cloud (Mandlburger et al, 2017). When looking at the example of figure 6, it becomes clear that if the airplanes would fly more distant from each other and the overlap would decline the amount of occlusion would increase. As stated in the introduction it is said that for the new AHN4 point cloud there are flight hours made than its predecessor AHN3. This could indicate a higher level of occlusion and a less complete point cloud. Later on in this thesis this will be researched.
How many points there are and whether the point cloud is complete are not the only key performance aspects. Whether the points are accurate is also important. Accuracy of a point cloud refers to how closely the coordinates and attributes of individual points in the point cloud align with the true or reference values of the corresponding objects or surfaces in the real-world scene. A good accuracy is an essential requirement for a good 3D model.

2.2.3 Influence flight height on occlusion

The degree of occlusion is thus related to the angle of incidence in combination with the degree of overlap of acquisition. But does the flying altitude also influence the degree of occlusion? This paragraph will theoretically illuminate this aspect.

As mentioned earlier, there are indications that AHN4 may exhibit more gaps due to occlusion compared to AHN3. What is known is that the flight altitude during the acquisition of AHN4 is higher than that of AHN3. However, can this difference be attributed to the increase in flight altitude?

Airborne Laser Scanning is performed using a laser scanner with a specific angle for sending and receiving laser pulses. If an airplane flies at a higher altitude, it will scan a larger area on the ground. Nevertheless, the angle at which the pulses hit the ground remains constant. Figure 7 illustrates a scanner with a maximum scanning angle of 30 degrees, resulting in a maximum angle at the ground of 75 degrees. Whether the plane flies higher or lower, these angles do not change. Thus, this factor alone could not lead to increased occlusion.

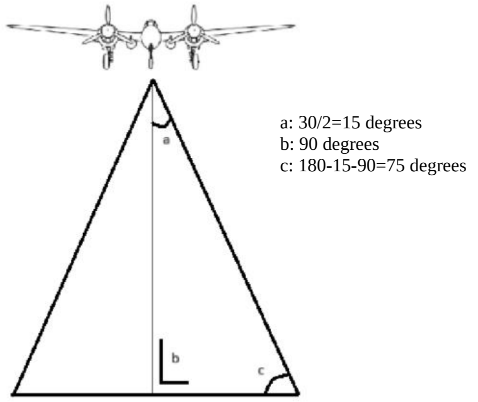


Figure 7 Example of scanning angle in ALS

That does not mean that flight height could indirect result in more occlusion. This has to do with the amount of overlap. When a airplane will fly higher with the same angle it will scan a larger surface. Often there is a certain amount of overlap to ensure that all area is scanned. However overlap also will decrease the amount of occlusion, because the these areas are scanned from two sides.

If the overlap stays the same but the flight height, and therefore the scanned area, increases it will result in a smaller percentage of the surface which is scanned from two angles. An example is shown in figure 8. There can be seen that in the left scenario, the overlap is 100 meters wide and the width of the scanning area in total is 1000 meters. This means that 80% of the surface has no overlap and 20% of the area has a overlap. The scenario on the right has the same angles but a higher flight height which results in a scanning width of 2000 meters. If the same overlap of 100 meters of overlap it will result in only 10% of the area which has a overlap. This means that in the left situation the 20% surface with a the smallest angle, and therefore scanned the most from one side, are also scanned from another angle. But in the right scenario just 10% of the surface is scanned from another angle which leaves the other 10 percent with a large amount of occlusion. So the flight height does not have to impact occlusion but then the overlap will have to increase with the increase in flight height.

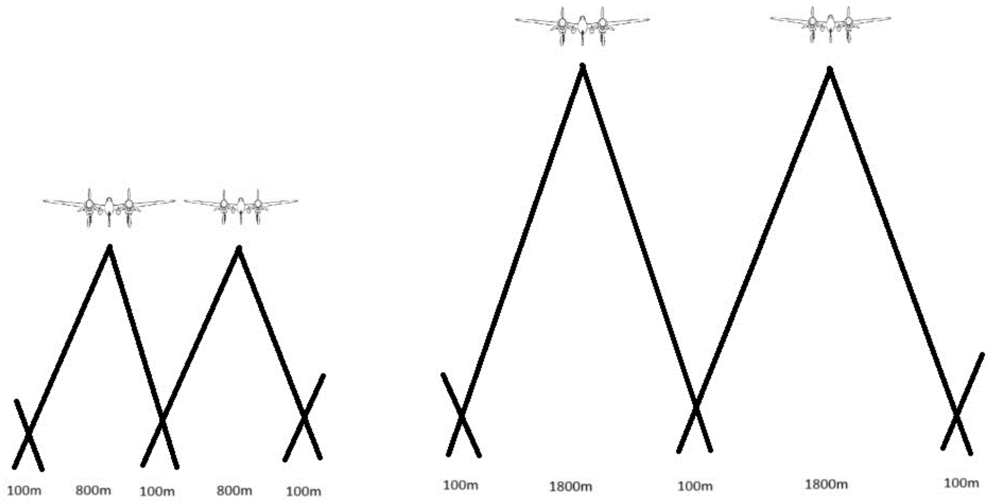


Figure 8 Example of increasing altitude with remaining width of overlap

2.2.4 Point Spacing

The method of calculating Point Density is either way (mean or local) a 2D analysis for calculating Point Density and does not take into account the Z-values of the point in a point cloud. Another way of calculating the density of a point cloud is by calculating Point Spacing. Point Spacing is defined as the distance between adjacent points. So instead of calculating the amount of points in a certain area, it will calculate the distance between adjacent points (Rohrbach, 2015).

Point Spacing can be used in different ways. Like Point Density, you can calculate the Mean Point Spacing and the Local Point Spacing. Mean Point Spacing assumes a regularly distributed point cloud. The Point Spacing is the square root of the average area per point, see figure 9. This method of Mean Point Spacing still does not take into account the Z-values of the points. The real benefit of Point Spacing over Point Density when looking at the Local Point Spacing.

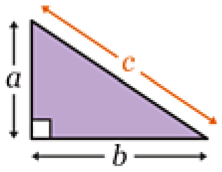
$$point\ spacing = \sqrt{\frac{1}{point\ density}}$$

Figure 9 formula for Mean Local Spacing (Rohrbach, 2015)

Like stated earlier Point Spacing calculates the average distance to adjacent points. Local Point Spacing calculates this average distance for every point and therefore does not assume that all points are distributed regularly. However, no sources were found that explain Local Point Spacing in detail.

Local Point Spacing calculates the average distance to a given amount (K) of adjacent points. This amount of points can be a fixed amount, for instance 5. In this case the distance to the 5 adjacent points will be calculated and summed and then divided by 5. The amount (K) of adjacent points do not necessarily have to be fixed. A method that can be used is by generating a mesh, for instance by applying Delaunay triangulation. In this mesh each point is connected to a variable amount of points. These points can be used as adjacent points, and therefore this will decide what amount of adjacent points (K) there are.

Take into account that Local Point Spacing can be used in both 2D and in 3D, this is due to the fact that you can use both the 2D distance, but also the 3D distances. A method to calculate the 3D distance between two 3D points is by using the formula of Pythagoras ($a^2 + b^2 = c^2$), see figure 10. First you will need to calculate the 2D distance where “a” is the offset on the x-axis between the points (abs(X1-X2)) and “b” is the offset on the y-axis (abs(Y1-Y2), “c” is then the 2D distance between the points. The same formula can be used to calculate the 3D distance, then “a” is the offset on the Z-axis (abs(Z1-Z2)) and “b” is the 2D distance. This results in “c”; the 3D distance between two points.



$$a^2 + b^2 = c^2$$

Figure 10 Formula of Pythagoras

The benefits of using the Local Point Spacing instead of Point Density is that it calculates the population of points on a 3D surface whether than a 2D surface. Point Density does not account for the height differences in the surface, and therefore the gaps in 3D. In Figure 11 is an extreme example of the difference between Point Density and Point Spacing. In this figure, which is a cross section of a fictional surface, the Point Density is equal everywhere (therefore there seem to be no gaps). However, the sides of this flat has no points at all. Point Spacing looks at the distance between the adjacent points and will therefore detect the gap on the sides of the flat correctly.

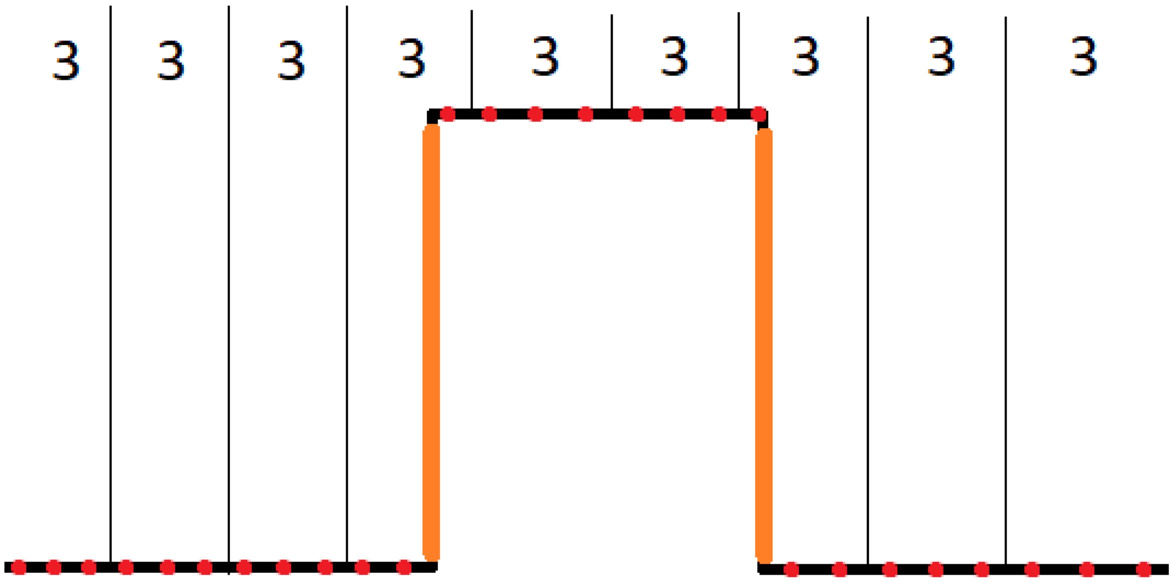


Figure 11 Equal Point Density (3) but gaps on the sides (orange)

2.3 Perceptional quality and object recognition

When looking specifically at perception of objects it is important to look at what a point cloud is.

A point cloud is a set of points with x,y and z values. Some of these points together can represent an object, but this cannot be found in the data itself. In order to know which points form an object, and what kind of object these points represent, it is crucial to recognise this object in the point cloud.

2.3.1 Human Visual System

When you look out of the window, your brain is recognising an enormous number of objects like cars, bikes, streetlights or a slide in the playground. Like other natural tasks that our brains perform effortlessly. This recognition is done by the Human Visual System (Liu et al, 2023).

In the literature there is a strong focus on applications of 3D models. However, there is little research about the perceived quality of these models and how to test it This is also mentioned by Zhang et al (2014). In their review they acknowledge that it is difficult to review the perceptual quality of a point cloud. The article states that there are two methods, objective and subjective, to evaluate the quality of 3D models depending on the evaluation subject. More specifically, if the subject is a human being, the method is called subjective evaluation. If the evaluation result is computed by machine, the method is called objective evaluation. The findings of a subjective evaluation can however be used for making a objective evaluation that can indicate certain aspects that could harm the perceived quality for the Human Visual System.

An example of such findings come from an subjective test that is performed by Zhang et al (2014) in which they down sampled a point cloud and letting people rate what the perceived quality is of the object the point cloud is representing. See figure 12 for an example of the down sampling. They found that the resolution change on a 3D model is almost linearly correlated with the visual perception of the Human Visual System. This means that a higher Point Density will most often lead to a higher perception of the object. Another research states that in terms of object recognition an higher point density does not affect the perceived quality much as long as the presentation is plausible for human viewers (da Silva Cruz, 2019).

Figure 12 Monkey point cloud models in different resolution by down-sample (Zhang et al, 2014)

The outcomes of these subjective researches indicate that a higher point density will provide a higher perceived quality of a point cloud for the Human Visual System.

2.3.2 Computers algorithms

Recognition at a level of the Human Vision System has turned out to be difficult to reproduce in artificial systems like computers. Computers do not work the same as the human brain. They use algorithms and mathematical models made by humans to process data like point clouds, with these models the computer is limited by the quality and quantity of point cloud data available. They may also struggle with objects or scenes that have complex structures.

The segmentation of point clouds is a fundamental step in processing 3D point clouds. Given the set of point clouds, the objective of the segmentation process is to cluster points with similar characteristics into homogeneous regions. For example when you want to calculate how many solar panels can be fitted on a roof of the building, it is needed to know where the building is and what the surface is of the roof (Nguyen, 2013).

In order to do 3D analysis on buildings it is common to reconstruct 3D objects out of point clouds. By deriving objects from the point cloud, such as buildings, the data can be organized and structured in a way that facilitates efficient 3D analysis. Most of the reconstruction methods work well if both the data is complete and the objects in the scene fit to assumptions made in the algorithm. However, the quality of 3D building models is actually determined by those situations where either data is lacking or the objects do not fit to general assumptions (Elberink & Vosselman, 2011). Relatively simple buildings do not need a high Point Density in order to let the algorithms reconstruct a reasonable 3D building. The Point Density has the most effect on the level of detail of the building such as chimneys on the roofs. In figure 13 the effect is shown of reducing the Point Density on relatively simple buildings.

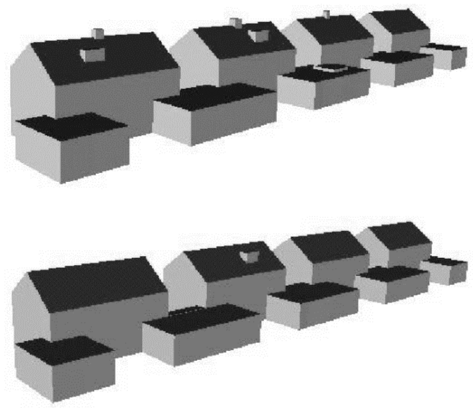


Figure 13 Effects of reducing the Point Density from 5-6 points per m2 (top) to 1.25 - 1.5 points per m2 (bottom) on the amount of reconstructed details (Vosselman & Dijkman, 2001).

The relation between the desired Level of Detail (LoD) should not only be related to the average Point Density of the dataset. For feature extraction purposes the variation in Point Density is often more important than the mean Point Density (Elberink & Vosselman, 2009).
Especially with more complex shaped buildings the details are very important. A example of such type of buildings are mills. To show how hard this kind of complex buildings are for computers to work with, and how easy it is for us humans to recognise the type of building are below two mills with a photo representation (to show how the buildings look in real life) compared to reconstructed 3D building and its corresponding point cloud are below figure 14 and 15. The data used are from 3DBAG made by the TU Delft (TUDelft3D, 2023) and the point cloud of AHN3 which is used to create the 3DBAG.

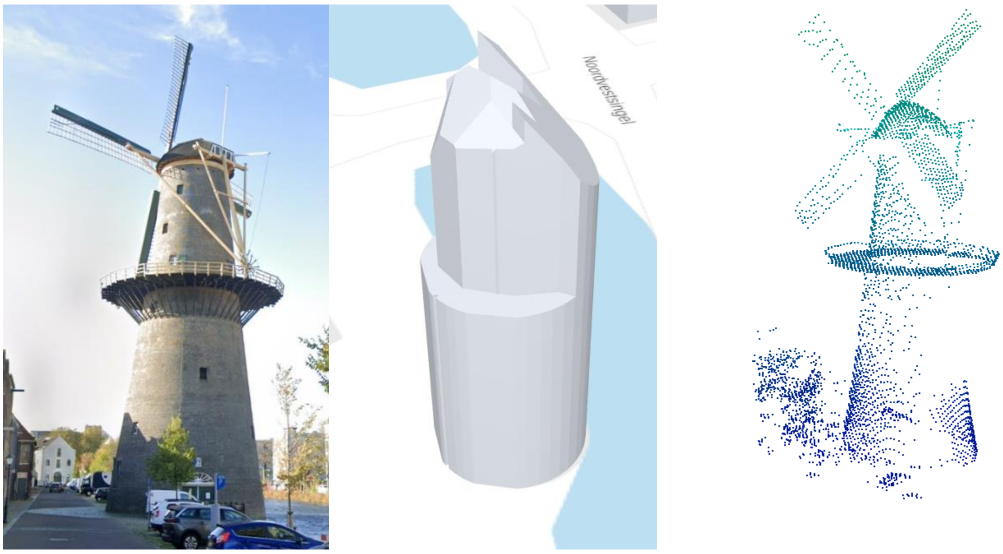


Figure 14 Mill in Schiedam represented as picture, 3DBAG object and corresponding AHN3 point cloud

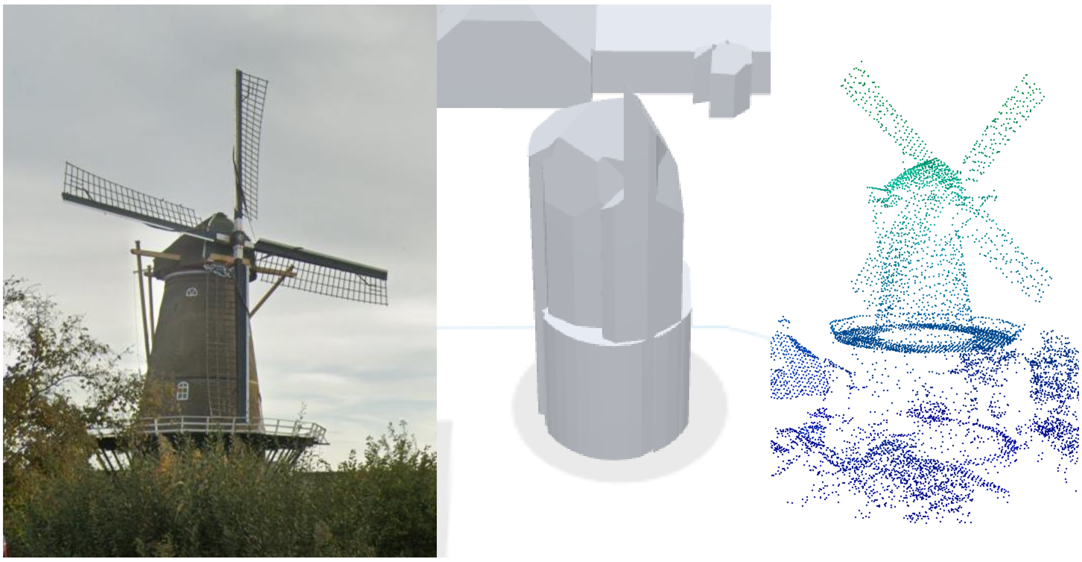


Figure 15 Mill in Bergambacht represented as picture, 3DBAG object and corresponding AHN3 point cloud

In these figures (14 and 15) can be seen that the AHN3 point clouds are very recognizable as mills to the Human Vision System. The blades of the mills are clearly recognisable. However, when a computer tries to apply algorithms that, such as Elberink & Vosselman (2011) mentioned earlier, apply predefined assumptions, it will give a very unrealistic 3D representation which is not even close to the real entity. The blades of the mill are missing and the balcony around the mill cannot be seen in the 3D object. It can be assumed that if the point cloud would have been incomplete, for example due to occlusion, that it would have assumably have a big impact on complex structures due to the error when filling in the gaps by algorithms.

2.4 Importance Key Performance Aspects for Human Visual System and Computer Algorithms

The first research question was *"What are the key performance aspects of a point cloud that affect the perceived quality for the Human Vision System, and how does this affect compare to the perceived quality for Computer Algorithms?"*.

There is not that much literature previously written about key performance aspects. However, it can be stated that Point Density is a key performance aspect for the perceived quality of a point cloud for both Human Vision System and Computer algorithms. Where the perceived quality of point clouds is linearly correlated for the Human Vision System according to Zhang et al. (2014), it is not the case for Computers algorithms. For computer algorithms Point Density is especially important to increase the level of detail, but more so the consistency of the Point Density.

For object recognition in a Computer Algorithm is the completeness of the point cloud more important than Point Density in order to fit the algorithms used to derive objects out of the unstructured point cloud. When the completeness or Point Density is not high enough to fit the algorithm, or the algorithm itself has to generalized assumptions computers will have lots of difficulties to derive 3D buildings, as shown in figure 14 and 15.

In this research the completeness of a point cloud will be researched in two ways. First will be looked at gaps in the point cloud. Secondly will be looked at the point coverage on different elements of a building.

2.4.1 Importance point coverage on walls and facades for 3D representation

With the increasing point density of laser scanners and the rising resolution of digital imagery, data-driven reconstruction techniques now yield more accurate and robust models. Smaller roof details can be detected, and geometric constraints can be applied with greater confidence. However, certain challenges persist, and new laser scanners have not completely addressed the issue of areas lacking data due to occlusion (Oude Elberink, 2008).

In Figure 6, it can be observed how this problem predominantly affects the vertical areas, specifically the walls and facades of buildings. The accurate derivation of walls is crucial for applications such as 3D building modeling, and a notable drawback of (airborne) laser scanning is the incomplete representation of walls (Rutzinger et al., 2009).

Walls are typically relatively homogeneous, forming vertical flat surfaces. In most cases, a high point coverage on walls is not essential. However, there are applications where walls, especially facades, play a vital role. Details on the facades of a building, such as bay windows and door frames, are crucial for achieving a higher level of detail in representing the building. While these elements may not be essential for recognizing a 3D representation of a building, they are important for the Human Visual System to identify the specific building being represented. For computer algorithms, the edges of surfaces are vital for understanding sharp angles. For instance, if a facade has only a few points on a bay window, it might lead the algorithm to interpret the entire wall as being inclined rather than recognizing the small bay window protruding at a 90-degree angle.

2.4.2 Quantification of the Key Performance Aspects

Point Density and Completeness are crucial Key Performance Aspects that significantly influence the qualitative perception of a point cloud, albeit in different ways for the Human Vision System and Computer Algorithms. In this paragraph will be given a short summary on the importance of the Key Performance Aspects for both the Human Vision System and Computer Algorithms.

In figure 16 you can see an overview of all the key performance aspects and how important they are for either the Human Vision System or Computer Algorithms. In this figure goes from "+" which means minor influence up to "++" which means that it has a major influence on the perceived quality of the point cloud.

Key Performance Aspects:	Human Vision System	Computer Algorithms
Point Density	++	+
Gaps	+	++
Completeness (Walls/Facades)	++	+
Completeness (Roof)	+	++

Figure 16 Impact Key Performance Aspects on the perceived quality of a point cloud

2.5 Key Performance Aspects in specifications AHN3 and AHN4

In this research, AHN3 and AHN4 will be compared based on the key performance aspects outlined in Figure 16. The aim is to assess what can be anticipated from these datasets in accordance with their specifications. Additionally, the research seeks to determine the extent to which it surpasses the specifications of AHN3 and AHN4.

According to the official website of Actueel Hoogtebestand Nederland, AHN3 lacks a specific point density specification. Nevertheless, it is mentioned that the point density typically falls within the range of 6 to 10 points per square meter. In contrast, AHN4 provides a point density specification ranging from 10 to 14 points per square meter. However, it is noteworthy that for the Schiphol area, the point density can reach as high as 20-24 points per square meter (AHN, 2023). It's important to note that these point densities represent averages over broad areas, and beyond this global point density, there are no specifications regarding the completeness of the point cloud.

Earlier in this literature review was already discussed what the limitations are of using solely point density as quality indication of a point cloud. It should be noted however that AHN was never intended to be used for the 3D representation of buildings. It is created for the waterboards, provinces and Rijkswaterstaat and primarily used for water(-barrier) management by these organizations for applications like floodmaps and monitoring dykes. Because the Netherlands is a relatively flat country, in most cases the landscape will not have lots of big height differences which could trigger high amounts of occlusion. So it is not strange that completeness in terms of gaps in 3D is less of a goal and therefore there are no specifications in terms of completeness.

Buildings, however, often cause occlusion, particularly taller structures in urban environments. Therefore, it is pertinent to investigate the completeness of AHN3 and AHN4 in terms of 3D representation, with a specific focus on buildings.

This research aims to extend beyond the specifications of AHN3 and AHN4. While point density is only defined globally, it fails to account for significant local variations due to its average nature. This study will examine the local point density of buildings, allowing for a detailed assessment of whether AHN3 or AHN4 exhibits a higher point density and the consistency of this density across structures. Additionally, the research will evaluate the completeness of point clouds by analyzing the presence of gaps and the distribution of points on a building, determined through the calculation of point spacing.

Furthermore, the investigation will explore point density disparities between Zuid-Holland and specific areas in Friesland. This comparative analysis aims to identify major differences in point density within the Netherlands.

3 Methodology

In the Methodology section, we will describe the approach used in this research. We will begin by giving an overview of the methodology. After that will be looked at the study area of this research. Then the datasets and software used for this research will be discussed. Finally will be looked more extensively at the methods used in this research by providing a detailed, step-by-step explanation of how the research will be conducted.

3.1 Setup of the research

The main objective of this research is to determine whether AHN4 is an improvement over AHN3, The first step is to identify the key performance aspects of a point cloud and assess whether they vary based on the application type (Human Vision System versus Computer Algorithms). This will be achieved by addressing the research question: *"What are the key performance aspects of a point cloud that affect the perceived quality for the Human Vision System, and how does this affect compare to the perceived quality for Computer Algorithms?"*. The literature review conducted in the previous chapter provides the answers to this research question.

According to the literature review, the key performance aspects of a point cloud are Point Density and completeness. Point Density is a quantitative performance aspect, which will be investigated through a GIS analysis using a point on area overlay. The completeness, including the distribution of points and the presence of gaps, will also be examined by calculating the Point Spacing using a GIS analysis.

This last analysis involves first converting the point clouds of the buildings into a mesh. The number of triangles and their area provide an indication of the completeness of the point cloud. If the area of a triangle exceeds a certain threshold, it can be identified as a gap. The slope of the triangles can indicate whether they represent a wall, sloped roof, or flat roof.

As previously discussed, there are concerns regarding the alleged decrease in flight hours for acquiring AHN4 compared to AHN3. To address this, the distance between a building and the nearest flight line will be calculated. The impact of this distance on the key performance aspects, such as Point Density and completeness, will be examined by searching for a correlation. Therefore, this analysis combines both GIS and statistical analyses.

Once these questions are answered, it will be possible to evaluate the performance of AHN3 compared to AHN4 in representing 3D buildings. The impact of this performance will be assessed for both the Human Vision System and Computer Algorithms. The methods behind every analysis will first be discussed with the parameters that are used in conducting the analysis. After that will be shown how these methods are applied in the FME workspaces in this research.

3.1.1 Study Area

For this research the study area will be limited by the amount of data that needs to be processed for this research. It is not feasible to study the entire Netherlands in the given time with the equipment at hand. Therefore a single province of the Netherlands is chosen to research. Since this research is supervised by the TU Delft the province of choice is Zuid-Holland. Due to time constraints only a selection of 30 municipalities (out of 50) is analyzed in the completeness analysis. These 30 municipalities are shown figure 15 as green where the rest of the municipalities in Zuid-Holland are shown in orange. The province of Zuid-Holland, that is displayed in figure 17, has an area of around 3300 square kilometers (including +/- 600 square kilometers water) and consists around 3,8 million inhabitants (Centraal Bureau voor de Statistiek, 2023).

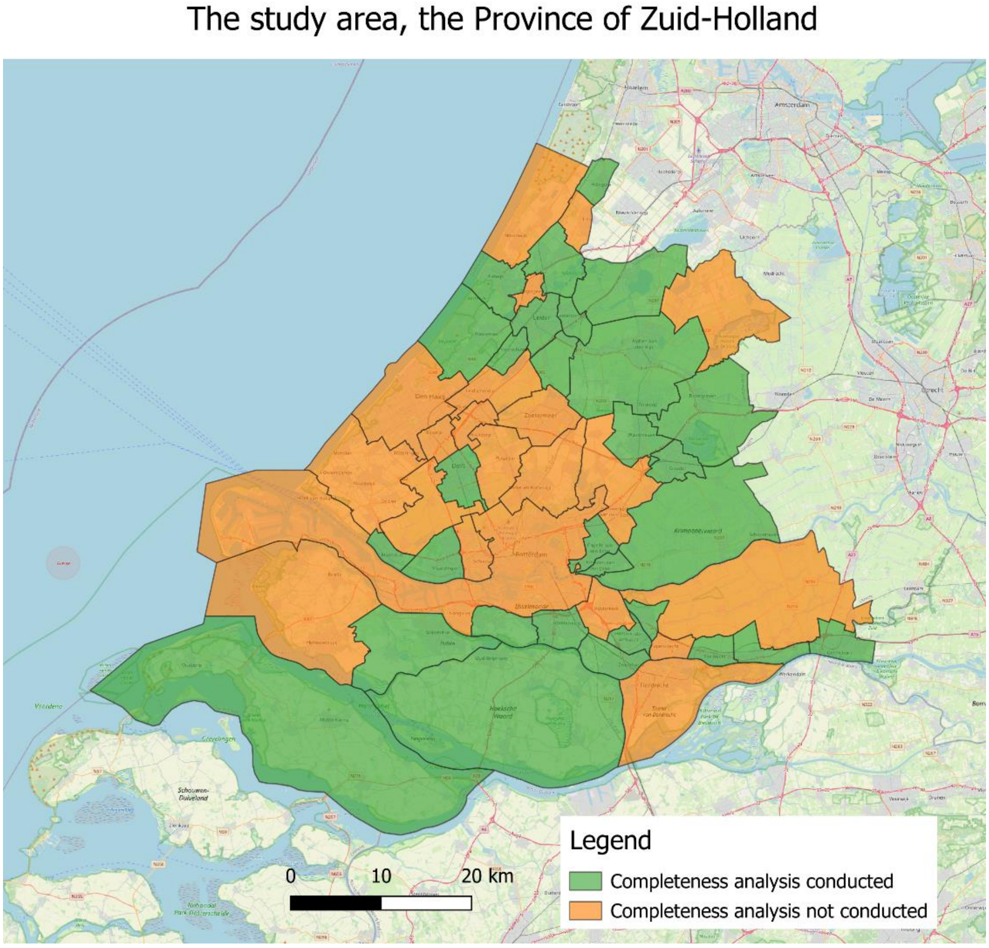


Figure 17 Map of the Province Zuid-Holland with the analyzed municipalities being green

3.1.2 The datasets

The primary dataset in this research is the Algemeen Hoogtebestand Nederland (AHN), which has been introduced in the problem statement. The research will focus on the point clouds of AHN, specifically by comparing the AHN3 point cloud with the AHN4 point cloud. The AHN3 dataset can be downloaded from PDOK as tiles, while the AHN4 dataset was not available on PDOK yet in December, but could be accessed through the Esri data portal using the same tile names as AHN3. As of January 1st, the entire Netherlands is covered by the AHN4 dataset and currently AHN4 can also be downloaded from PDOK.

As mentioned in the problem statement there are concerns about the (alleged) decline of flightlines in AHN4 in comparison of AHN3. In order to analyse this, it is needed to have the flightlines of both AHN3 and AHN4. These lines are not publicly available. However, these datasets are acquired after a mail exchange with Rijkswaterstaat (via the contact page of www.ahn.nl). The acquisition of AHN3 started in 2014 and ended on 2019. The acquisition of AHN4 was much shorter, taken from 2020 until April 2022 (AHN, 2023).

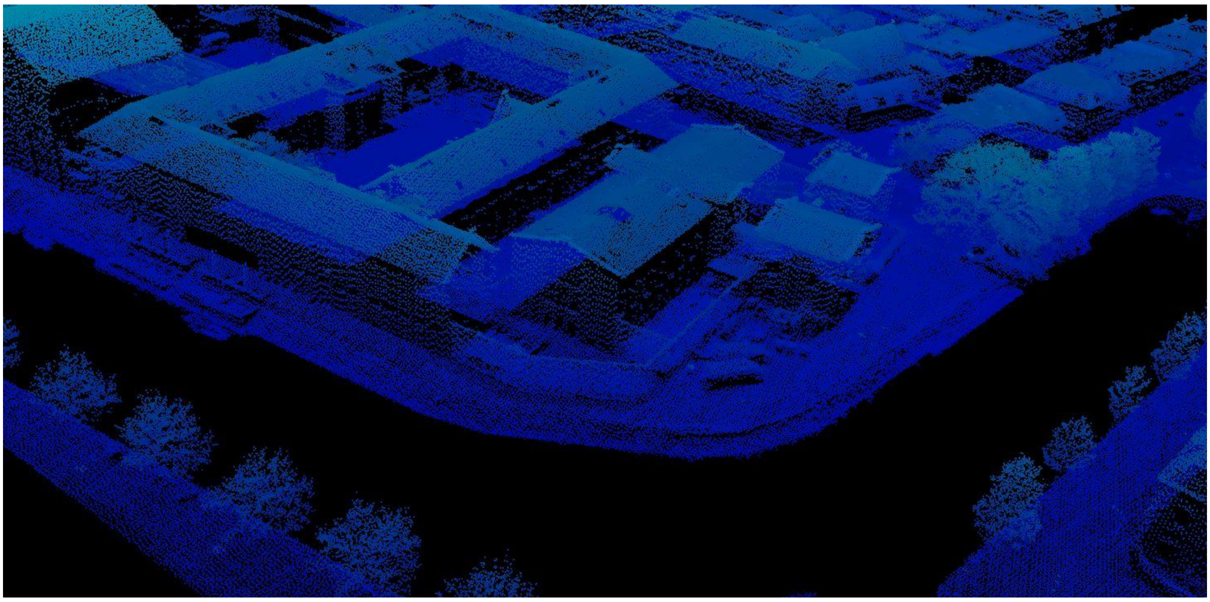


Figure 18 AHN4 point cloud

The primary focus of this research is the 3D representation of buildings as either a point cloud or as a 3D vector object. Therefore most of the analysis will therefore be conducted at a building level. To find out where buildings are located, a dataset of buildings is required, which can be found in the Basisregistratie Adressen en Gebouwen (BAG). The BAG contains all registered buildings in the Netherlands. The BAG dataset is publicly available via multiple open GIS services, and in this research, will be downloaded from the network share of the University of Utrecht.

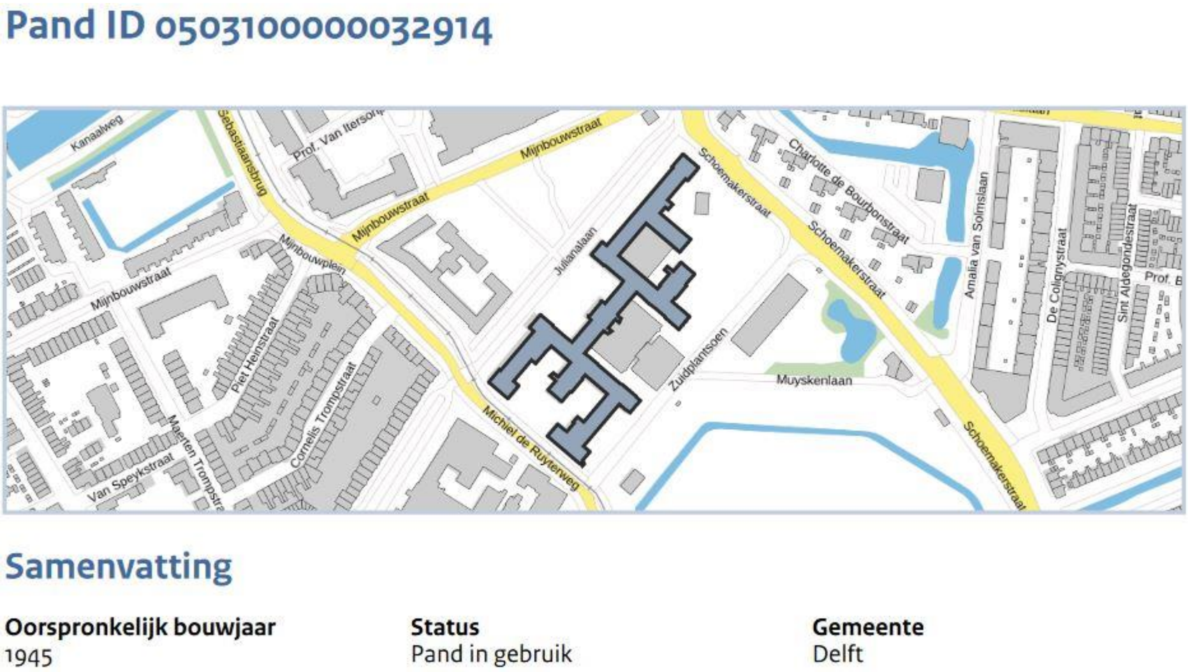


Figure 19 Building in the Basisregistratie Adressen en Gebouwen (BAG)

In order to make a good comparison between the AHN3 and AHN4 point clouds it is important that the BAG buildings which will be researched have not changed in between the acquisition dates of the AHN point clouds. Therefore the BAG data of both January the first 2014 and of November 2022 is used in this research in order to look which buildings are unchanged between the making of the two datasets.

The AHN datasets are large, making analysis on a consumer PC challenging. Although the research PC (i7 4790K CPU with 8 threads at 4.4 GHz and 32 GB of RAM) is adequate, it is advisable to divide the datasets into smaller parts. This has several advantages, including improved performance through parallel processing, which allows the 8 threads to handle smaller processes simultaneously, instead of one large process using a single thread. Additionally, if a computer crash or other disruption occurs during the analysis, the saved results can still be used, and the analysis can be resumed instead of starting from scratch. The AHN datasets will be divided into individual point clouds for each building outline (with a 0.5m buffer), then analysed and stored on a municipality-level.. The dataset that will be used for the borders of the municipalities is the “Bestuurlijke Gebieden” dataset from PDOK. This is obtainable through a Web Feature Service (WFS).

3.1.3 The software

The software that is used for most of the geographic analyses is FME. FME is an ETL (Extract Transform Load) software package made by the Canadian company Safe Software. It is specialised in handling large geographic datasets. In the case that FME is not able to do a certain analysis it can be that Python will be used. Python scripts can be added to the FME workflow.
For the statistical part of the research will JASP be used. This is an open-source statistical software program based on R supported by multiple universities including Universiteit Utrecht and Universiteit van Amsterdam.

3.1.4 Data preparation

Before conducting the analyses, it is necessary to prepare the data beforehand. This section will discuss how this is done.

3.1.4.1 Filtering the unchanged BAG buildings

In order to make a comparison between AHN3 and AHN4 it is necessary to know which buildings have not changed in the time between the acquisition of AHN3 and AHN4.

So first will be looked at what buildings have not changed in the BAG data of both January the first 2014 and of November 2022. In this dataset will be looked at what buildings are still there, with a status “in gebruik (in use)” and have not changed in geometry.

These buildings will then be given a buffer of 20 centimeters. This is done by two reasons. The first reason is simply because you want to be sure that you include the walls of the building. And there often is overhang on a building (for example a gutter) which would otherwise maybe would not be included in clipping.

The second reason of using a 20 centimeters buffer is because it allows to dissolve buildings which are connected to each other. Dissolving makes a single polygon out of all polygons that overlap. These new dissolved polygons will be handled as a single building in this research. The reason for this method is that false positive gaps on the touching area of the buildings are eliminated. For example, of you take a terraced house then it will have gaps on the sides where the next terraced house is. Dissolving eliminates this problem by considering the linked terraced houses as a single building.

FME workspace

In the FME workspace in figure 20 the municipalities are first loaded and related to the buildings of both dates by using the geometry of the municipality as a spatial query. Then is looked whether the buildings have changed by using a ChangeDetector. This is a transformer which can look whether a attribute or geometry is changed for a given attribute (in this case BAG identification).

After that the geometry is checked and a buffer is made of 0.2 meters (to make sure that the wall of the building are included when using for the clipping of the point clouds) they are being dissolved. All the unchanged buildings are then written to a OGC Geopackage for every municipality.

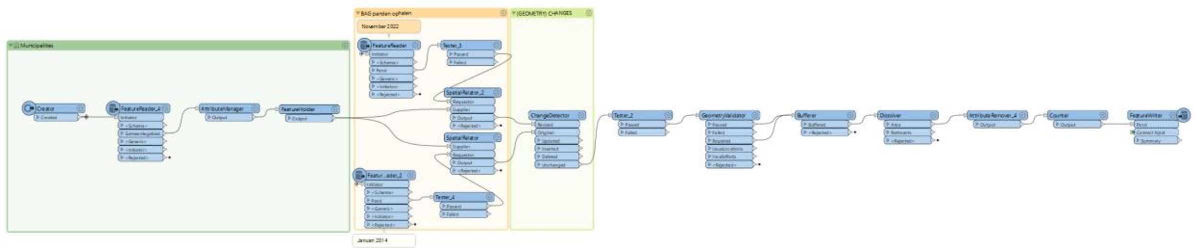


Figure 20 FME Workspace for filtering the unchanged BAG buildings

3.1.4.2 AHN download

Downloading the AHN pages is time consuming because of its size. For the study area (province of Zuid-Holland) it is 257GB for AHN3 and 882GB for AHN4. Because internet connection can sometimes fail it is wise to download the AHN pages one by one. If you would download them all at the same time and the connection fails halfway you have to start all over again but if you download them one by one you can go further where you left off. Another benefit of downloading the AHN pages one by one is that if the first page is downloaded it is possible to analyze that page immediately instead of having to wait till all pages are downloaded.

FME workspace

Because it is time consuming to trigger the download by hand, a FME workspace was made to download the AHN pages automatically. This FME workspace can be seen in figure 21. The workspace first loads the province you want to download, in this case Zuid-Holland, and then loads all the AHN pages that are located in that municipality. Then it automatically downloads the pages by filling the page name into the download link. For example for the page 31FN2, the download URL will be https://download.pdok.nl/rws/ahn3/v1_0/laz/C_**31FN2**.LAZ for AHN3.

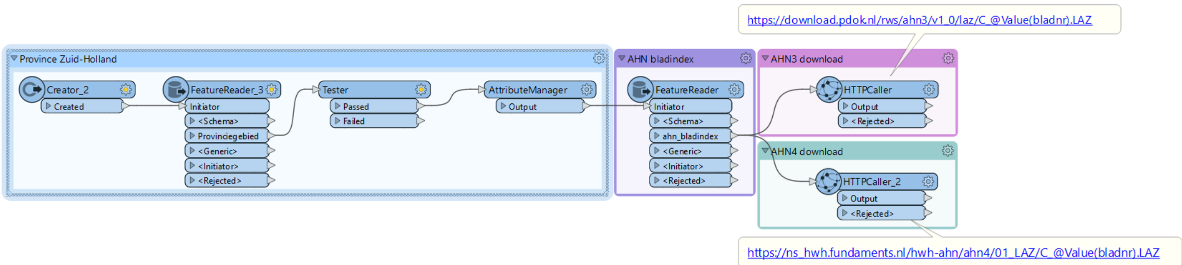


Figure 21 FME Workspace which downloads automatically all AHN3/4 pages in a province

3.1.4.3 AHN classifications

After downloading all the AHN pages needed to conduct the research, these point clouds need to get filtered. AHN has a classification in which is stated what a certain point represents. The classification classes are shown in figure 22. This research focuses on buildings so only those points are needed. Some of the points do not have a classification, because it is not safe to say what the points represents. Since this could also represent buildings this classification is also used for this research, that means that for this research the points with a classification code “1” and “6” are used.

Code	Meaning
0	never classified
1	unclassified
2	ground
3	low vegetation
4	medium vegetation
5	high vegetation
6	building
7	low points(noise)
8	reserved
9	water

Figure 22 Classifications AHN (Ledoux et al., 2020)

FME workspace
In figure 23 is the part of the FME workspace shown which loads, filters and writes the AHN point clouds. This part is within a Custom Transformer (hence the green input and output blocks) so that parallel processing can be conducted, so that multiple AHN pages can be filtered at the same time.

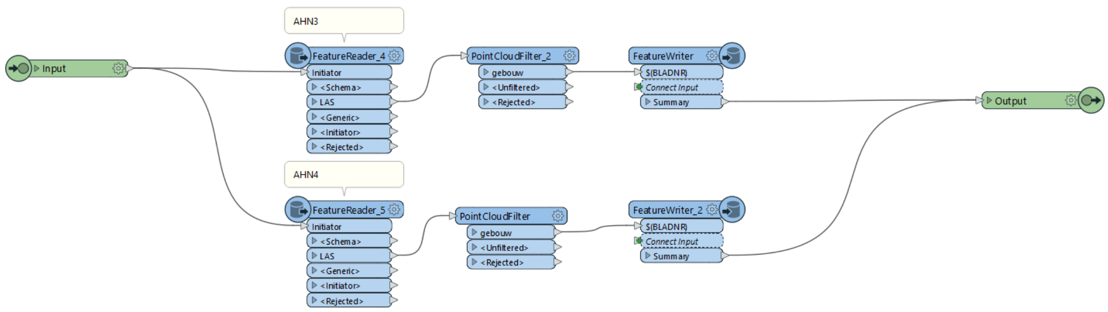


Figure 23 Part of FME Workspace which filters on classification

3.2 Point Density Analysis

The quantitative performance aspect of a point cloud is the Point Density, which refers to the number of points per unit area. This research focuses on the representation of 3D buildings, thus the Point Density will be calculated at a building-level. The sub question that needs to be answered is: *"What is the average Point Density of buildings in AHN3 and AHN4, and how do these densities compare between the two datasets?"*. A point-on-area overlay method will be applied to determine the number of points that represent each building. This method will be applied to both AHN3 and AHN4 point clouds.

FME workspace
In figure 24 the FME workspace part that overlays the LAZ files with the buildings. For the overlay the Clipper is used. This not only overlays, but clips the point cloud into smaller point clouds that represent the building in question. These are stored to be used later. The amount of points that represent a building is determined by using a VertexCounter which counts all the points in the point clouds.

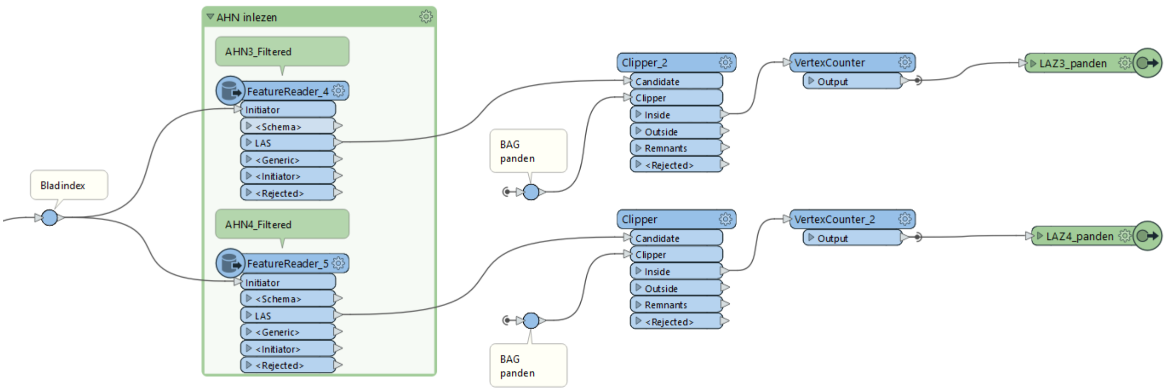


Figure 24 Part of the FME Workspace which splits the Point clouds on building level and determines the Point Density on that level

3.3 Completeness and Distribution Patterns

Point Density is quite easy to calculate and research. Due to its quantitative nature it is the case of more is better. The completeness of a point cloud however is much more difficult to determine. The completeness will be researched by looking at the distribution pattern of the point clouds. Not only the consistency of the distribution will be researched, but also the number of gaps in the point clouds will be calculated. The sub question that will have to be answered is: *"What is the Point Spacing on building elements in AHN3 and AHN4, and how do the patterns in Point Spacing compare in terms of completeness and gaps in the point clouds?"*.

3.3.1 Finding gaps in point clouds

There are multiple methods for detecting gaps in a point cloud. One of which is raster analysis, which involves counting the number of points in each grid cell. However, this method has limitations as it only considers the 2D distribution of points and may not accurately identify gaps, particularly those that are vertical in nature. For instance, a square flat rooftop may have a high density of points, but if the sides of the building have no points, it would still be considered a gap in the 3D point cloud. However, the raster analysis would not consider the vertical sides of the building and would not identify it as a gap. Therefore, this method is not suitable for detecting all types of gaps in a point cloud.

Like mentioned in the literature review, a better way of looking at the (lack of) point coverage is by looking at the point spacing. Point spacing is measures the average distance from a point to all adjacent points.

The most uniformly way of connecting adjacent points is by creating a mesh. In this research the method of making a mesh is using Delaunay triangulation. Delaunay triangulation guarantees that each point in the mesh is directly connected to its natural neighbors, i.e., the points that are closest to it. This property is beneficial when calculating point spacing because it ensures that the distances considered are between points that are geometrically close, providing a more accurate representation of the local point distribution. Delaunay triangulation ensures an even distribution of smaller triangles in areas with dense points and larger triangles in areas with fewer points. This method generates a mesh that covers the entire building, as seen in figure 25. By calculating the surface area of each triangle in 2.5D, considering the slope of the surface, it is possible to determine the Point Spacing on different parts of the building.

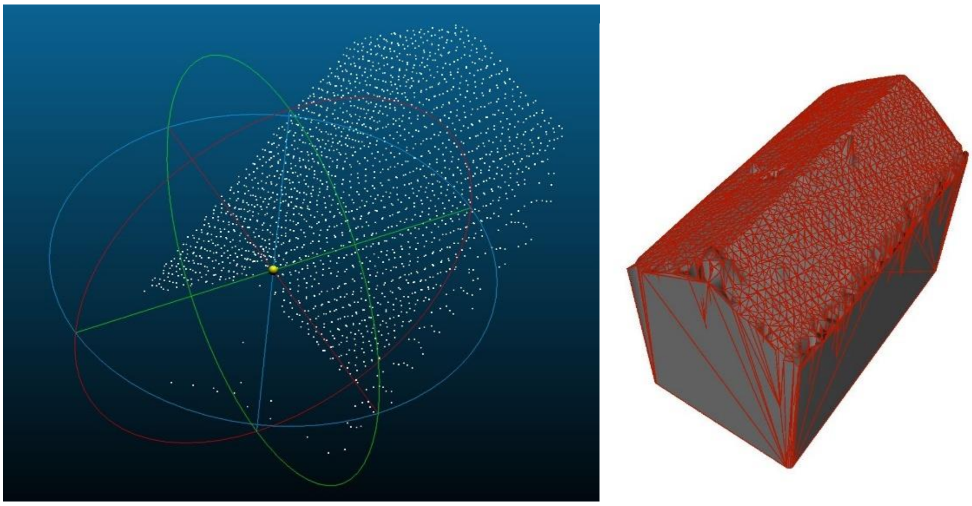


Figure 25 AHN3 point cloud of a building (left) versus corresponding Delaunay Mesh (right)

There is, however, a downside to using Delaunay Triangulation: it couples points in a 2D space, meaning that points close to each other in a 2D projection might actually be far apart in the third dimension. This poses challenges in capturing vertical surfaces uniformly because the algorithm constructs triangles based on the 2D positions of points, potentially neglecting their vertical separation. As a result, the triangulated mesh predominantly represents a 2.5D structure where horizontal relationships are well-maintained, but accurately preserving vertical surfaces becomes problematic due to the disparity in the third dimension.

Despite this limitation, it's worth noting that this issue isn't a significant problem. This is because even though the triangulation may not perfectly capture the vertical surfaces due to the 2.5D representation, the inclusion of slope in the calculations for the area of the triangle accounts for some of the 3D characteristics. In Figure 25 an extreme example is being displayed, it can be observed how the sloped roof is nicely displayed with relatively uniform triangles. However on the sides of the building many "stretched" triangles can be seen, illustrating the challenges associated with vertical surface representation. Despite the stretched triangles, it's important to note that they still represent a 3D surface and provide a good indication of point spacing within the model.

In order to convert the point cloud of a building in a mesh that goes all the way down to the ground, so that it covers the entire surface of the building, the mesh is draped down to the polygon edge from the BAG. These do not have any Z-values.

FME workspace
In figure 26 can be seen how the lowest Z-value of the point cloud is given to the polygon in order to make a solid foundation for the mesh. The TINGenerator then transforms the point cloud and the polygon into a mesh by using the Delaunay method.

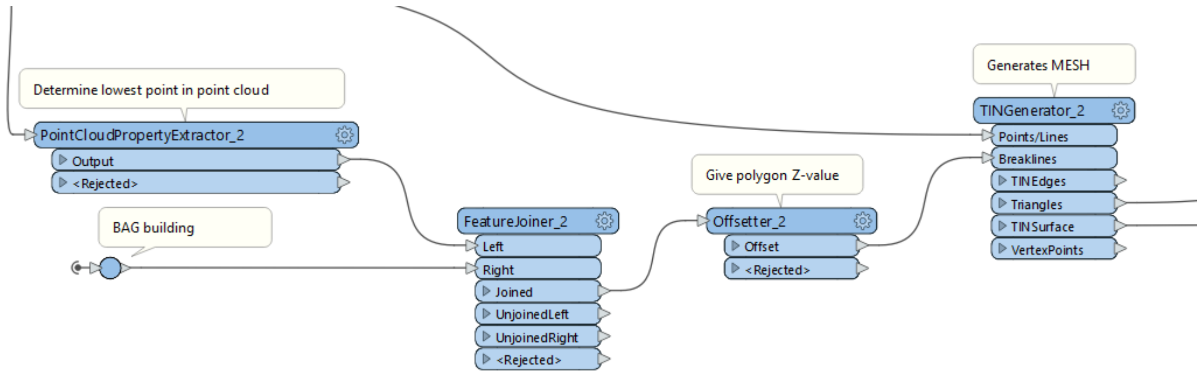


Figure 26 Part of FME workspace that generates the Delaunay Mesh

The mesh generated in Figure 26 is then used to analyse the completeness of the point cloud. In figure 27 is the FME Workspace which analyses the meshes. But just a part of it is used for finding any gaps.

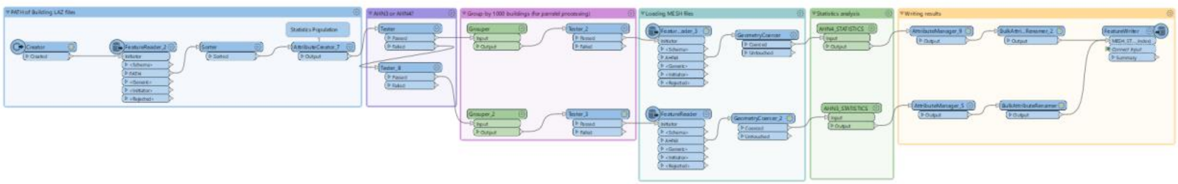


Figure 27 FME Workspace to analyse meshes on the completeness of the point clouds

The mesh needs to be split into separate triangles in order to analyse the completeness of the point cloud. Then the surface area of these triangles is calculated, this area is calculated with slope in mind so that it takes the 3D characteristics of the triangles into account. The part of the FME workspace that is responsible for this is shown in figure 28.

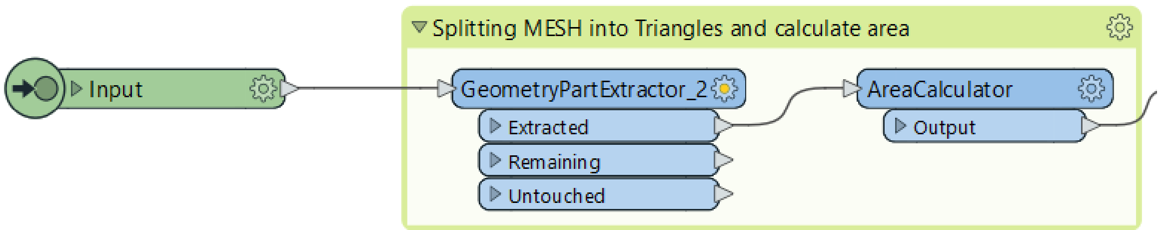


Figure 28 Splitting mesh into separate triangles and calculating the surface area.

Each triangle that has a bigger area than the threshold will be filtered and counted as shown in figure 29. It is hard to find how big a area for a gap should be, since it is arbitrary. For the threshold a minimum local Point Density of 5 points per square meters (in 3D) is chosen. In multiple sources including Rohrbach (2015) it is stated that for a basic 3D model a Point

Density of 5 points per square meters is needed. This method does mean that there are different thresholds for AHN3 and AHN4. A triangle in the mesh has a area and consists of three points, this results in a threshold of 0.6 square meter per triangle.

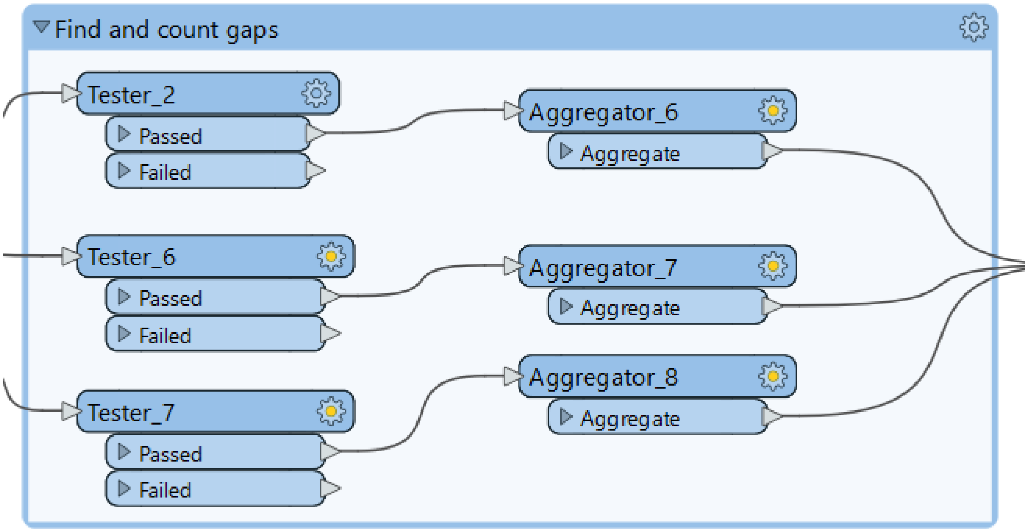


Figure 29 Filtering the triangles that are gaps

3.3.2 Distribution patterns of points on building elements

The areas of the triangles in the previously mentioned method can also be used to analyse the consistency of the Point Density in the point cloud. This can be used to evaluate the spread and standard deviation of the triangle areas. It would also be interesting to investigate the relationship between Point Density and building aspects (such as walls, sloped roofs, or flat roofs) in the point clouds. This could potentially reveal differences in point distribution.

By analysing the vertical orientation (slope) of triangles on a building, it is possible to determine which parts of the building they represent. For example, horizontal orientations would likely indicate walls, oblique orientations sloped roofs, and vertical orientations flat roofs. However, this method is relatively simplistic and may not always produce very accurate results, it could give a good indication when enough buildings would be analysed due to the law of large numbers. This is a statistical principle that states that the larger the sample of objects on which you perform a measurement, the greater the chance that the average outcome of the measurement will be more accurate. This is because any errors made when measuring a limited number of objects will spread out over a larger sample, causing the errors to 'cancel out' and the results to be more accurate. Since this research makes use of a larger sample of objects this seems to be the best way to determine the distribution of point on a building.

There are alternative methods for determining the distribution of points on a building. One other method is using a raster grid and looking how many points there are in a cell with many different Z values, this could indicate a wall (many points more or less above each other). However, this is very unreliable and uses the amount of points as a indication whether it is a wall or a roof. This makes it a unsuitable method to determine how the amount of points for a certain face of a building has changed. Also, it does not take into account the Point Spacing, or the area that the point need to be representing. Therefore this method is not used.

When a triangle has a slope between 85 and 90 degrees it is representing a wall. If a triangle has a slope of less than 5 degrees it is assumed to be a flat roof. If a triangle has a slope between 5 and 85 degrees then it is assumed to be a sloped roof.

FME workspace
In figure 30 is the part of the FME Workspace that calculated the slope for each triangle of a mesh.

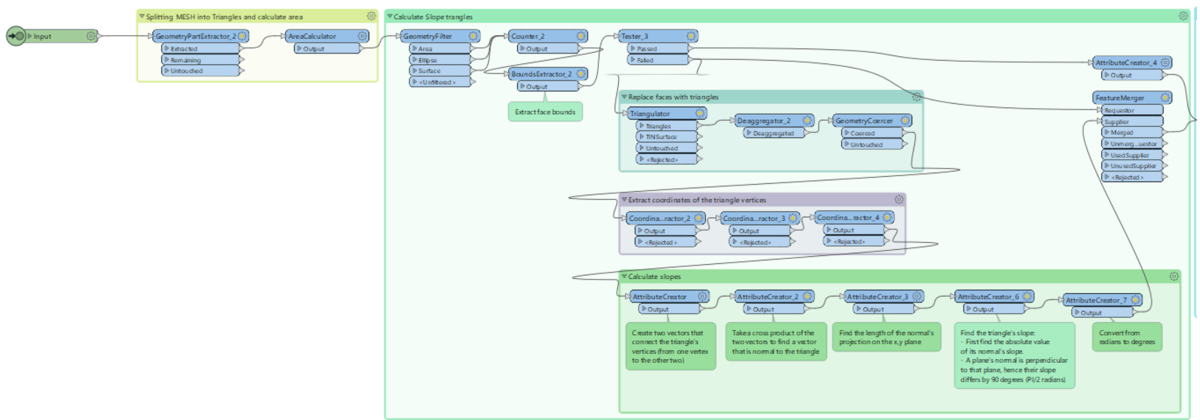


Figure 30 Part of FME Workspace that calculates the slope of all triangles in mesh

3.4 Flight lines analysis

3.4.1 Analysis on the distance to flight lines

The last sub question that needs to be answered is*: "How do the quantitative and qualitative differences between AHN4 and AHN3 relate to the distance from a building to the nearest flight line?".* To answer this sub question there will be looked at how far the flight lines of AHN3 and AHN4 are from a building and what the effect is on the completeness and density of the point clouds.

The flight lines for AHN4 have highly consistent metadata, including the altitude flown, the laser scanner model used, and the maximum scanning angle. This allows for accurate determination of any overlap in the scanned areas. The metadata for the flight lines of AHN3 is however inconsistent, with multiple scanner models used but often lacking information about which model was used in a specific area. Additionally, flight altitude is not recorded. Some areas do have information about the scanned area, but this is limited.

Ideally you would want to look at the scanning area and for any building in the scanning area you would calculate the distance (and preferably angle) from the building to the flight line. In areas with overlap this would result in two distances and angles. However, this analysis cannot be conducted for the entire province of Zuid-Holland because some AHN3 flight line metadata is unavailable. Therefore, the sub-question will be answered by calculating the distance from each building in the BAG dataset to the nearest flight line, using a shortest path analysis.

The Shortest Path analysis is a method that is utilized to find the most efficient route between two nodes in a graph, minimizing the total cost or distance travelled. In the context of geographic information systems (GIS), this method is widely applied to calculate distances between spatial features, such as lines and polygons. When the cost between features is fixed at zero, a Shortest Path analysis coincides with the direct geometric distance between the objects, whether they are points, lines, or polygons.

The distance attribute will be evaluated statistically to investigate whether it has a significant correlation with previous results, specifically Point Density, Completeness, and the presence of gaps in the point cloud.

Additionally, a case study will be conducted in a smaller study area where sufficient AHN3 metadata is available to determine the scanned area. In this area, all flight lines that can scan a building will be examined, and by calculating their distance, it will provide further insight into whether the alleged decrease in flight lines in AHN4 has an impact on the completeness and Point Density, in particularly in the overlap area.

FME workspace

In figure 31 is the FME workspace that calculates the distance to the nearest flight line for both AHN3 and AHN4. Also the Point Density from figure 24 is stored in the results of this FME Workspace. The Shortest Path Analysis

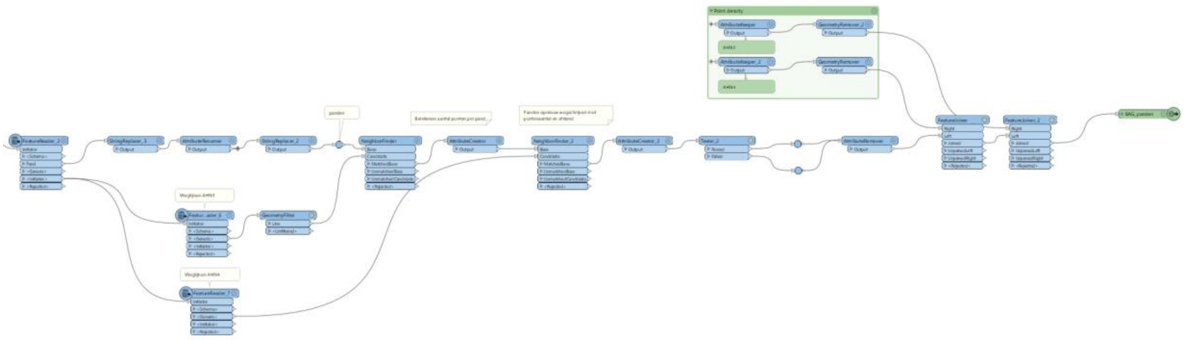


Figure 31 Calculating the distance to the nearest flight line for both AHN3 and AHN4.

Correlation of distance to flight lines with performance aspects

Since there are so many triangles (for just the municipality of Delft there are almost 125 million triangles) the data cannot easily be handled on a triangle level by most statistical programs, including JASP which is based on R which means that it will load all the data into the memory before doing the analyses. This leads to very quick results when the sample size is not that big but in this case the 32GB ram in the PC of the researcher was not sufficient. Therefore the resulting data will be aggregated on building level.

FME Workspace

The workspace in figure 32 combines all the data about the triangles, and calculates averages. The resulting statistics include:

- Average area (wall, sloped roof, flat roof, gaps and the total average area).
- Standard Deviation of the area (wall, sloped roof, flat roof, gaps and total)
- Number of triangles (wall, sloped roof, flat roof, gaps and the total number of

triangles).
- Percentage of area (wall, sloped roof, flat roof and gaps)
- Percentage of the triangles (wall, sloped roof, flat roof and gaps)

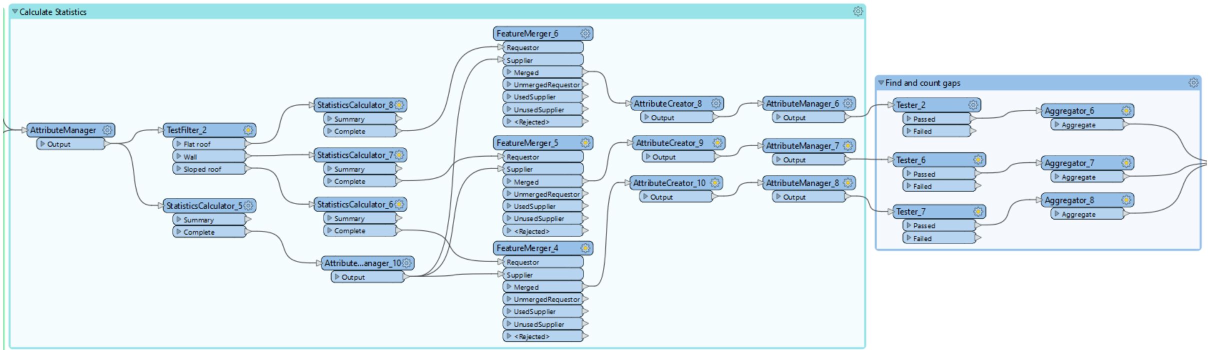


Figure 32 Deriving statistics from analyzing the slope, area and amount of triangles.

3.4.2 Exploring Scan Area Overlaps Between Flights: A Case Study Analysis

This case study will look more thoroughly at the flight methods and the impact it has on the resulting point cloud representation of buildings. The flight lines dataset that was provided was very limited however. AHN4 was consistent with the flight height, type of scanner data as polylines. The AHN3 data however had only the required data in some areas. Most of the Netherlands, including Zuid-Holland, only had the flight lines (polygon lines) without additional data such as flight height or type of scanner. This case study will look at some areas that do have enough data to analyse the amount of overlap and the affects it has on the corresponding point clouds.
The goal of this case study is to find out find out whether there are different types of flight methods used in AHN3 and AHN4, or even within the same dataset. Also what impact the amount of overlap has on the quality of the 3D representation of buildings will be researched. Also will be looked how the quality of the point clouds are in Zuid-Holland, in the Randstad, compared to more rural regions of the Netherlands.

Analysis of overlaps

The areas which are suitable for this case study is limited to the regions where there is sufficient flight data accessible for AHN3. The regions where there was sufficient data can be seen in figure 33.

Regions where AHN3 scanning area data is available

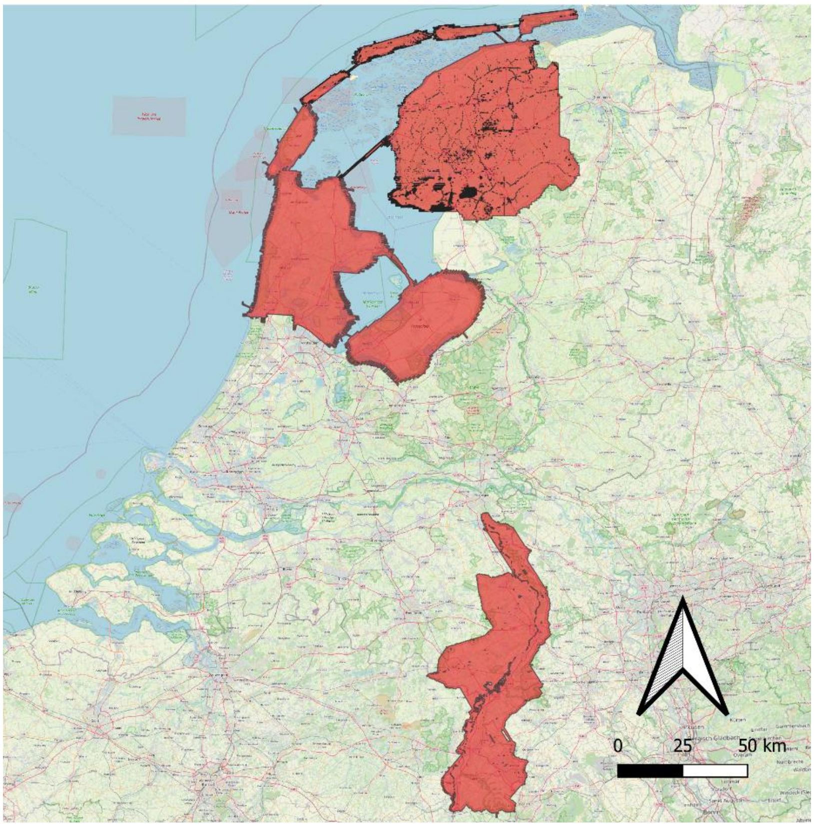


Figure 33 Regions with information scanning area AHN3

These areas are the provinces of Limburg, Friesland, Flevoland and a large part of Noord-Holland. In some of these areas it can be seen that there were extra flights conducted for AHN3, probably because the Point Density was insufficient in these sectors. These areas have lots of overlaps which could decrease the amount of occlusion in these sectors. Therefore is chosen to research an area where there are no additional flights and an area where there were additional flights.

For the case study areas is chosen to use the two AHN pages with different flight patterns. The AHN pages that will be researched are 06cz1 and 10hn1. 06cz1 covers mostly the city of Leeuwarden and for a large part there is an additional flight flown, as can be seen in de coverage area map of figure 34. 10hn1 covers mostly the city of Sneek, here no additional flights are conducted and can be seen in figure 35.

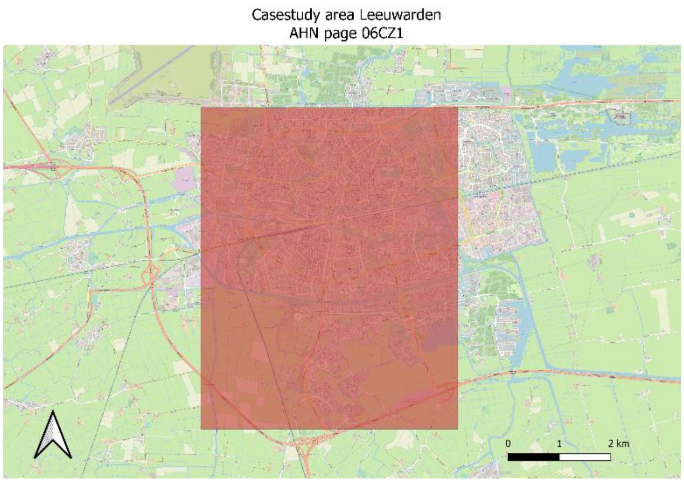


Figure 34 Study area region Leeuwarden

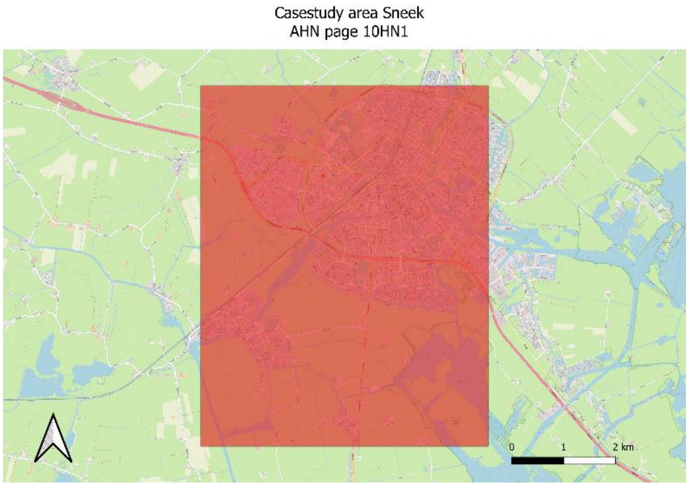


Figure 35 Study area region Sneek

In contrast to AHN3 does AHN4 not have any data for the scanning area, just polylines. Yet the scanning area can be calculated because of the fact that the flight height is known (1402 meters) and the scanning angle is known (58 degrees). If you apply geometry calculations then this results in an scanning area of 777,14 meters to each side, which is the buffer that needs to be used for the scanning area.

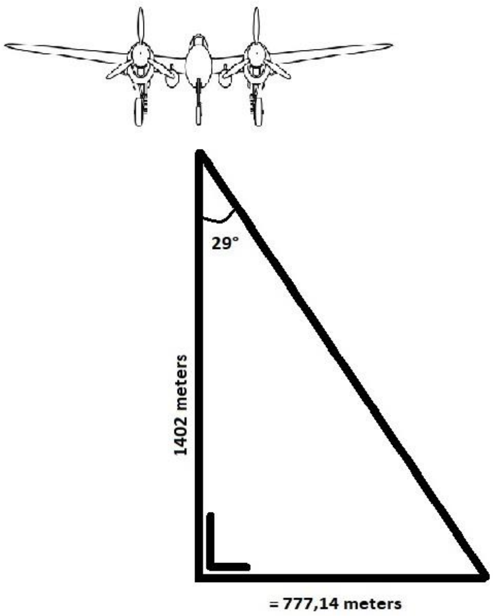


Figure 36 Scanning area from flight height and scanning angle

Analysing the overlaps in these two regions is done by first clipping the boundary of the AHN box and clipping the scanning areas with it. After that a Polygons and Polygons Intersection is being used to get a dataset with for every location how many scans there are used for that specific area.

FME workspace

The FME workspace that calculated the overlaps of the scanning areas can be seen in figure 37. For the Polygon and Polygon Intersection the AreaOnAreaOverlayer transformer is being used. This performs the intersection on all features that go into the transformer, in this case all scanning areas.

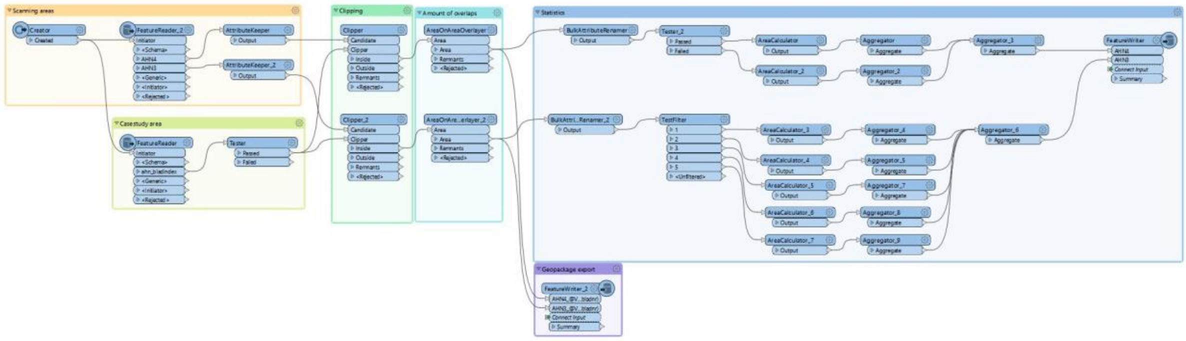


Figure 37 The FME Workspace that calculates the amount of overlaps in the case study

Analysis of the impact of overlaps on quality representation building

In these two regions in Friesland the amount of overlaps of scanning areas there are for a building will be investigated. The buildings which are partly in separate scanning areas in AHN3 or AHN4 will be filtered out. Out of the resulting buildings four will be chosen in each area at random, so four buildings in the Leeuwarden area and four buildings in the Sneek area. These buildings will then be analysed more deeply with the human eye on:

- How high is the Point Density?
- How are the points distributed?
- Are there any differences in the 3D representation of the buildings?
- Is there a relation visible between the amounts of overlaps and the quality of the 3D representation?

Comparing case study area with Zuid-Holland

In order to compare the case study area in Friesland with the study area in Zuid-Holland the same analyses are used as in Zuid-Holland. This means that the point clouds are clipped on a building level and then these point clouds are being transformed into a mesh. After that the triangles in the mesh area analyzed in order to know the completeness of these point clouds (amount of gaps and the distribution of points on building elements).

4 Results

4.1 Point Density

In total there are 572.963 unchanged buildings in the province of Zuid-Holland. The average Point Density per building is 4.183 point per building for AHN3 and 9.605 point per building for AHN4. This means that in AHN3 there are around 2.4 billion points representing these buildings. AHN4 however has more than twice as much at around 5.5 billion points. This higher Point Density can be seen in figure 39.

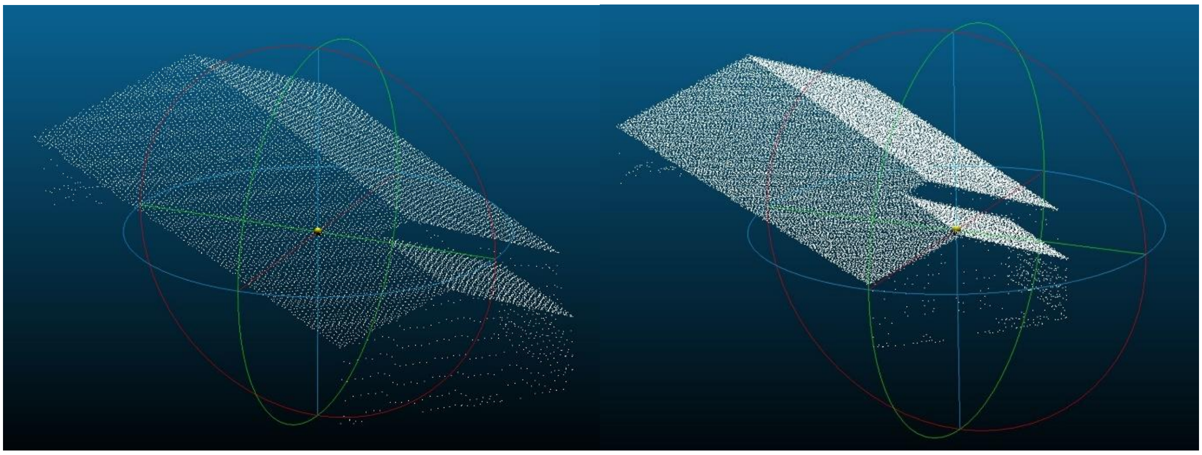


Figure 39 Building in AHN3 (left) and building in AHN4 (right) with higher Point Density

Out of all 572963 buildings there are 531.098 buildings where AHN4 has a higher Point Density compared to AHN3. 92.6% of buildings have a higher Point Density in AHN4 in comparison to AHN3. We can conclude that the AHN4 dataset indeed has much higher Point Density.

4.1.1 Point Density Zuid-Holland compared to Friesland

When comparing the average point density of Zuid-Holland with the point density of the case study area in Friesland it can be seen that the point density in both AHN3 and AHN4 are higher in Zuid-Holland than it is in Friesland as can be seen in Figure 40. There is nothing stated about this in de specifications of AHN3 and AHN4, but for AHN3 the point density is 49.5% higher in Zuid-Holland. For AHN4 it is even 69,2% higher in Zuid-Holland.

Zuid-Holland

	AHN3	AHN4
Mean	13.925	31.344
Std. Deviation	7.636	15.826

Friesland

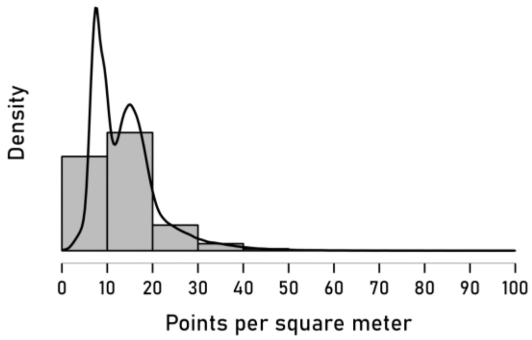
	AHN3	AHN4
Mean	9.317	18.520
Std. Deviation	3.575	6.624

Figure 40 average point density Zuid-Holland compared to Friesland

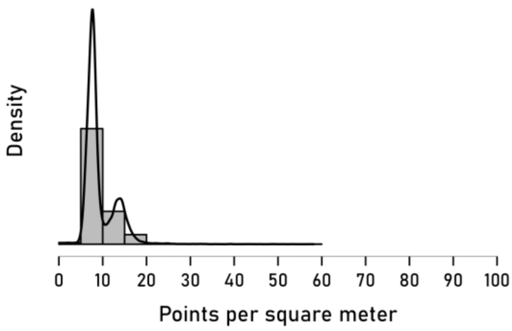
When looking at the distribution of point densities on buildings in figure 41 it can be seen that in Zuid-Holland the point density is way more inconsistent and diverse in AHN4 compared to AHN3 with a less distinct peak around a certain point density.

In Friesland it is also clear that the point density increased but way less than in Zuid-Holland. Also the distribution is way more comparable and consistent in AHN4 in Friesland compared to Zuid-Holland.

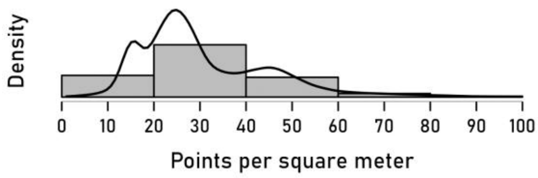
The reason why these differences are occurring is unclear but what can be concluded that there are large regional differences in point density. Also can be concluded that AHN4 has a higher point density than AHN3 in both Zuid-Holland as in Friesland but that AHN4 has more variation in point density and is less equal distributed as AHN3.



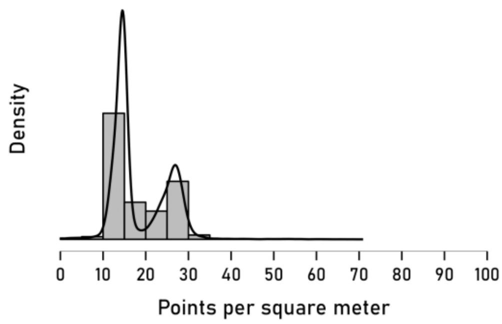
AHN3 Zuid-Holland



AHN3 Friesland



AHN4 Zuid-Holland



AHN4 Friesland

Figure 41 Point Density distribution AHN3 and AHN4 in Friesland and Zuid-Holland

4.2 Completeness

The completeness of the point clouds is analyzed for 30 municipalities in Zuid-Holland. First, some general results about the analysis of the triangles will be presented.

Across these 30 municipalities, there are a total of approximately 833 million triangles for AHN3 and nearly 1,45 billion triangles for AHN4. The triangles in the AHN3 dataset have an average area of 0.44 square meters, while the triangles in AHN4 have an average area of 0.25 square meters. This difference appears to be consistent with the variance in Point Density.

Another intriguing aspect to consider is the sum of the areas of all triangles. Since both AHN3 and AHN4 represent the same object surfaces, the total surface area should be fairly consistent. If there was a significant discrepancy, one could argue that the method employed using the triangles of a mesh to assess point cloud completeness might not be precise enough. This is not the case. The total area covered by AHN3 is 368,545,373 square meters, and for AHN4, it is 365,145,694 square meters. This difference of 0.92% is relatively small, especially given the methodology employed.

4.2.1 Amount of gaps in point cloud

The number of gaps in AHN3 is just over 109 million compared to just over 90 million in AHN4. This does mean that AHN3 has more gaps, and if you look at the total amount of triangles then the gaps are about 13% of the triangles in AHN3 compared to 6.2% in AHN4. When looking at how much surface area is a gap then it is more of the same, with 68.9% for AHN3 and 58.9% for AHN4. What can be concluded from this is that AHN4 does not have more gaps than AHN3 on average, it has significantly less.

4.2.2 Distribution of points

When looking at the distribution of point on a building it is interesting to start with the surface area. Since in both AHN3 and AHN4 the same objects (buildings) are analyzed, it should be the case that the distribution of surface area on the different elements of the building (wall, flat roof and sloped roof) are comparable. In figure 42 can be seen that that is the case with only marginal differences between the two datasets.

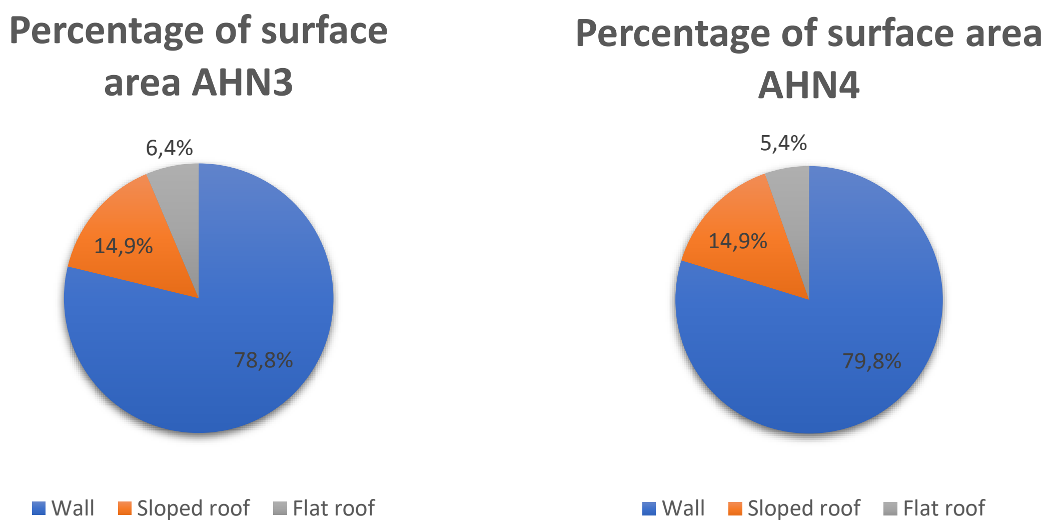


Figure 42 The distribution of surface area on buildings for AHN3 (left) and AHN4 (right)

If we then look at the percentage of triangles (and therefore points) that represent these areas the difference becomes clear as can be seen in figure 43.

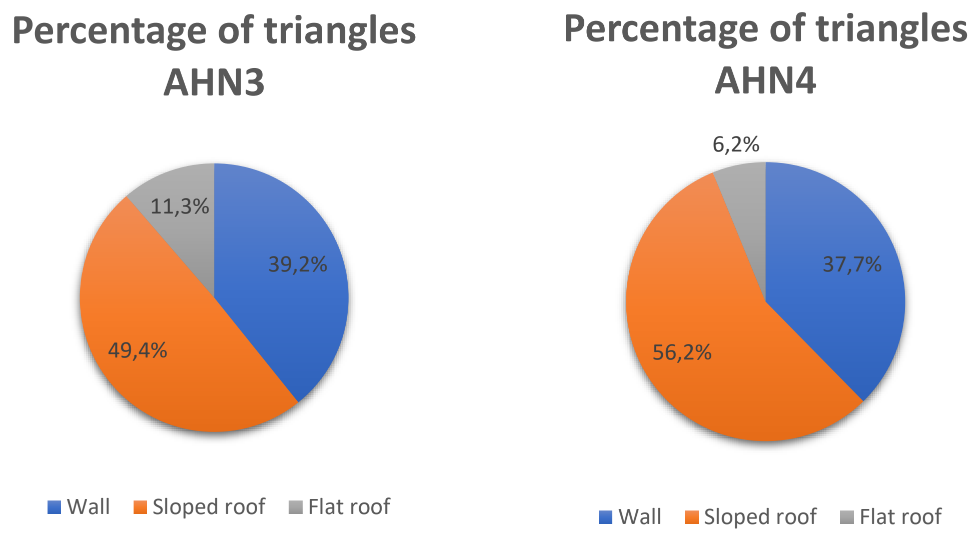


Figure 43 The distribution of triangles on buildings for AHN3 (left) and AHN4 (right).

When comparing figure 42 and figure 43 in both AHN3 and, to an even bigger extend, AHN4 the wall elements are underrepresented. About 79% of the buildings consist of walls but in AHN3 only 39,2% and in AHN4 37,7% of the triangles represents these walls. This also implies that both types of roof (sloped and flat) are over represented in both AHN3 and AHN4.

4.3 Influence of flight methods

This research also looked at the flight methods of both AHN3 and AHN4. In this chapter will be researched what the influence of the flight methods is on how well a point cloud performs. In Zuid-Holland, the analysis initially focused on the distance to the nearest flight line. This approach was chosen due to the absence of additional flight information for AHN3 in this region. Subsequently, the results of a case study are presented, where an area in Friesland was investigated. This specific area was chosen because of the detailed information about the scanning area, including the amount of overlap, was available. The research also evaluates the performance of both AHN3 and AHN4 in this Friesland case study area and compares it to their performance in Zuid-Holland.

4.3.1 Distance to flight lines

There are 572.963 buildings in the province of Zuid-Holland there are 479.475 buildings that lie closer to the closest AHN3 flight line than to the closest AHN4 flight line. This is about 83.7% of the buildings. The average distance to a AHN3 flight line is 74.7 meters with a maximum distance of 230.7 meters. For AHN4 the average flight line is 254.28 meters away, which is more than the maximum distance to a AHN3 flight line. The maximum distance to a AHN4 flight line is 626.9 meters. We can conclude with certainty that there are less flights flown for the acquisition of AHN4 compared to AHN3.

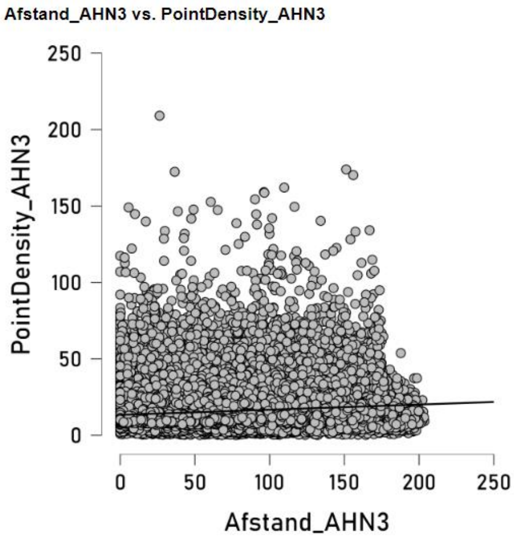
Correlation distance flight lines to the Point Density and completeness

Firstly the correlation between the Point Density and the distance to the nearest flight line was investigated.

Correlation distance to flight lines with point density AHN3

Pearson's Correlations

			Pearson's r	p
Afstand_AHN3	-	PointDensity_AHN3	0.196	< .001



Correlation distance to flight lines with point density AHN4

Pearson's Correlations

			Pearson's r	p
Afstand_AHN4	-	PointDensity_AHN4	0.201	< .001

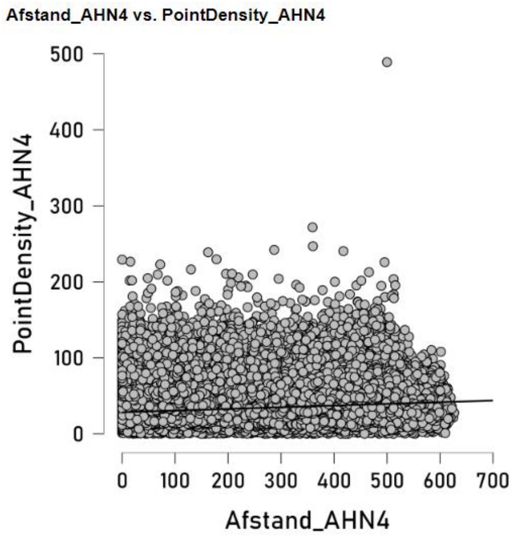


Figure 44 The correlation between distance to flight lines and Point Density

As shown in figure 44 there is a significant very weak positive correlation between distance to flight line and Point Density for both AHN3 and AHN4. This means that if the distance increases the Point Density slightly increases. This could be the case because of the fact that at a longer distance there are overlaps between different scans which results in a higher Point Density.

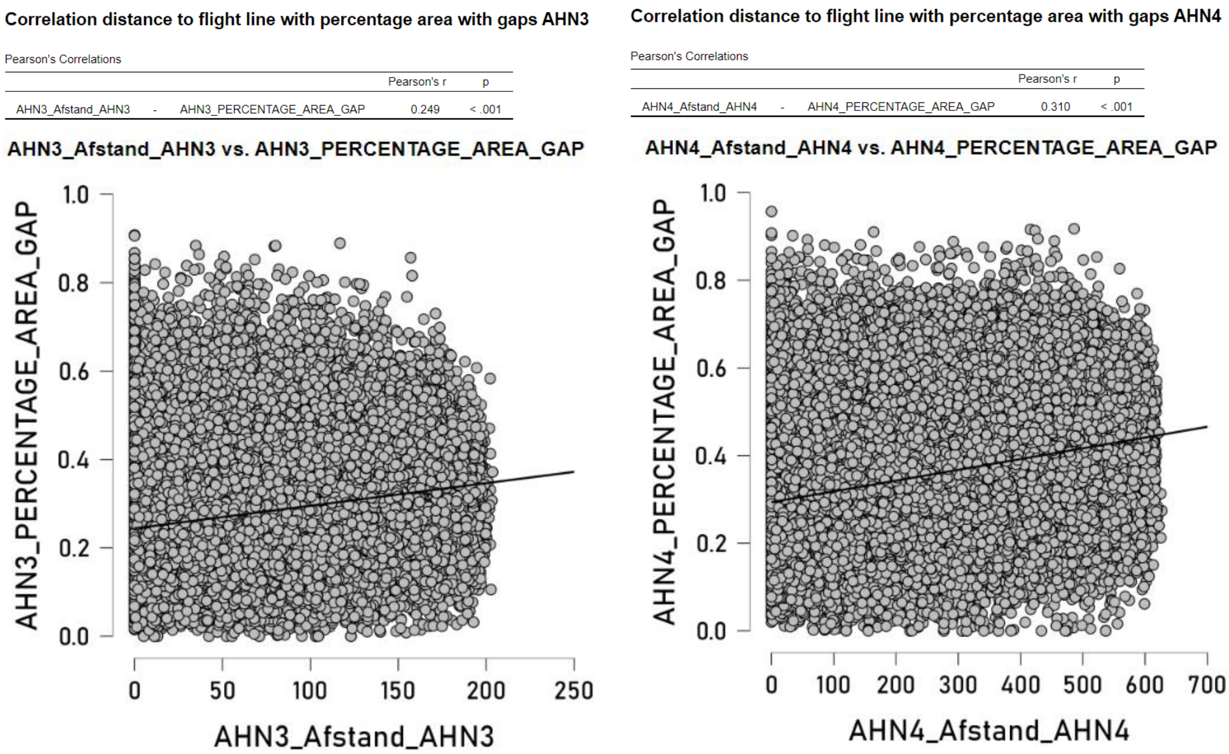


Figure 45 The correlation between distance to flight lines and area which is a gap.

The relationship between the distance to the nearest flight line and the percentage of the area with gaps in the point cloud is depicted in figure 45. It is evident that there is a noticeable yet very weak positive correlation between flight lines and the percentage of gaps in the AHN3 dataset. In the case of AHN4, this correlation is even more pronounced, indicating a significant but still weak positive correlation. These results imply that both AHN3 and AHN4 exhibit more gaps in their point clouds as the distance to the nearest flight line increases. One plausible explanation for this trend is that a larger scanning angle, which occurs when the distance to a flight line is greater, likely leads to increased occlusion.

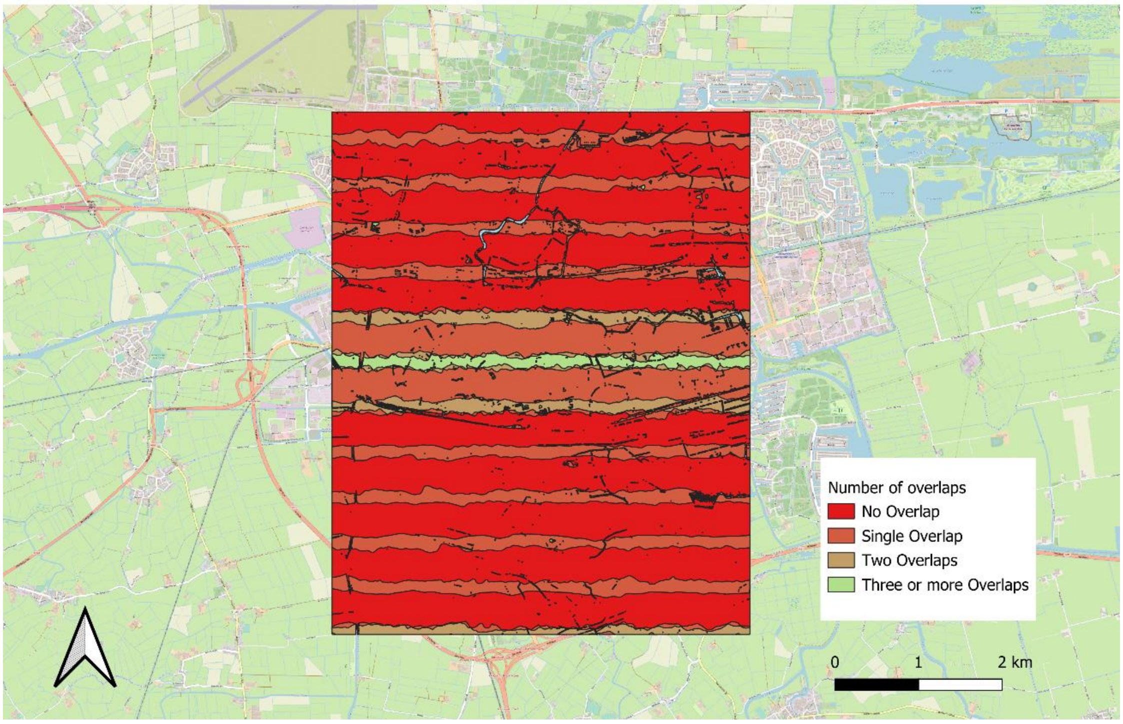
4.3.2 Case study: impact amount of overlaps on quality

The case study looks deeper in the effect of flight methods for the performance of point clouds in representing 3D buildings. Firstly the amount of overlaps for both AHN3 and AHN4 for the case study areas will be looked at. Then will be looked more closely to some buildings in these areas and whether the amount of overlaps has an impact on the quality of the 3D representation. Then also will be looked whether the case study area is comparable with the study area in Zuid-Holland.

Amount of overlaps

The amount of overlaps region around Leeuwarden is shown in figure 46. AHN4 is flown very consistently with a high coverage of area with one scan. AHN3 has smaller coverage and two extra flights where flown resulting in many overlaps in the middle of the study area.

Number of overlaps in Leeuwarden area AHN3



Number of overlaps in Leeuwarden area AHN4

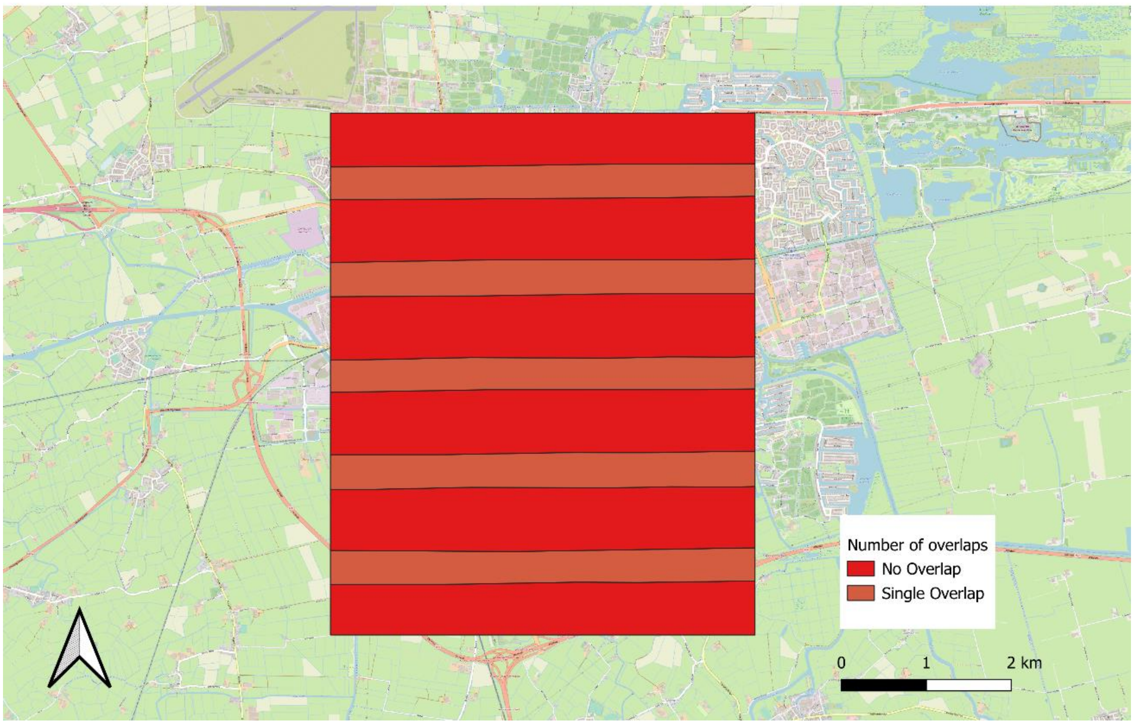
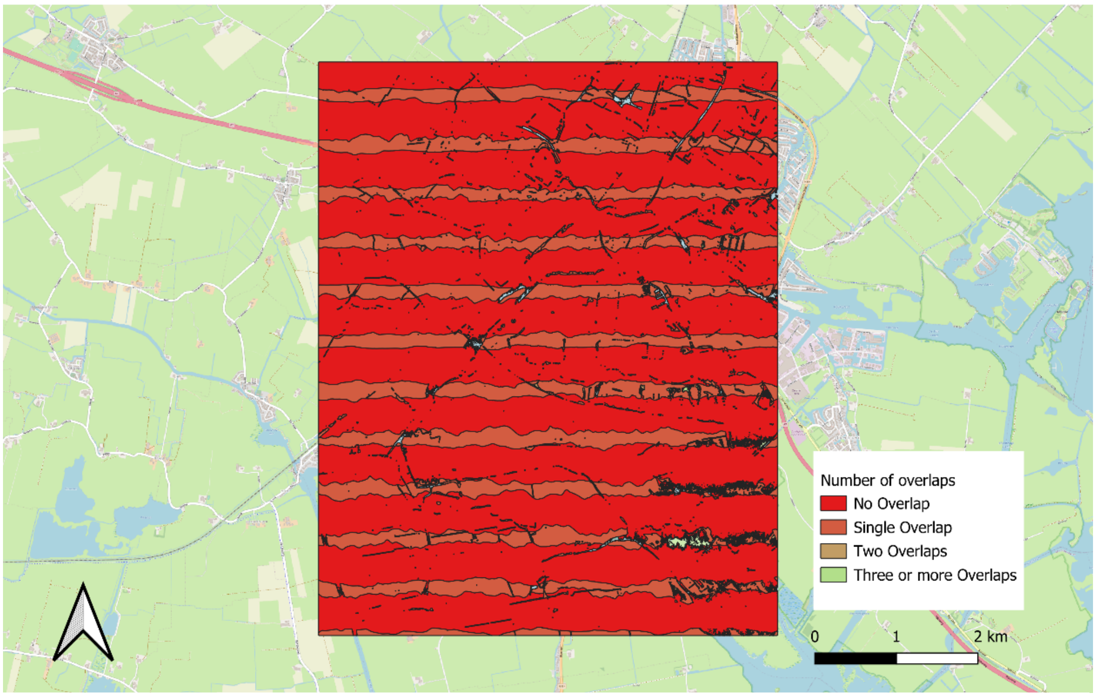


Figure 46 Leeuwarden overlaps AHN3 versus AHN

The amount of overlaps in the Sneek region can be seen in figure 47. AHN3 has no additional flights and that AHN4 is very similar to the Leeuwarden area in the scanning area pattern.

Number of overlaps in Sneek area AHN3



Number of overlaps in Sneek area AHN4

Figure 47 Sneek overlaps AHN3 versus AHN4

When examining the extent of overlaps across the entire study areas, Figure 48 reveals that the Sneek region exhibits minimal areas with two or more overlaps. Additionally, it becomes evident that AHN4 holds an advantage, as approximately 76% of the area in AHN3 is scanned only once (with no overlap), whereas this proportion decreases to 66% in the case of AHN4. This indicates that the width of the overlap did not increase with the increase in flight height which could potentially result in more occlusion in 10% of the surface.

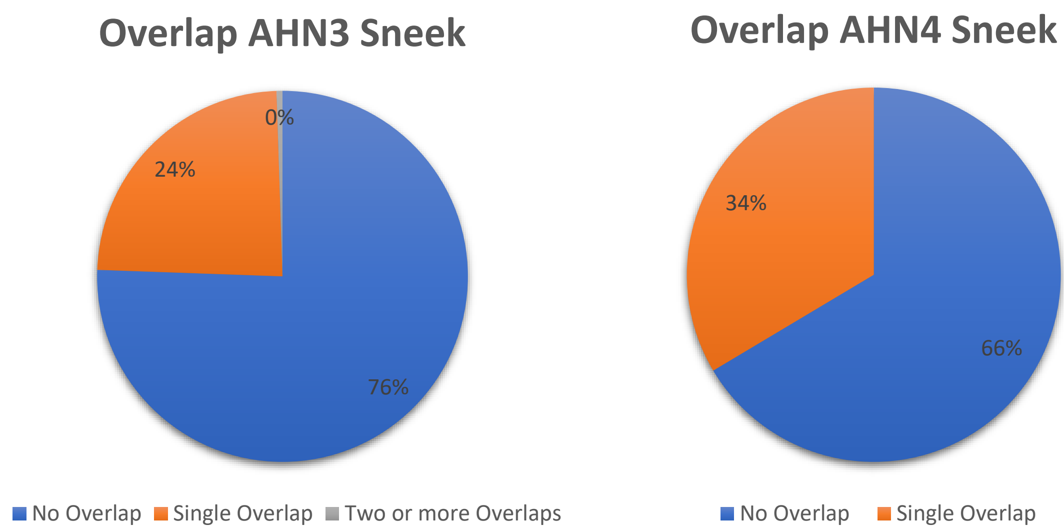


Figure 48 Amount of overlap in the Sneek region

Turning our attention to the Leeuwarden area in Figure 49, it becomes apparent that AHN4 showcases a very similar scanning distribution to AHN4. However, a notable contrast emerges when examining AHN3. In this case, a significant disparity compared to the Sneek region is evident. Here, a smaller area lacks overlap, accounting for only 60%. Additionally, the impact of the supplementary flights is observable within the distribution, contributing to a 9% coverage of the area with two or more overlaps. What is similar between Sneek and Leeuwarden however is the fact that about 8% less area has a overlap indicating that the width of overlap did not increase with the increase in flight height for AHN4.

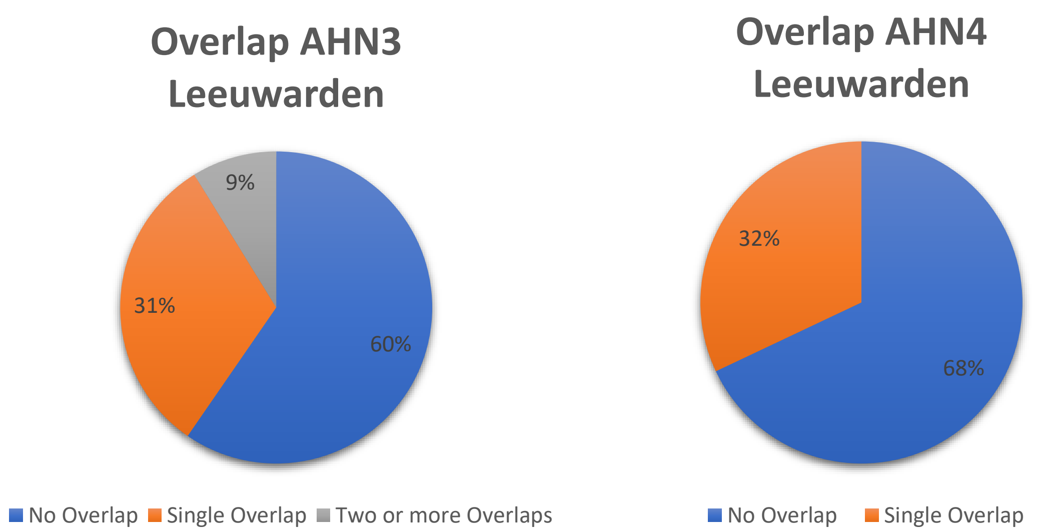


Figure 49 Amount of overlap in the Leeuwarden region

Impact amount of overlaps on quality 3D representation buildings

What the impact is of this difference between the Sneek region and the Leeuwarden region will be looked at know. As explained earlier in the method there are four buildings near Leeuwarden and four buildings near Sneek randomly picked. How many overlaps there were on that particular building for both AHN3 and AHN4 will be investigated first. The results can be found in figure 50 below, along with the amount of points on that specific building for both AHN3 and AHN4.

	Overlaps AHN3	Overlaps AHN4	Points AHN3	Points AHN4
Building Leeuwarden 1	2	0	2457	2612
Building Leeuwarden 2	3	1	10046	17577
Building Leeuwarden 3	1	1	1182	3755
Building Leeuwarden 4	3	1	6830	12521
Building Sneek 1	1	0	12921	11723
Building Sneek 2	0	0	4515	8573
Building Sneek 3	1	1	907	1684
Building Sneek 4	1	1	7895	15150

Figure 50 Overlaps buildings and the amount of points

Upon examining Building 1 in the Leeuwarden area, a slight increase in points is observed within the AHN4 representation in comparison to the AHN3 counterpart. It is noteworthy that AHN3 exhibits two instances of overlapping scanning areas, which are absent in AHN4. Despite the lower Point Density in AHN3, its representation is deemed superior due to the greater number of points along the sides of the building, contributing to a more comprehensive depiction. Notable examples include the side of the garage and the left side of the building.

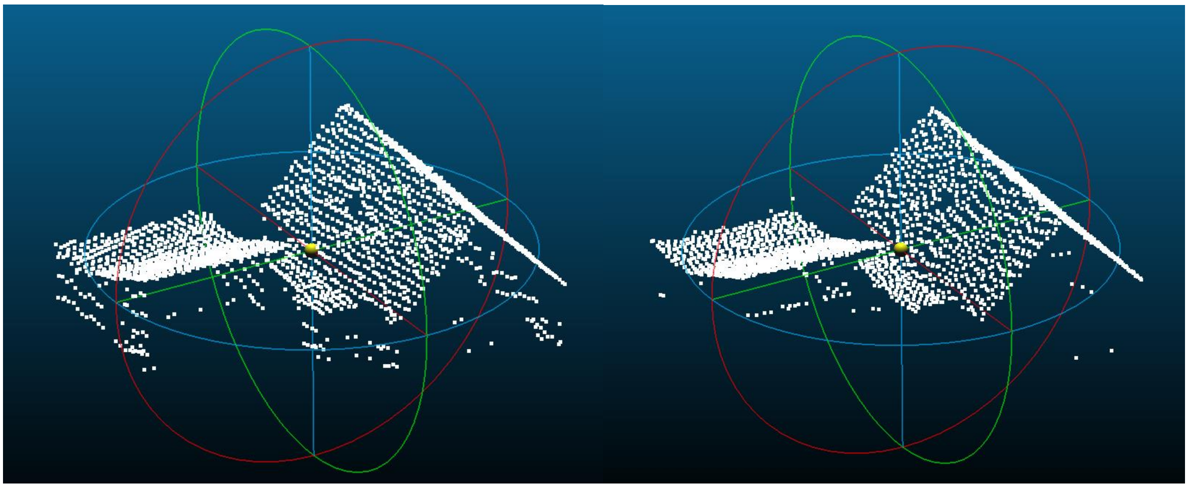


Figure 51 Building 1 Leeuwarden Pointclouds AHN3 and AHN4

Building two in the Leeuwarden region has in both AHN3 and AHN4 at least one overlap. AHN4 has a single overlap and AHN3 has three overlaps. The amount of points is a lot higher in AHN4, but with to human eye there is not that many differences between the two representations. They both do not have large gaps and both have a good 3D representation of the building.

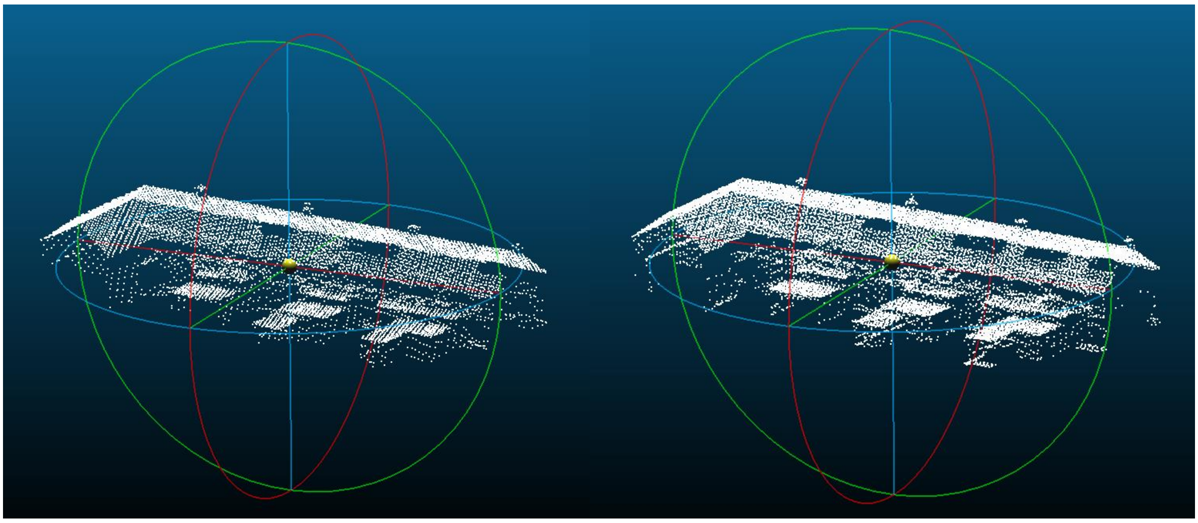


Figure 52 Building 2 Leeuwarden Pointclouds AHN3 and AHN4

For building 3 in the Leeuwarden area there is an interesting phenomenon. Both AHN3 and AHN4 have a single overlap. But the 3D representation of AHN4 is much better due to more points but also a way better distribution. AHN3 has very little points on the sides and face of the building whereas AHN4 does have points on these places.

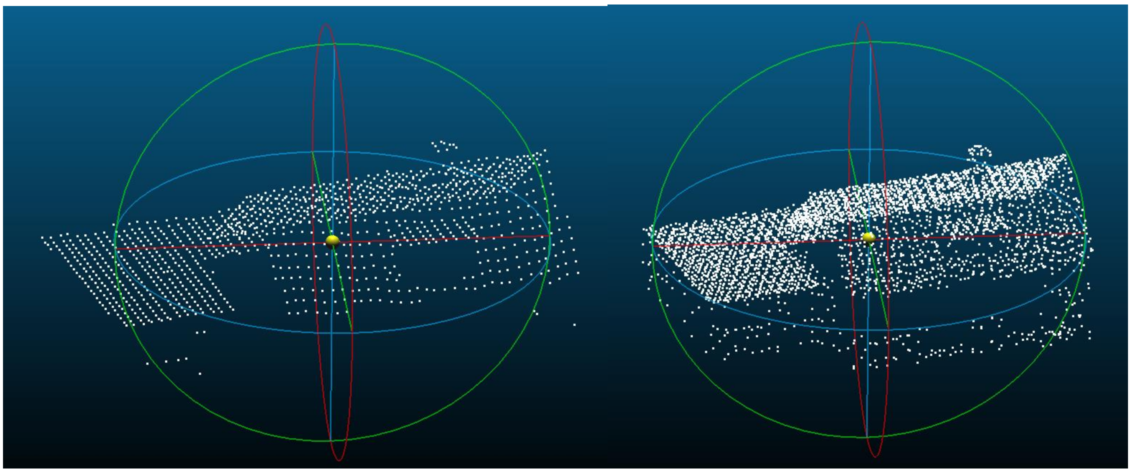


Figure 53 Building 3 Leeuwarden Pointclouds AHN3 and AHN4

Building 4 in the Leeuwarden region has three overlaps for AHN3 and a single overlap for AHN4. What is interesting with this building is that AHN4 has around twice the amount of points for representing the building. However AHN4 has gaps on the left side of the building where AHN3 does have points. So in terms of Point Density AHN4 is better but the points of AHN3 are better distributed and therefore the point cloud is more complete.

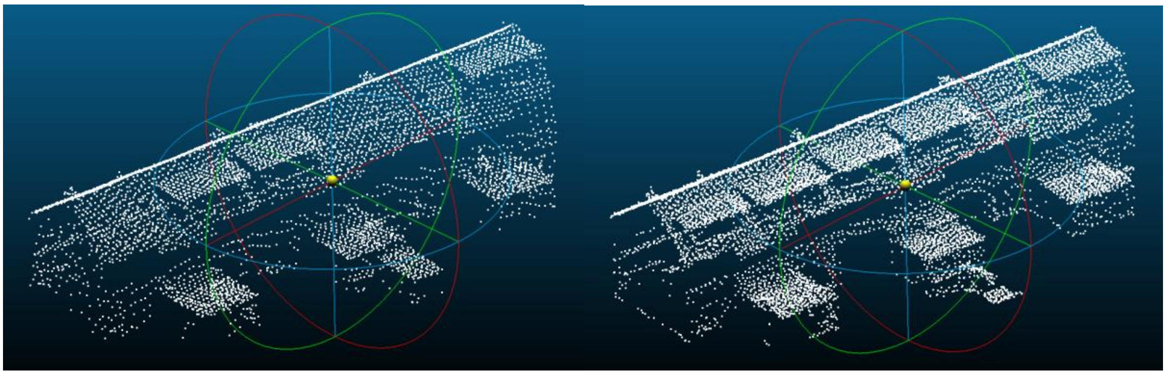


Figure 54 Building 4 Leeuwarden Pointclouds AHN3 and AHN4

When moving to the Sneek region we find that Building 1 has a single overlap in AHN3 where AHN4 has none. The impact of that absence of overlap can be seen in figure 55. The AHN3 representation has a very consistent distribution of points over the entire building. AHN4 however suffers from huge gaps on one side of the building. The building results in occlusion on one side of the building and due to the fact that that single scan was the only one it results in huge gaps. Also this is the only building of the eight where AHN3 has a higher Point Density than AHN4.

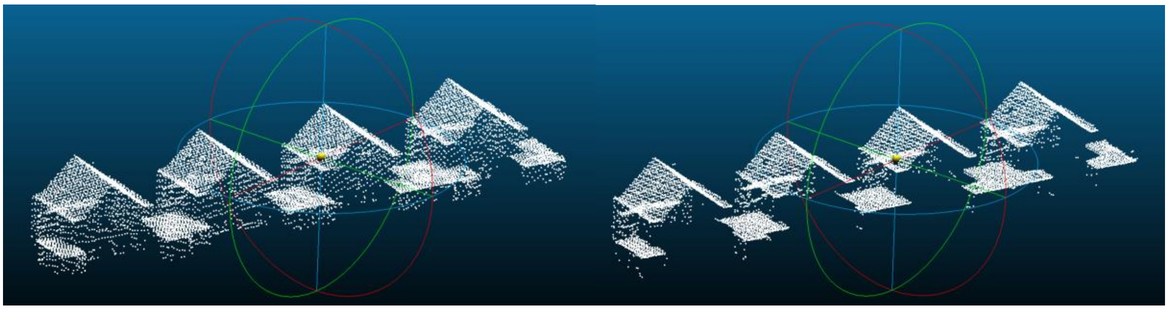


Figure 55 Building 1 Sneek Pointclouds AHN3 and AHN4

Building 2 in the Sneek region has no overlaps for both AHN3 and AHN4. AHN3 does have slightly more points on one of he sides of the building but the difference in completeness is very little. AHN4 did scan almost twice as many points in that single scan compared to AHN3.

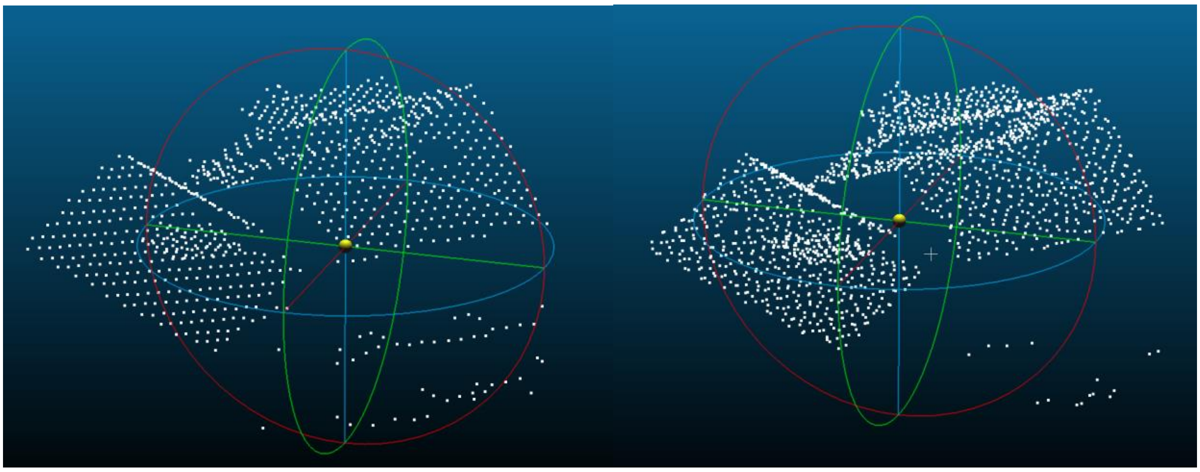


Figure 56 Building 2 Sneek Pointclouds AHN3 and AHN4

For building 3 in the Sneek area there area single overlaps for both AHN3 and AHN4. The Point Density is a lot higher for AHN4 but the overall representation seems very similar due to almost no gaps.

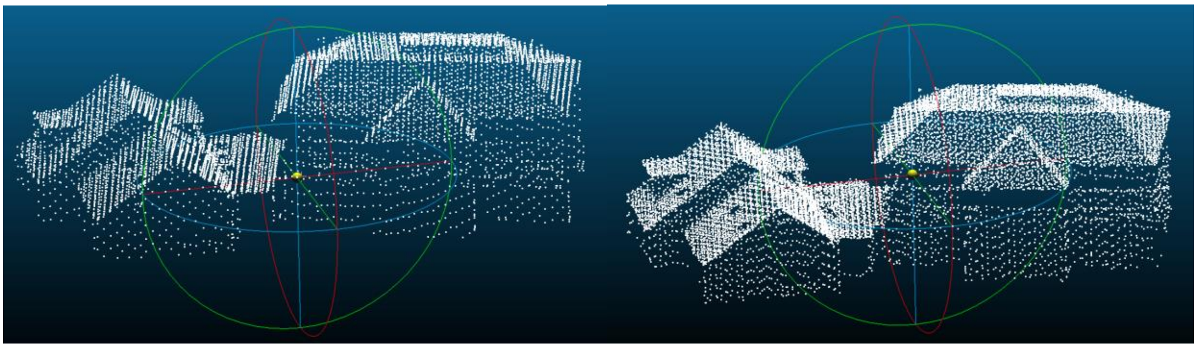


Figure 57 Building 3 Sneek Pointclouds AHN3 and AHN4

Building 4 in the Sneek region is very similar to building 3, both AHN3 and AHN4 have a single overlap and the Point Density of AHN4 is a lot higher. However the overall representation of the building is fairly similar due to the good distribution of points and little gaps in the point clouds.

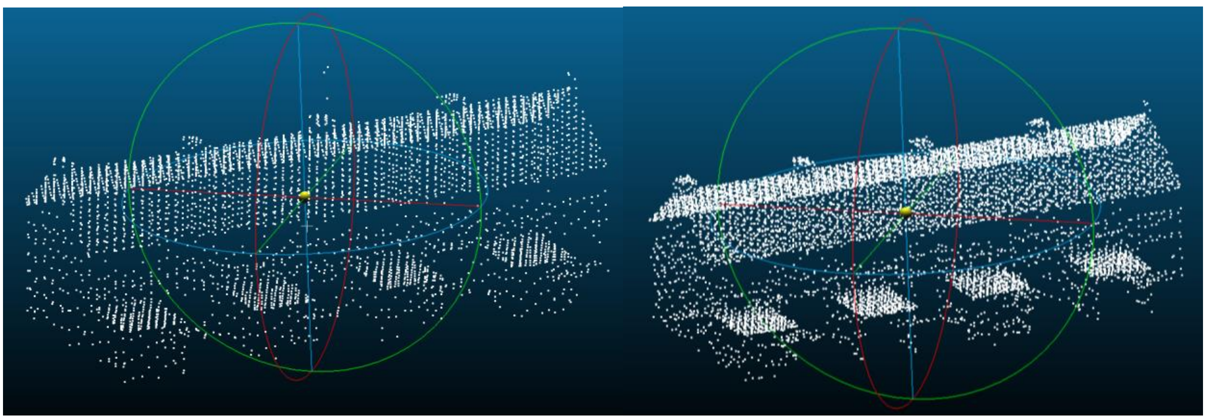


Figure 58 Building 4 Sneek Pointclouds AHN3 and AHN4

When we analyze the outcomes of these eight buildings, it becomes evident that, in the majority of cases, AHN4 exhibits significantly higher Point Density even when the quantity of overlaps is identical. Conversely, AHN3 typically features more overlaps, yielding fewer disparities in terms of Point Density between the two datasets. However, the perspective changes when considering completeness. In the Leeuwarden area, AHN3 encompasses a substantially larger portion of overlapping regions. This often culminates in a comparatively more complete representation or, at the very least, a representation similar to that of AHN4.

Differences Zuid-Holland with Friesland

The distribution of points of AHN3 in Zuid-Holland and the two case study areas in Friesland are shown in figure 59. In these charts can be seen that both the surface area as the distribution of triangles (points) are very similar for these in these three areas.

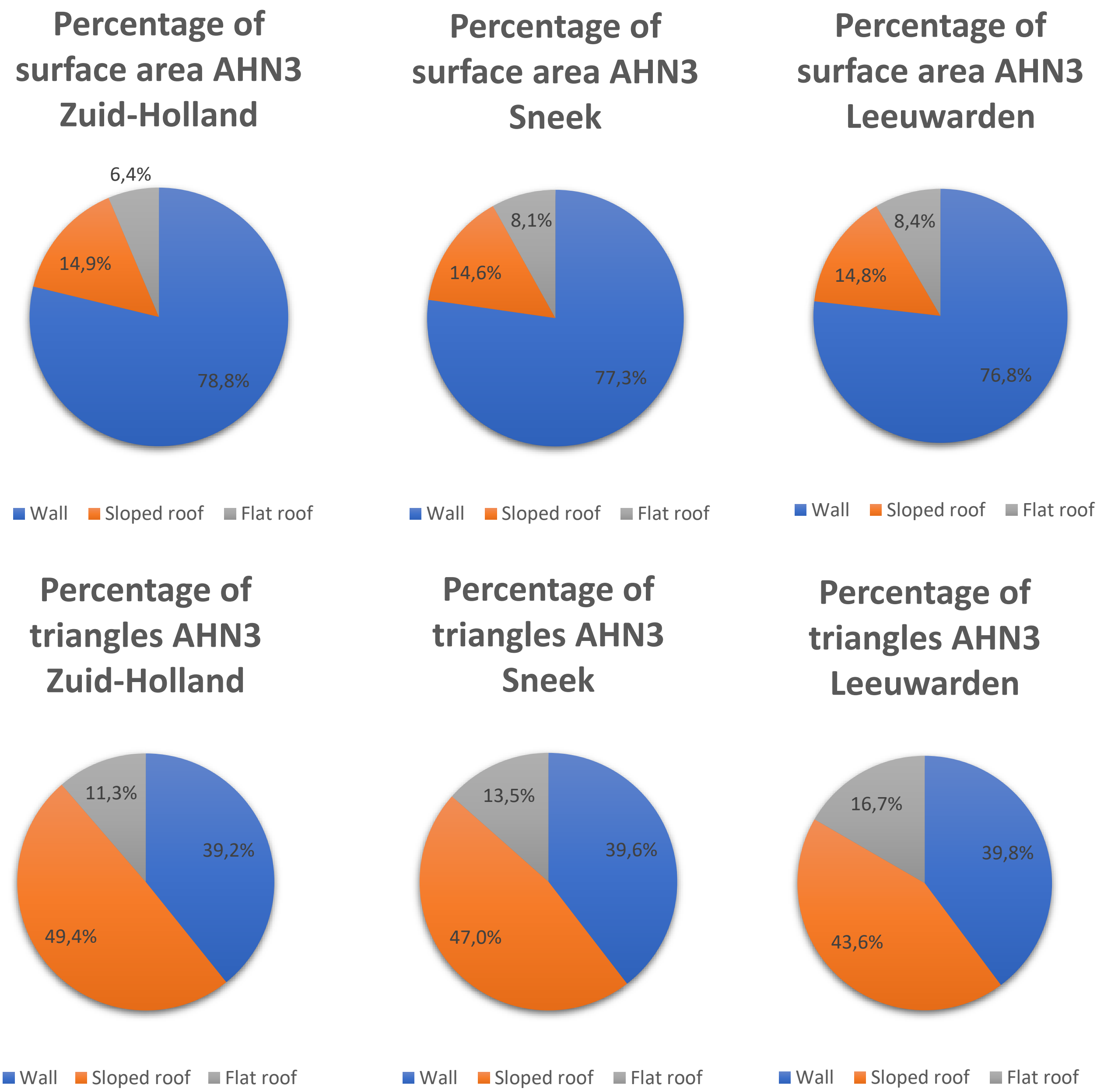


Figure 59 Differences in triangles distribution AHN3

When looking at AHN4 there is not a big difference in triangle distribution between the three regions. The areas in Friesland have much larger percentage of the triangles on a flat roof in both AHN3 and AHN4.

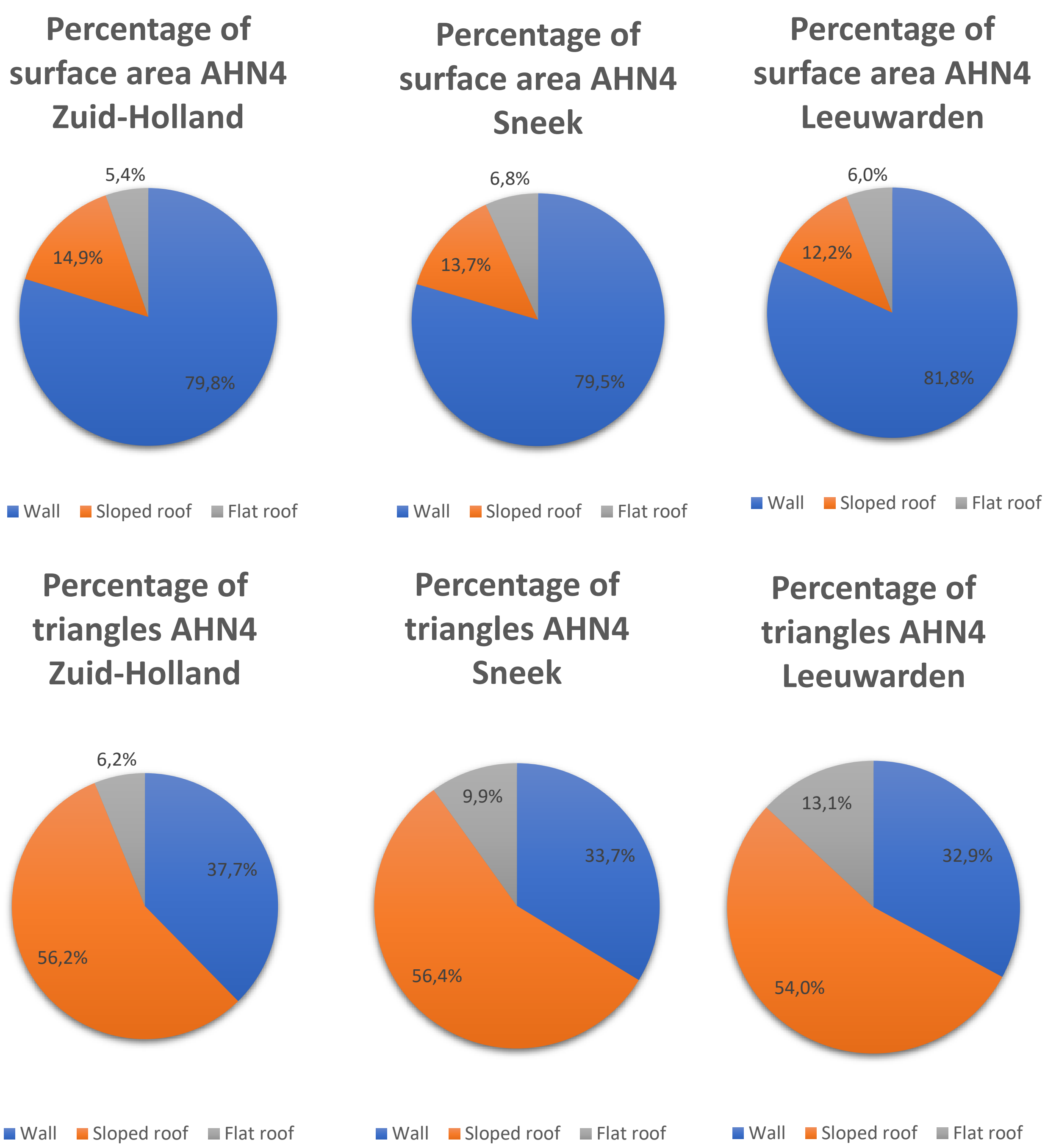


Figure 60 differences in triangles distribution AHN4

Observing the gap percentage in Figure 61, a substantial disparity becomes apparent when comparing Zuid-Holland to the two regions in Friesland. In all three areas, it is evident that AHN4 exhibits a reduced gap area in comparison to AHN3. While the two Friesland areas portray a remarkable similarity to each other in terms of AHN4, a notable enhancement is visible in AHN3. This improvement is likely attributed to the additional flights that were conducted.

	Zuid-Holland	Sneek	Leeuwarden
Percentage Gap AHN3	68,89%	79,70%	76,49%
Percentage No- Gap AHN3	31,11%	20,30%	23,51%
Percentage Gap AHN4	58,86%	71,04%	72,16%
Percentage No- Gap AHN4	41,14%	28,96%	27,84%

Figure 61 Differences in area gaps between Zuid-Holland and the two areas in Friesland.

However, the differences in area that is a gap are fairly large between Zuid-Holland and both areas in Friesland. Zuid-Holland also performs a lot better even though Leeuwarden has extra flights and therefore more overlaps, there were still a lot more gaps than Zuid-Holland. The amount of overlaps in Zuid-Holland is not available so either there were a lot of extra flights flown above Zuid-Holland or they used a different LiDAR scanner for Zuid-Holland compared to Friesland for AHN3.

5. Conclusion; How AHN4 performs in representing 3D buildings compared to AHN3

In this concluding chapter, the outcomes of this study will be synthesized. This research aims to assess the performance between point clouds in the representation of 3D buildings. This research focused on comparing AHN3 and AHN4 in their performance to represent 3D buildings. The main question that needs to be answered is: *"To what degree does the AHN4 point cloud enhance performance in 3D building representation, as compared to the AHN3 point cloud?".* This will be answered according the research sub-questions.

First a literature review was conducted in order to find out what was previously researched around the performance of point clouds. The answering of the main question will be done for two types of applications, the Human Vision System and for Computer Algorithms.

Computer Algorithms are used to interpret and understand the characteristics, shape, and properties of a 3D object. The Human Vision System does this in a subjective matter. In this case the human brain is analysing the point cloud to interpret and understand the characteristics, shape, and properties of a 3D object. The sub-question that is researched in this literature review is: *"What are the key performance aspects of a point cloud that affect the perceived quality for the Human Vision System, and how does this affect compare to the perceived quality for Computer Algorithms?"* Two key performance aspects were found in this research, the Point Density and the completeness of the point cloud.

Point Density refers to the amount of points that there are in a point cloud for a given area. Point Density is important for both the Human Vision System as for Computer Algorithms. For humans there is a correlation between the perceptive quality of an object and the Point Density of the object. For Computers Algorithms the Point Density is mostly important to perceive and identify smaller details in a 3D model such as a chimney as can be seen in figure 13.

The second key performance aspect is the completeness of an object. Where Point Density refers to the amount of points on a object does completeness refer to the distribution of these points on the building. If a certain surface on a object has no points then it can be considered a gap. Completeness is of lesser significance for the Human Vision System, as the human brain adeptly extrapolates and fills gaps. However, Computer Algorithms struggle considerably when encountering gaps within a point cloud. Particularly with complexer shapes these gaps can lead to challenges in accurately identifying the shape of a building and result in errors within a derived 3D object.

Completeness is however a lot more difficult to determine compared to Point Density, which is just counting the number of points in a area. Point Spacing is used in this research to determine the completeness of a cloud. Point Spacing refers to the average distance from a point in a point cloud to all the adjacent points. When this distance is calculated in 3D then it is a good measurement for the amount of points on that piece of surface. When the Point Spacing is a high distance then it indicates a gap in the point cloud.

The first step after the literature review is calculating the Point Density in AHN3 and AHN4 on buildings. The corresponding sub-question is: *"What is the average Point Density of buildings in AHN3 and AHN4, and how do these densities compare between the two datasets?".* In AHN3 there are around 2.4 billion points representing these buildings. AHN4 however has more than twice as much at around 5.5 billion points. Therefore the average

Point Density is more than twice as high in AHN4 compared to AHN3. Out of all 572963 buildings in Zuid-Holland there are 531098 buildings where AHN4 has a higher Point Density compared to AHN3. That makes that 92.6% of buildings have a higher Point Density in AHN4 in comparison to AHN3. Also was found that the point density in Zuid-Holland was much higher than in Friesland for both AHN3 and AHN4. However the spread in point density increased significantly in Zuid-Holland for AHN4 which means that there is much regional differences. So the answer on the *question "What is the average Point Density of buildings in AHN3 and AHN4, and how do these densities compare between the two datasets?"* is that the AHN4 point cloud has much higher Point Density.

After that the completeness of both AHN3 and AHN4 were analyzed. The sub-question was *"What is the Point Spacing on building elements in AHN3 and AHN4, and how do the patterns in Point Spacing compare in terms of completeness and gaps in the point clouds?".* The Point Spacing was calculated by making a mesh with a Delauney triangulation. The area of each triangle in the mesh is the result of three sides which are the distance to adjacent points. If the area is above a certain threshold then it is considered a gap. The outcome is that AHN4 has fewer gaps compared to AHN3, and these gaps also represent less surface area. So AHN4 is an improvement over AHN3 when it comes to completeness. Also is looked at the distribution of triangles on a building. The slope of the triangle indicates whether it represents a wall, sloped roof or flat roof element. When comparing the distribution of triangles on the different building elements with the amount of surface area it seems that in both AHN3 and AHN4 the walls of a building are vastly underrepresented. About 79% of the surface area of buildings are walls but only 39,2% of the triangles represents that area.

The fourth and last sub-question focusses on the distance to flight lines. The sub-question is *"How do the quantitative and qualitative differences between AHN4 and AHN3 relate to the distance from a building to the nearest flight line?".* This is done by calculating the distance from a building to the nearest flight line. This distance is statistically compared to the Point Density on a building and the percentage of gaps on that building. The outcome is that it is evident that there are significantly fewer flights conducted to capture data for AHN4 compared to AHN3. This has however on average not resulted in more gaps or a lower Point Density. When looking at the statistics there were some (very) weak correlations found between the distance to the nearest flight line and the Point Density and amount of gaps. The further the flight line is the (slightly) higher the Point Density and (slightly) more gaps.

After that a case study was conducted. The case study area was in Friesland and consisted of two AHN pages. One near the city of Leeuwarden and one near the city of Sneek. These areas were selected because there was sufficient data available to know the scanning area of a flight and therefore also the amount of overlap in these scanning areas. The impact of these overlaps is researched in this case study. The area around Leeuwarden is specifically chosen because the flight method was different here. In this area extra were flights conducted resulting in even more overlaps in scanning areas. The case study found that in most cases AHN4 has a higher Point Density and is on average more complete in line with the results above. Even in the areas where there were extra flights and therefore more overlap in the scanning area it seems that the Point Density is on par with AHN4 at best, but still lower in most cases.

In terms of completeness the amount of overlaps is more important and in some specific areas where AHN3 has one or more overlaps and AHN4 it outperforms is successor. Especially the percentage of area with a overlap has decreased with about 10%. This means that the overlap width did not increase with the increase of flight height which could indicate why in some cases AHN4 does have much more occlusion compared to AHN3 as indicated by 3DBAG.

The main question of this research is *"To what extent does the AHN4 point cloud improve performance compared to the AHN3 point cloud in representing 3D buildings?"*. The answer to this question is that AHN4 consistently yields an enhancement for the Human Vision System, owing to its increased Point Density. For Computer Algorithms there is on average also an improvement, apart from some small areas where AHN3 is slightly performing better in terms of completeness due to more surface with overlap. For future iterations occlusion should be taken into account when making the requirements for the scanning. This research shows that there is a way to calculate Point Spacing instead of Point Density in order to calculate the amount of gaps and occlusion.

6. Discussion

Reflecting on the conducted research, there inevitably emerge aspects that could have been approached differently. These encompass elements that were constrained by time limitations, as well as areas of potential investigation that hold interest for future studies. This chapter comprehensively catalogs these considerations. It begins with critical observations pertaining to the executed research, followed by insightful propositions for prospective research endeavors.

6.1 Remarks about the research

Looking back at the research then there are some things that were done differently if the research had to be done over. The first thing that comes to mind is analysing the distance to flight lines. In the analysis done there is not taken into account that the flight height could differ or that there are overlaps on the far end. The case study did look at these aspects but in hindsight it would probably be better to look at hot spots of certain key performance aspects plotted into a map and compare these with the flight lines.

At first it was the intention to do this research for the entire Netherlands. This was however not feasible so it was chosen to just research the province of Zuid-Holland. In the end even that took too long to conduct the entire completeness analysis of making meshes and analysing them. Given the differences between the results in Zuid-Holland and the case study area it would have been better to research a selection of areas all around the Netherlands. This would have allowed for analysis of regional differences.

The method of making a mesh to calculate the distribution of points and the amount of gaps turned out to be a successful one. However it is not perfect. In order to let the mesh cover the entire building it is snapped to the ground polygon which has the Z-value of the lowest point on the building. The problem with this method is that it snaps on all the nodes of the ground polygon. So if the polygon only has four nodes this will result in a lower amount of triangles then when the polygon consist of hundreds of nodes. This effect was not taken into account in this research and if this research would be done again there should be thought of a way to improve on this aspect.

For the case study it was necessary to calculate the scanning area for a given flight line. The height that was mentioned in the data was 1402 meters. However it this height is probably not constant precisely 1402 meters so the scanning areas are maybe not everywhere accurate. It is however not expected that the height will differ much so the method used is good enough.

Another aspect that was not taken into account in this research but could have affected the results is the high dynamic regions in the Netherlands. Some areas in the Netherlands are scanned more due to its dynamic nature. This mostly consists of beaches where water boards want to inspects differences in dunes and beaches. These regions could have a different performance compared to the rest of the Netherlands but that is not researched in this thesis.

6.2 Future research

Apart from things that could be done differently in this research there are also things that could be interesting for future research. In the case study is found that there are big differences between Zuid-Holland and the areas in Friesland in terms of amount of gaps. This could be the result of many things including using more extra flights in Zuid-Holland or the use of different laser scanners for AHN3. For AHN4 there no other laser scanner used for Zuid-Holland but still there is a big difference in amount of gaps in these areas. It would be interesting to research what results in this difference between Zuid-Holland and Friesland.

This thesis mostly focusses on AHN3 and AHN4. But at the moment of writing AHN5 is already started being scanned by a company named Miramap. During the AHN congress in May this year they presented that the requirements are the same for AHN4 and AHN5 but that the method of scanning is different. AHN5 will be scanned using a conical scanning pattern instead of a sweeping pattern used in AHN3 and AHN4. It would be interesting to research what the impact of this scanning pattern is to the amount of occlusion in the AHN5 point cloud. This research can be used as a basis for comparing this new AHN5 point cloud with its predecessors.

Another interesting idea for a research is using existing 3D model to predict the amount of occlusion for a given flight. If you know the height at which a plane flies and at what angle it is able to scan, then you can estimate the amount of occlusion for a given flight plan. This could also be an interesting method when used for existing point clouds like AHN4 to compare these estimations to the actual situation.

7 References

3DBag. (2023). Data sources. Accessed on 06-11-2023, from https://docs.3dbag.nl/en/overview/sources/

Actueel Hoogtebestand Nederland. (2023). AHN4. Accessed on 07-06-2023, from https://www.ahn.nl/ahn-4

Actueel Hoogtebestand Nederland. (2023). Kwaliteitsbeschrijving. AHN. Accessed on 07-06-2023, from https://www.ahn.nl/kwaliteitsbeschrijving

Ahokas, E., Kaartinen, H., & Hyyppä, J. (2008). On the quality checking of the airborne laser scanning-based nationwide elevation model in Finland. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 37, 267-270.

Baauw, M. (2021). Maintaining an up to date Digital Twin by direct use of point cloud data (Master's thesis) http://www.gdmc.nl/publications/2021/MScThesisMarcBaauw.pdf

Batty, M. (2018). Digital Twins. Environment and Planning B: Urban Analytics and City Science, 45(5), 817– 820. https://doi.org/10.1177/239980831879641

Bregt, A. K., Grus, L., van Beuningen, T., & van Meijeren, H. (2016). Wat zijn de effecten van een open Actueel Hoogtebestand Nederland (AHN)?. Wageningen University & Research.

van Bochove, D. P. (2019). Combining MLS & ALS Point Cloud Data (Master's thesis). http://www.gdmc.nl/publications/2019/MSc_thesis_Derek_van_Bochove.pdf

Centraal Bureau voor de Statistiek. (2023). Regionale kerncijfers Nederland. Retrieved on 15-06-2023 from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?ts=1686862005168

da Silva Cruz, L. A., Dumic, E., Alexiou, E., Prazeres, J., Duarte, R., Pereira, M., Pinheiro, A., & Ebrahimi, T. (2019). Point cloud quality evaluation: Towards a definition for test conditions. In Proceedings of the IEEE International Conference on Quality of Multimedia Experience (pp. 1-6). IEEE.

Elberink, S. O., & Vosselman, G. (2011). Quality analysis on 3D building models reconstructed from airborne laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing, 66(2), 157-165.

Fernandez-Diaz, J.C.; Carter, W.E.; Shrestha, R.L.; Glennie, C.L. Now You See It… Now You Don't: Understanding Airborne Mapping LiDAR Collection and Data Product Generation for Archaeological Research in Mesoamerica. Remote Sens. 2014, 6, 9951–10001.

Glennie, C. L., Carter, W. E., Shrestha, R. L., & Dietrich, W. E. (2013). Geodetic imaging with airborne LiDAR: The Earth's surface revealed. Reports on Progress in Physics, 76(8), 086801.

Heidemann, H. K. (2014). Lidar base specification version 1.0: US Geological Survey techniques and methods. US Geological Survey.

Jo, H. C., Sohn, H. G., & Lim, Y. M. (2020). A LiDAR Point Cloud Data-Based Method for Evaluating Strain on a Curved Steel Plate Subjected to Lateral Pressure. Sensors, 20(3). https://doi.org/10.3390/s20030721

Kadaster. (2022). Over BAG. Accessed on 19-6-2022, from https://www.kadaster.nl/zakelijk/registraties/basisregistraties/bag/over-bag

Ketzler, B., Naserentin, V., Latino, F., Zangelidis, C., Thuvander, L., & Logg, A. (2020). Digital Twins for cities: A state of the art review. *Built Environment*, *46*(4), 547-573.

Lafarge, F., & Mallet, C. (2012). Creating Large-Scale City Models from 3D-Point Clouds: A Robust Approach with Hybrid Representation. International Journal of Computer Vision, 99(1), 69–85. https://doi.org/10.1007/s11263-012-0517-8

Ledoux, H., Arroyo Ohori, K., and Peters, R. (2020). Computational modelling of terrains.

Leusink, J. (2019, April 16). AHN4 in drie jaar! . . . en daarna? Retrieved March 10, 2023, from https://www.ahn.nl/_flysystem/media/ahn4_in_drie_jaar._en_daarna_-_jeroen_leusink_hwh.pdf

Liu, Q., Su, H., Duanmu, Z., Liu, W., & Wang, Z. (2023). Perceptual Quality Assessment of Colored 3D Point Clouds. IEEE Transactions on Visualization and Computer Graphics, 29(8), 3642-3655. https://doi.org/10.1109/TVCG.2022.3167151

Loenen, B. van, Donker, F.W., Braggaar, R. (2016). De stand in opendataland 2016, Kenniscentrum Open Data, TU Delft.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). Geographic information science and systems. John Wiley & Sons.

Mandlburger, G., Wenzel, K., Spitzer, A., Haala, N., Glira, P., & Pfeifer, N. (2017). Improved topographic models via concurrent airborne LiDAR and dense image matching. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 4.

Nguyen, A. (2013) 3d Point Cloud Segmentation: A Survey. In Proceedings of the IEEE 6th International Conference on Robotics, Automation and Mechatronics (RAM 2013), Manila, Philippines, 12–15 November 2013.

Okyay, U., Telling, J., Glennie, C. L., & Dietrich, W. E. (2019). Airborne lidar change detection: An overview of Earth sciences applications. Earth-Science Reviews, 198, 102929.

van Oosterom, P., Martinez-Rubi, O., Ivanova, M., Horhammer, M., Geringer, D., Ravada, S., ... & Gonçalves, R. (2015). Massive point cloud data management: Design, implementation and execution of a point cloud benchmark. Computers & Graphics, 49, 92-125.

Oude Elberink, S. (2008). Problems in Automated Building Reconstruction based on Dense Airborne Laser Scanning Data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(3A), 93-98.

Oude Elberink, S. J., & Vosselman, G. (2009). 3D information extraction from laser point clouds covering complex road junctions. The Photogrammetric Record, 24(125), 23-36.

Peters, R., Dukai, B., Vitalis, S., Van Liempt, J., & Stoter, J. (2021). Databronnen. 3DBAG. Accessed on 19-6-2023, from https://docs.3dbag.nl/nl/overview/sources/

Peters, R., Dukai, B., Vitalis, S., van Liempt, J., & Stoter, J. (2022). Automated 3D Reconstruction of LoD2 and LoD1 Models for All 10 Million Buildings of the Netherlands. Photogrammetric Engineering & Remote Sensing, 88(3), 165–170. https://doi.org/10.14358/pers.21-00032r2

Peters, R., Gao, W., Dukai, B., & Stoter, J. (2023, mei 9). 3DBAG-AHN4 en andere ontwikkelingen. Presented on the AHN – BM Congres, Amersfoort, Nederland.

Petras, V., Petrasova, A., McCarter, J. B., Mitasova, H., & Meentemeyer, R. K. (2023). Point Density Variations in Airborne Lidar Point Clouds. Sensors, 23(3), 1593. https://www.mdpi.com/1424-8220/23/3/1593

Pradhan, B., & Sameen, M. I. (2020). Laser Scanning Systems in Highway and Safety Assessment. Springer International Publishing. https://doi.org/10.1007/978-3-030-10374-3

Poullis, C., & You, S. (2009). Automatic reconstruction of cities from remote sensor data. 2009 IEEE Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.2009.5206562

Richter, R. (2018). Concepts and Techniques for Processing and Rendering of Massive 3D Point Clouds. (Doctoral dissertation, University of Potsdam). https://doi.org/10.25932/publishup-42330

Rohrbach, F. (2015, October 14). Point Density and Point Spacing. Accessed on 07-05-2023, from https://felix.rohrba.ch/en/2015/point-density-and-point-spacing/

Rupnik, B., Mongus, D., Žalik, B. (2015). Point Density Evaluation of Airborne LiDAR Datasets. Journal of Universal Computer Science, 21(4), 587-603. DOI: 10.3217/jucs-021-04-0587

Schrotter, G., & Hürzeler, C. (2020). The Digital Twin of the City of Zurich for Urban Planning. PFG - Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 88(1), 99–112. https://doi.org/10.1007/s41064-020-00092-2

Smit, B. (2020). Creating Remote Situation Awareness of Indoor First Responder Operations using SLAM (Master's thesis). http://www.gdmc.nl/publications/2020/MScThesisBart-PeterSmit.pdf

Stoter, J., Arroyo Ohori, K., Noardo, F. (2021). Digital Twin: comprehensive solution or hopeful vision? GIM International, bit.ly/ComprehensiveSolutionorHopefulVision

Stoter, J., Arroyo Ohori, K., & Noardo, F. (2021). Digital Twin: Een allesomvattende oplossing of een (te) veel belovend concept? Geo-Info, 18(4). https://geoinformatienederland.nl/tijdschriften/geo-info-4/

Sun, S., & Salvaggio, C. (2013). Aerial 3D building detection and modeling from airborne LiDAR point clouds. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6(3), 1440-1449.

Tse, R., Gold, C., & Kidner, D. (2007). Using the Delaunay Triangulation/ Voronoi Diagram to extract Building Information from Raw LIDAR Data. 4th International Symposium on Voronoi Diagrams in Science and Engineering (ISVD 2007). https://doi.org/10.1109/isvd.2007.40

Tomljenovic, I., Rousell, A., & Strasse, B. (2014). Influence of point cloud density on the results of automated object- based building extraction from ALS data. AGILE Digital Editions. https://repositori.uji.es/xmlui/bitstream/handle/10234/98911/26agile2014_130.pdf?sequence=1&isAllowed=y

TUDelft3D (2023) Sources Documentation 3D BAG. Accessed on 08-01-2023, from https://docs.3dbag.nl/en/overview/sources/

TUDelft3D (2023) Sources Documentation 3D BAG. Accessed on 06-11-2023, from https://docs.3dbag.nl/en/overview/sources/

Vosselman, G., & Dijkman, S. (2001). 3D building model reconstruction from point clouds and ground plans. International archives of photogrammetry remote sensing and spatial information sciences, 34(3/W4), 37-44.

Wang, Q., Tan, Y., & Mei, Z. (2020). Computational methods of acquisition and processing of 3D point cloud data for construction applications. Archives of computational methods in engineering, 27(2), 479-499.

Warchoł, A. (2019). The concept of LIDAR data quality assessment in the context of BIM modeling. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 61-66.

Wu, J., Yao, W., Chi, W., & Zhao, X. (2011). Comprehensive quality evaluation of airborne lidar data. In International Symposium on Lidar and Radar Mapping 2011: Technologies and Applications (Vol. 8286, pp. 30-37). SPIE.

Zhang, J., Huang, W., Zhu, X., & Hwang, J. N. (2014). A subjective quality evaluation for 3D point cloud models. In 2014 International Conference on Audio, Language and Image Processing (pp. 827-831). IEEE.

Zonatlas. (2023). Zo werkt Zonatlas Accessed on 07-06-2023, from https://www.zonatlas.nl/start/over-ons/