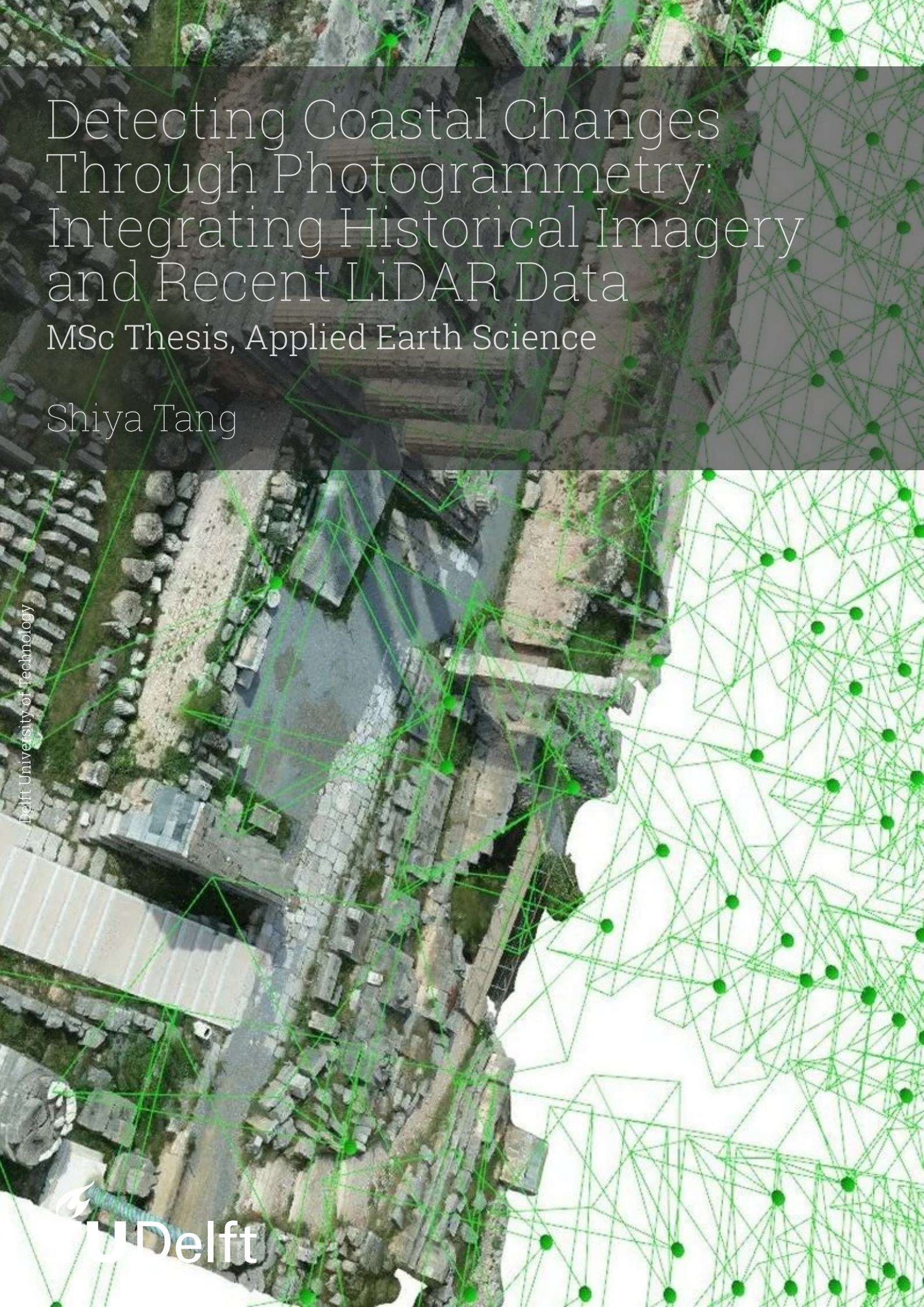
Detecting Coastal Changes Through Photogrammetry: Integrating Historical Imagery and Recent LiDAR Data

MSc Thesis, Applied Earth Science

Shiya Tang



Delft University of Technology

TU Delft

Detecting Coastal Changes Through Photogrammetry: Integrating Historical Imagery and Recent LiDAR Data

MSc Thesis, Applied Earth Science

by

Shiya Tang

Shiya Tang
5914671

Instructor:	R.C. Lindenbergh
Teaching Assistant:	José A. A. Antolínez
Project Duration:	January, 2025 - September, 2025
Faculty:	Faculty of Civil Engineering and Geosciences, Delft

Style: TU Delft Report Style, with modifications by Daan Zwaneveld

TU Delft

Abstract

This study evaluates the feasibility of applying photogrammetry techniques to reconstruct historical coastal topography and assess decadal scale coastal change from historical aerial images, focusing on the Dutch coastal of Westkapelle. The workflow was validated on two benchmark datasets, Benchmark Toronto and Benchmark Westkapelle, to verify registration accuracy under ideal acquisition conditions before being applied to the Westkapelle test dataset for change detection.

The reconstruction pipeline employed Structure from Motion and Multi-View Stereo algorithms, followed by a two stage point cloud registration. Coarse alignment was achieved through Sample Consensus Initial, and fine registration used the Iterative Closest Point algorithm. In the two benchmark datasets, registration achieved sub-metre mean C2C distances: 0.790 m for Toronto, 0.626 m for Westkapelle. In the test dataset, the mean C2C distance improved from 10.287 m before registration to 2.025 m afterwards, with 95% of points within $\sim$6 m. DEM differencing, supported by JARKUS cross-shore transect profiles, revealed systematic elevation gains of up to $+10$ m along foredune ridges, primarily resulting from a combination of documented coastal nourishment and natural process between 1990 and 2020. However, limitations of the historical dataset, including sparse image coverage, strongly oblique viewing geometry, lack of vertical imagery, and poor GCP distribution, introduced geometric distortions and inconsistencies. The resulting orthophoto contained substantial voids and warped features, particularly in urban areas and low texture surfaces, underscoring the challenges of dense stereo matching under suboptimal imaging conditions.

Despite these constraints, this study demonstrates that meaningful reconstructions of past coastal environments are achievable when supported by careful preprocessing, robust registration, and multi source validation. The proposed workflow offers a transferable approach for extracting geomorphic insights from historical imagery in other coastal settings.

Contents

1

Introduction

1.1. Research Background

For a long time, the acquisition of high-resolution digital elevation models (DEMs) has mostly relied on airborne or ground-based LiDAR (LiDAR). LiDAR is a mature technology and can obtain high-quality elevation models without complex post-processing, so it has been widely used in various fields such as coastal zones, urban construction, and ecological environment monitoring [51]. However, LiDAR acquisition is usually costly and the platform deployment flexibility is limited, which restricts its use in small-scale or frequent monitoring scenarios.

With the popularization of Unmanned Aerial Vehicles (UAVs) and high-resolution commercial cameras, image-based photogrammetry is rapidly developed. Compared with the traditional LiDAR system, photogrammetry can acquire multi-view aerial images and use Structure from Motion (SfM) and Multi-View Scanning (MVS) algorithms to generate 3D point cloud data on the ground surface, thereby construct digital elevation models with a high degree of flexibility and economy [15, 11, 14].

The Dutch coastline represents about half of the national boundary line, which makes the monitoring and management of the coast particularly important. In recent decades, sea level rise due to climate change has been an imminent threat to the Netherlands due to its low-lying topography[19]. Over the past few decades, the Netherlands has implemented a series of large-scale hydraulic projects, such as the Zuiderzee Works and the Delta Plan, to reduce the risk of storm surges and long-term overflows. [66, 60]. Coastal change, as a long-term process, relies on managers to understand the shoreline, coastal elevation changes and coastal landscape changes of the coast at different times to synthesize and examine the [38]. Modern LiDAR and UAV photogrammetry can provide near real-time, centimeter-level accuracy data, but they only go back as far as the 1990s. Further back in time, historical imagery is one of the few valuable resources available to understand the landscape at the time [10].

Attempts to integrate historical aerial imagery with modern SfM and LiDAR data allow for the reconstruction of DEMs on interdecadal scales and thus quantitatively analyze coastal evolution over decades.

1.2. Research Problem Statement

With the continuous development of photogrammetric technology, the reconstruction of 3D models using historical aerial images has become one of the valuable tools for studying historical landforms. In particular, in periods when LiDAR data were unavailable, Structure-from-Motion (SfM)-based 3D modeling methods offer a practical and feasible solution for recovering historical DEMs by making use of archival imagery.

In order to accurately embed the generated 3D models into the real-world coordinate system, the traditional photogrammetry workflow usually involves two means of georeferencing. One is direct georeferencing based on the known external parameters of the camera, and the other is indirect georeferencing with the help of the recognizable ground control points (GCPs) within the image[52, 57]. However, for

most historical aerial images, both of the means are difficult to satisfy-camera parameters are usually missing and it is challenging to identify control points that can be precisely geolocated.

Besides the problems mentioned above, there remains two chanllenges to solve. In most cases, there is a temporal gap between historical imagery and modern LiDAR data. During this period, landscapes may have changed significantly due to natural processes or human activities. In addition, the quality of point clouds generated from historical imagery often depends on the original image resolution scanning quality, and viewing angles, resulting in irregular and sparse point distributions. Modern LiDAR data, on the other hand, have centimeter-level precision and uniform point cloud density. This difference in time span and acquisition makes it challenging for point cloud spatial alignment. Furthermore, registration errors may be difficult to be distinguished from geomorphological changes, complicating the validation of the registration workflow and results.

Based on these challenges, this study focuses on a coastal area in the Netherlands, aiming to combine selected photogrammetric methods with modern point cloud registration techniques. In addition to the main experimental dataset, two representative benchmark datasets are introduced to evaluate the sensitivity of registration accuracy across different time spans and geomorphic settings. To ensure the reliability of the DEM comparison and to distinguish real topographic changes from reconstruction artifacts as much as possible, long-term coastal monitoring data such as JARKUS cross-shore profiles were additionally consulted.

1.3. Research Questions

The research question for this study is: **How can photogrammetry applied to historical images enhance our understanding of coastal changes?**
To answer the lead research question, the following sub-questions need to be addressed:

1. **Coastal Changes**
   - What are the common and significant changes in coastal areas over the last few decades?
   - How have natural and anthropogenic factors contributed to these observed coastal changes?
2. **Availability and Suitability of Historical Images**
   - What are the factors that contribute to determine the availability of archival photographs?
   - What are the factors that contribute to determine the suitability of archival photographs?
3. **Point Cloud Generation**
   - What are the photogrammetry workflow involved in generating point clouds from historical images for this study?
   - How is the quality of the point cloud assessed and ensured?
4. **Point Clouds Registration**
   - How to achieve the spatial alignment between the photogrammetric point cloud and LiDAR point cloud?
   - How is the quality of registration assessed?
5. **DEMs Generated from Point Cloud**
   - What is the workfow to generate generating DEMs from point cloud?
   - How is the quality of these DEMs assessed and ensured?
   - What kind of coastal changes can be derived by comparing the DEMs?

2

Background

2.1. Coast Change

The Dutch coast faces multiple dynamic changes driven by both natural environmental forces and human activities in the past few decades. Given the country's low-lying topography and high population density in coastal regions, understanding these changes is essential for sustainable coastal management. In the context of this research, dutch coastal changes are categorized into 3 types: coastline change, elevation change and landscape change.

2.1.1. Coastline Change

The coastline is the dividing line between the oceanand the mainland, and is one of the basic elements of topographic maps and charts[56]. Coastline change refers to the change of a coastline's position over time. This change can result from natural coastal processes like wave action, tide and sediment transport etc. In the context of the Netherlands, due to its densely populated low-lying coastal zones and the dependence on artificial coastal defenses, coastline change is a particular concern. The shoreline dynamics is related to a phenomenon - coastal squeeze, which presents a significant threat to these storm-dominated shores [27]. It happens when sand beaches are trapped between the sea-level rising erosion and the coastal development made by human from the inland side [12].

2.1.2. Elevation Change

Elevation change refers to the variation of land surface height over time, particularly in coastal dunes, beaches, and tidal flats. These changes are driven by both marine and aeolian processes.

One important driver of elevation change is the changes in relative sea level (RSL), both rise and fall [44]. When relative sea level rises, frontal erosion of dunes becomes more frequent, especially during storm events. This erosion can move sand, which is then redeposited inland, leading to the formation of secondary dune features such as blowouts and parabolic dunes. In Netherlands, this mechanism has been linked to the formation of " Young Dunes" [24], which is a second-generation coastal dune system on the west coast of the Netherlands, with a bandwidth of up to about 5 km and a relative height of about 25 m. And they have been formed since the beginning of the Middle Ages (ca. 1000 A.D.) by the inward movement of fresh crustacean sand from the beaches, blown inwards by the prevailing westerly winds, and overlaying the earlier Older Dunes.[21].

2.1.3. Landscape Change

Landscape change involves the conversion of natural habitats to urban or agricultural use. For example, the natural dunes can be either remained because of the increasing vegetation or disappeares because of construction. In the Netherlands, this may include the urbanization of dune landscapes, the expansion of port and industrial infrastructure, or conversely, re-naturalization efforts that allow vegetation to stabilize sandy environments.

2.2. Digital Elevation Model

Digital elevation model (DEMs) is simply the continuous representation of terrain surface contains XYZ coordinates [26]. DEMs are widely used in geomorphology, hydrological modeling, coastal monitoring, land use planning, and other applications that require surface elevation information.

There are generally two types of DEMs, depending on whether surface features are included, as shown in Fig. 2.1:

- **Digital Surface Model (DSM):** A DSM represents the elevation of a terrain surface including all objects present on it, such as buildings, trees, and other above-ground structures.
- **Digital Terrain Model (DTM):** A DTM represents the bare-earth surface, with vegetation, buildings, and other man-made structures removed.

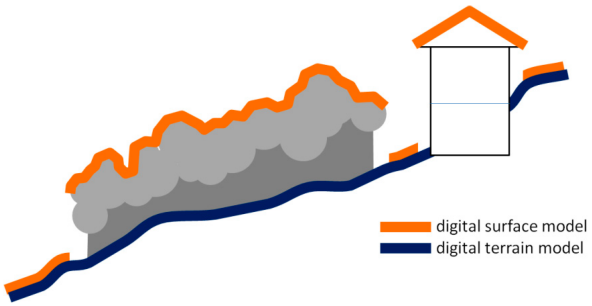


Figure 2.1: Different types of DEM [43]

2.3. Photogrammetry

Photogrammetry is a 3D coordinate measuring technique that uses photos as the medium for measurement and the fundamental principle of it is triangulation. In recent years, photogrammetry has transformed from traditional analogue methods to fully digital and automatic workflows. Modern photogrammetry makes use of the development in computer vision, particularly through algorithms such as Structure from Motion (SfM) and Multi-View Stereo (MVS)[65]. These techniques make the reconstruction of 3D models from overlapping images possible even in the absence of camera calibration.

Photogrammetry is generally categroized into two main types: aerial photogrammetry and close-range photogrammetry. Close-range photogrammetry is widely used for small-scale objects and aerial photogrammetry is more suitable for capturing large-scale terrain and topographic features. Considering that coastal monitoring requires wide-area elevation mapping and shoreline detection, aerial photogrammetry provides a higher degree of spatial coverage and adaptability, making it a more appropriate choice for this research.

2.3.1. Aerial Photogrammetry

The basis of aerial photogrammetry originated from the development of aerial photography. The technique appeared in the late 19th century, and it experienced significant development during World War I due to its application for military purpose. After the war, aerial photography is gradually used in commercial fields. During World War II, this technology was further developed and used again for military purpose.

Based on camera orientation and imaging purpose, aerial photographs can be categorized into 3 main types[61, 62]:

- **Vertical Aerial Photographs:** They are taken with the camera axis aligned as closely as possible to the vertical (nadir) direction. These are the most common type used for topographic mapping, digital elevation model (DEM) generation, and orthophoto production, due to their geometric suitability for accurate measurements.
- **Oblique Aerial Photographs:** They are taken with the camera tilted from the vertical axis. Depending on the tilt angle, these are further divided into high oblique and low oblique images.

Oblique images provide a more intuitive view of the landscape, making them suitable for landscape documentation, archaeology, and real estate.

- **Orthophotos:** They are geometrically corrected aerial images in which distortions caused by terrain relief, lens aberrations, and camera tilt are removed. Unlike raw aerial photographs, orthophotos maintain uniform scale and can be used as accurate base maps in Geographic Information Systems (GIS). See Section 2.8 for further explain, technical details and procedures.

The images used in aerial photogrammetry involves capturing photographs of the Earth's surface from airborne platforms such as aircraft, unmanned aerial vehicles (UAVs), satellites, or rockets. These platforms are equipped with optical sensors that take overlapping images from different locations. When captured with sufficient overlap, these images can be used by photogrammetry.

2.3.2. Photogrammetry Fundamentals

Photogrammetry aims to obtain accurate geometric information about the physical environment from 2D images. The core mathematical principles are grounded in projective geometry, collinearity constraints, and optimization techniques[58, 17, 35].

Pinhole Camera Model

The pinhole camera model is the most commonly used model in photogrammetry because of ability to describe the geometric relationship between 3D world points and their 2D image projections. The model assumes an ideal camera with an infinitesimally small aperture (the "pinhole") through which light rays pass. All light rays from a scene point converge at the camera center and intersect the image plane, forming a projection of the object.

Based on this physical principle, the perspective projection model provides a mathematical description of the projection process. A 3D point $\mathbf{X} = (X, Y, Z)^T$ is projected onto the image plane as a 2D point $\mathbf{x} = (x, y)^T$, according to the following equations:

$$x = f \cdot \frac{X}{Z}, \quad y = f \cdot \frac{Y}{Z} \tag{2.1}$$

where f is the focal length, and the projection assumes that the camera is located at the origin and oriented along the Z-axis. In practice, more complete formulations include the camera intrinsics (e.g., focal length, principal point, and pixel scale) and extrinsics (rotation and translation), which together define the full camera projection model.

The full projection model can be written in matrix form as[17]:

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{R} \mid \mathbf{t}] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.2}$$

where:

- $\mathbf{K}$ is the intrinsic camera matrix,
- $\mathbf{R}$ is the rotation matrix (extrinsics),
- $\mathbf{t}$ is the translation vector (extrinsics),
- s is a scale factor.

Camera Calibration

If we want to restore the real-world position from a 2D image formed by projection, we need to understand the transformation between the two. One of the main purposes of camera calibration is to determine this transformation, which means solving for the intrinsic and extrinsic parameter matrices. In addition, another goal of camera calibration is to estimate distortion coefficients to correct the distortion caused by the camera's perspective projection.

Intrinsic parameters describe the internal characteristics of the camera, such as focal length, principal point, and lens distortion. Extrinsic parameters define the camera's position and orientation in the real world, and are needed to transform coordinates between the camera and object space. In traditional photogrammetry, calibration is performed using known physical targets (such as a checkerboard pattern) to recover parameters like focal length, principal point, and lens distortion.

However, the photogrammetry in this study is based on historical aerial imagery, which means that most of the images were taken before the introduction of GPS and inertial navigation systems. Therefore, the external orientation parameters are either unknown or highly uncertain, and there is no prior internal calibration data. Due to this limitation, this study did not conduct camera calibration in the traditional sense, but instead estimated the intrinsic parameters automatically through SfM, and manually added ground control points (GCPs) to the images to provide real-world coordinates and assist in estimating the extrinsic parameters.

In short, SfM-based self-calibration combined with GCP-assisted alignment provides a practical alternative to traditional calibration, allowing relatively accurate 3D reconstruction and orthophoto generation from historical imagery.

Epipolar Geometry

Epipolar geometry describes the geometric relationship between two images taken from different viewpoints of the same scene. As shown in Fig. 2.2, when a scene is observed from two distinct camera positions C_1 and C_2, the geometry of the cameras and a 3D point M together define the epipolar geometry.

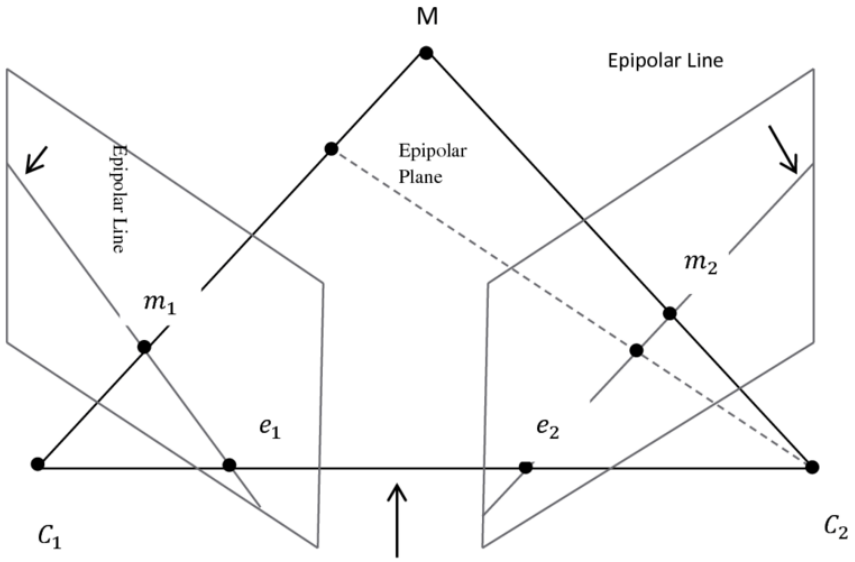


Figure 2.2: Epipolar geometry [7]

In the figure, m_1 and m_2 are the projections of the 3D point M onto the image planes of cameras C_1 and C_2, respectively. The lines joining the two camera centers and the point M defines a unique epipolar plane. This plane intersects each image plane along lines known as epipolar lines, and the projections of one camera center onto the other camera's image plane are called epipoles, denoted e_1 and e_2.

To describe the mathematical relationship between corresponding image points under the epipolar constraint, a matrix can be introduced to express this constraint concisely as an equation. This matrix can either be the essential matrix, defined in the camera coordinate system, or the fundamental matrix, defined in the image (pixel) coordinate system.

In the camera coordinate system, let $\mathbf{p}_l$ and $\mathbf{p}_r$ denote the normalized image coordinates, which are the 2-D coordinates you obtain after removing all camera-intrinsic parameters so that every image

point simply represents the direction of the incoming light ray in the camera reference frame, of a 3D point projected into the left and right cameras, respectively. These points satisfy the following epipolar constraint:

$$\mathbf{p}_l^\top \mathbf{E} \mathbf{p}_r = 0 \tag{2.3}$$

Where, the essential matrix $\mathbf{E}$ describes the relationship between projection points in two camera views, expressed in their respective camera coordinate frames. The constraint imposed by the essential matrix makes it possible to recover the relative motion of the cameras from two views. If the transformation between the two cameras is represented by a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{T}$, the essential matrix can be formulated as:

$$\mathbf{E} = [\mathbf{T}]_\times \mathbf{R} \tag{2.4}$$

where $[\mathbf{T}]_\times$ is the skew-symmetric matrix, which means is any square matrix whose transpose equals its own negative, representing the cross product with $\mathbf{T}$.

In contrast, the fundamental matrix $\mathbf{F}$ relates image points in the pixel coordinate system. Let $\mathbf{c}_l$ and $\mathbf{c}_r$ be the pixel coordinates of the same 3D point projected into the left and right image planes. Then the epipolar constraint becomes:

$$\mathbf{c}_l^\top \mathbf{F} \mathbf{c}_r = 0 \tag{2.5}$$

Typically, the intrinsic matrices of the cameras are required to solve for the essential matrix. However, in practical applications the intrinsics may be unknown, in which case one can resort to the fundamental matrix. The fundamental matrix accounts for both the relative motion between the cameras and the intrinsic parameters of each camera. It is related to the essential matrix through the intrinsic matrices $\mathbf{K}_l$ and $\mathbf{K}_r$ of the left and right cameras:

$$\mathbf{F} = \mathbf{K}_l^{-\top} \, \mathbf{E} \, \mathbf{K}_r^{-1} \tag{2.6}$$

2.4. Structure from Motion (SfM)

Structure from Motion (SfM) is a field within computer vision that focuses on recovering the 3D structure of a scene from a sequence of images using algorithms. Structure from Motion (SfM) builds upon the principles of photogrammetry introduced earlier, such as the perspective camera model and the collinearity constraint. While those concepts focus on the geometric relationship between 3D points and their image projections, SfM extends these ideas to recover both camera poses and 3D structure from unordered or sequential images.

2.4.1. Feature Detection and Matching

A feature is a piece of information that can be consistently recognized in images and scenes and associated with solving a specific type of computational task. For an image, features can be represented as points, lines, edges, or entire objects. In photogrammetry and computer vision fields, feature detection usually refers to the identification of certain interest points in an image that can be well interpreted and express the texture of an object. These interest points often correspond to areas where there is a sharp change in object boundaries or at the intersections of multiple edge segments. Once the detection of interest points is completed, feature descriptors need to be computed for these points to achieve quantifiable comparison and matching. A feature descriptor encodes distinguishing information about a point, allowing for robust identification across different views. For the same point of interest, the feature descriptors remains invariant as the image changes, so that the same point can be matched across multiple images. Descriptors are typically categorized into two types: local descriptors, which capture information within a small neighborhood around the point, and global descriptors, which describe the entire image. However, in general, global descriptors are less robust and reliable compared to local descriptors in most applications.

Among various feature detection algorithms, Scale-Invariant Feature Transform (SIFT) is one of the most representative and widely used methods. SIFT performs multi-scale feature detection by constructing a Gaussian pyramid and computes histograms with with gradient distributions for each key point in order to obtain descriptors, which are designed to be invariant to changes in scale and rotation, making SIFT particularly effective in matching features across diverse images[29].

After feature detection is completed, the next critical step is feature matching. Feature matching refers to establishing correspondences between images of the same scene or object. A common approach involves detecting a set of interest points from the image data and computing a corresponding descriptor for each point. Once feature points and their descriptors are extracted from the images, initial correspondences can be established between them.

The performance of interest-point-based matching methods largely depends on the properties of the detected points and the type of descriptor used. Therefore, it is essential to select appropriate feature detectors and descriptors according to the specific content of the images in practical applications.

2.4.2. Camera Pose Estimation

Camera pose estimation aims to determine a camera's position in a global coordinate frame by means of computer-vision algorithms. Traditional estimation method relies on feature detection and matching: by using the geometric relationships among matched features across images, the camera's extrinsic parameters can be inferred. With the development of machine learning, neural-network models have been proposed to estimate the pose directly from images[34, 53], but these methods usually require large, accurately labelled training sets[36]. Thus, this research adopts the traditional feature-based method.

Mathematically, the problem can be written as recovering the rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$ of the camera from a set of correspondences between 3D points $\mathbf{X}_j$ and their 2D projections $\mathbf{x}_j$:

$$\mathbf{x}_j \;\sim\; \mathbf{K}\,[\,\mathbf{R}\mid\mathbf{t}\,]\,\mathbf{X}_j, \qquad j=1,\ldots,n, \tag{2.7}$$

where $\mathbf{K}$ is the intrinsic matrix and $\sim$ denotes equality up to scale in homogeneous coordinates. The solution of this equation is based on the epipolar geometry introduced earlier.

2.4.3. Bundle Adjustment

Bundle adjustmentis the problem of refining a visual reconstruction to produce jointly optimal 3D structure and viewing parameter (camera pose and/or calibration) estimates[58]. The principle of this method is to consider all image point observations, camera external orientation elements (pose) and 3D point coordinates as a whole, and optimize them simultaneously through least squares iterations to achieve a global minimum of the reprojection error.

In multi-view reconstruction, the same 3D point can be observed from different viewpoints. Define X_j as the spatial coordinates of the j-th 3D point, P_i denotes the projection matrix of the i-th camera, and x_j^i denotes the observed coordinates of the j-th 3D point X_j in the i-th image. Under ideal conditions, the projection point $P_i X_j$ would coincide exactly with x_j^i. In practice, however, due to various errors there is normally some deviation. Bundle adjustment addresses this by jointly optimising all camera parameters and 3D point positions to minimise the sum of squared reprojection errors:

$$\min_{P^{,i}{}_{,,}X_j} ;\sum_{i,j} d\big(P^{,i}X_j,;x_j^i\big)^2 \tag{2.8}$$

where, $d\big(P^{\,i}X_j,\;x_j^i\big)$ represents the distance between the projection point $P^{\,i}X_j$ and the observed point x_j^i. And the optimisation variables include the camera projection matrices $P^{\,i}$ and the positions of the 3D points X_j.

2.5. Multi-view Stereo (MVS)

Structure-from-Motion (SfM) can estimate the relative poses and the intrinsic and extrinsic parameters of the cameras from multiple views, thereby producing a sparse 3-D point cloud. Multi-view stereo (MVS) can then use the camera poses and intrinsics obtained from SfM to estimate the depth of every pixel in each image, thus reconstructing a dense 3-D point cloud.

2.5.1. Two-view stereo

MVS is based on the same principles as two-view stereo, we first introduce the basic concepts of two-view stereo. Two-view stereo is founded on the stereo matching algorithm, which can estimate the

depth of every pixel in the first image by using the second image. The process involves several steps:

Rectification

The goal of rectification is to transform a stereo image pair so that corresponding points lie on the same horizontal scanline. In the special case where the two cameras have parallel image planes, equal focal lengths, and their centres are at the same height relative to the baseline, the images are already rectified. For more complicated cases, these conditions do not hold, and the raw images are not geometrically aligned. Thus, image rectification must be performed with the matrix computation method proposed by Loop and Zhang [30].

Correspondence Search

After rectification and after the epipolar lines have been obtained, all pixels on the epipolar line in the second image are examined to find the pixel that best matches the pixel in the first image, thus establishing correspondences. A variety of similarity metrics can be used, such as the sum of squared differences (SSD) and normalized cross-correlation (NCC).

Consistency Constraints

After correspondences are obtained, several lightweight checks are applied to delete mismatches:

- Each pixel in one image should correspond to a unique pixel in the other image.
- The left-to-right ordering of points along an epipolar line should be preserved, except in cases of true occlusions.
- Disparity values must lie within a plausible range based on scene geometry and camera setup.
- A left-right cross-check is enforced to detect outliers. Specifically, for a pixel at position x in the left image with disparity $d_L(x)$, and the corresponding pixel at position $x' = x - d_L(x)$ in the right image with disparity $d_R(x')$, the following consistency condition must hold:

$$\left|d_L(x) + d_R\big(x - d_L(x)\big)\right| < \tau, \tag{2.9}$$

  where τ is a predefined threshold that tolerates small errors due to noise or calibration inaccuracies.

Pixels violating any of these conditions are re-examined or marked as invalid to improve overall matching reliability.

2.5.2. Multi-view Stereo

The difference in multi-view stereo is that a global view selection must first be performed: based on the 3-D point cloud extracted by SfM, the number of SIFT matches, camera poses, and other criteria, a subset of views adjacent to the reference view is chosen from all available views for stereo matching. After that, local view selection is carried out. When processing a specific pixel, the algorithm once again examines the neighbouring views selected in the previous step and, according to the depths and surface normals already estimated for nearby pixels, as well as the disparities between the reference view and each neighbouring view, selects a small subset of truly useful views for computing the depth of that pixel. The pixel's depth is finally estimated from these few views in order of matching confidence.

2.6. Feature Descriptor

As mentioned earlier in feature detection, feature descriptors can describe points of interest and encode that description in the form of multidimensional feature vectors in a vector space. Feature descriptors can record a range of information about different points of interest and allow humans to effectively distinguish between them. Features are important data for accomplishing the computations required for a particular application. Features in an image can be specific elements such as points, edges or objects.

2.6.1. Point Feature Histogram

The Point Feature Histogram(PFH) is a descriptor that captures the local geometric structure of a 3-D point cloud. For a query point p_q it analyses the pair-wise geometric relationships, angles, distances

and surface normals between p_q and all points inside its k neighbourhood, and encodes the result in a 33-bin multidimensional histogram[49]. By averaging the curvature around the point in this way, PFH provides a rich, high dimensional representation of the neighbourhood's geometry that is robust to variations in sampling density and moderate noise.

Figure 2.3 plots the PFH support region. The red query point p_q lies at the centre of a sphere of radius r and is fully connected to every neighbouring point whose Euclidean distance from p_q is less than r. The final descriptor is the joint histogram over all ordered point pairs within this support, giving the method a computational complexity of $O(k^2)$.

For any interest point p_q, the first thing to do is to gather all points that lie within a radius r, which are its neighbours. Then it is needed to build a small local coordinate frame: its first axis is the surface normal at p_q, and the other two axes are defined by the line that joins p_q to each neighbour. For every neighbour we then measure three angles($\alpha, \phi,$ and θ) plus the distance between the two points. Each angle range is split into eleven equal parts (bins). Counting how often α, ϕ and θ fall into their respective bins produces 33 numbers, which together form a histogram. Finally, normalise this histogram so that all its values add up to 1. The resulting 33 element vector is called the Point Feature Histogram (PFH) of p_q, which provides a compact but expressive summary of the local surface shape.

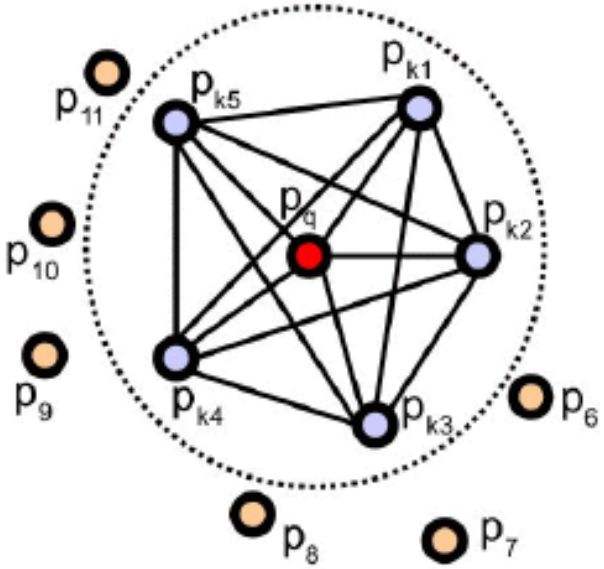


Figure 2.3: Point Feature Histogram. The red point is the query pointp_q; blue points lie within the spherical neighbourhood of radius r. All ordered point pairs inside this region are considered when building the histogram[42].

2.6.2. Fast Point Feature Histogram

Based on the original Point Feature Histogram (PFH), the Fast Point Feature Histogram (FPFH) simplifies the method and reduces the computational complexity to $O(k)$, saving significant computation while retaining most of PFH's descriptive power.

The first step is to compute the Simplified PFH (SPFH). For each query point p_q, only the angular triplet (α, ϕ, θ) between p_q and each of its neighbors is recorded. Unlike the full PFH, SPFH deliberately ignores the triplets between neighbor pairs and focuses solely on the relations from p_q to its neighbors.

Once all points' SPFHs are obtained, the final FPFH descriptor for each point is generated by revisiting its k nearest neighbors and forming a weighted combination of their SPFHs with that of p_q, as is shown in equation below. The weights are typically based on the Euclidean distance, so that closer neighbors have a greater influence on the result.

$$\mathrm{FPFH}(\boldsymbol{p}_q) = \mathrm{SPFH}(\boldsymbol{p}_q) + \frac{1}{k}\sum_{i=1}^{k}\frac{1}{\omega_i}\cdot\mathrm{SPFH}(\boldsymbol{p}_i) \tag{2.10}$$

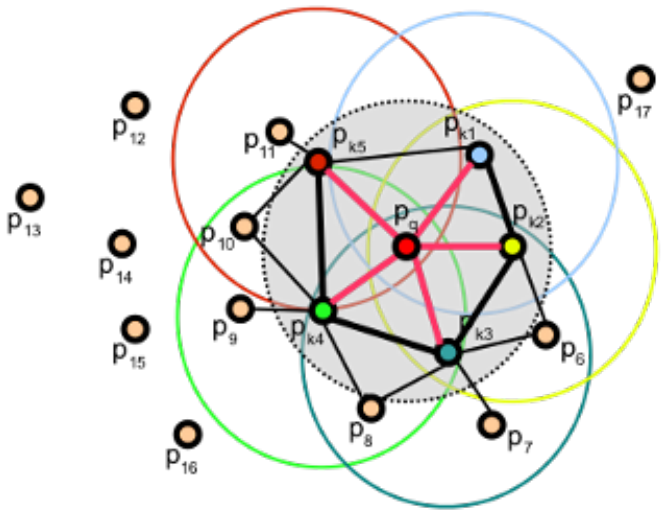


Figure 2.4: Fast Point Feature Histogram[41]

2.7. Point Cloud Registration

Point cloud registration is a key problem in computer vision applications, aiming to align two or more point clouds captured from different viewpoints or sensors into a common coordinate system. It involves finding a rigid transformation that maps one point cloud onto another such that their geometries align accurately [22].

A wide range of point cloud registration methods have been developed to date. Generally, these methods can be categorized into two stages: coarse registration and fine registration. When the initial pose difference between point clouds is large, coarse registration is typically used to provide an initial transformation estimate, which then serves as the starting point for the fine registration stage. Here in this research, the coarse registration algorithm we use is Sample Consensus Initial Alignment (SAC-IA) and the fine registration method we use is Iterative Closest Point (ICP) algorithm.

Early approaches mainly relied on classical geometric principles, estimating transformations by matching local features or minimizing geometric distances. With the rapid development of machine learning, researchers have increasingly explored deep learning-based registration methods, which have demonstrated excellent performance in recent years [4, 31]. In this study, we adopt traditional geometry-based rigid registration methods, which do not require large-scale annotated training datasets and are well-suited for aligning multi-source point clouds, such as those derived from LiDAR and photogrammetry.

2.7.1. Sample Consensus Initial Alignment

Sample Consensus Initial Alignment (SAC-IA) is a feature-based coarse registration algorithm commonly used in point cloud alignment tasks, particularly effective in scenarios where there are large initial rotations or translations between datasets. This method combines RANSAC (Random Sample Consensus) with an initial alignment strategy to estimate a rough rigid transformation by selecting a subset of corresponding feature points[48].

At its core, SAC-IA builds on the principles of RANSAC, an iterative outlier rejection algorithm widely used in machine learning and computer vision. RANSAC works by repeatedly selecting random samples from a dataset to fit candidate models, and evaluating how many data points are consistent with each model. The model with the highest number of inliers is then selected as the best estimate[48].

SAC-IA applies this idea to point cloud feature matching and alignment. While RANSAC is fundamentally a sampling-consistency model estimation method, SAC-IA can be interpreted as sampling-consistent initial alignment. Specifically, it first computes feature descriptors—typically Fast Point Feature Histograms (FPFH)—for each keypoint in the point clouds. Then, based on feature similarity, it constructs an initial set of candidate correspondences. The RANSAC framework is then employed to

iteratively sample and evaluate these correspondences, eventually identifying the most reliable subset to estimate the rigid transformation.

SAC-IA has been successfully applied in various domains, including autonomous robotics, object recognition, and 3D modeling [50, 28, 40]. Its main advantage lies in its robustness to large pose differences and its ability to provide reasonable initial alignment even under challenging conditions. However, its performance heavily relies on the quality of the feature descriptors. his makes SAC-IA a suitable first step before fine registration methods such as ICP.

2.7.2. Iterative Closest Point (ICP) Algorithm

The Iterative Closest Point (ICP) algorithm is a classic and widely used method for point cloud registration in computer vision. It is applicable to both 2D and 3D point clouds and has demonstrated excellent performance, particularly in 3D applications. ICP is commonly adopted as a fine registration technique. The algorithm was first formally introduced by Besl and McKay [9], and its core idea is conceptually simple: it iteratively refines the alignment between two point clouds by minimizing the spatial discrepancy—typically the sum of squared errors—between corresponding points, thereby estimating the optimal rigid transformation.

It is important to note that ICP is highly sensitive to the initial alignment between the point clouds. A poor initial pose may lead to inaccurate results or even prevent the algorithm from converging. The implementation of ICP begins with finding, for each point in the source point cloud, its nearest neighbor in the target point cloud under the current transformation. Since computing nearest neighbors for large-scale point clouds containing millions of points is computationally expensive, ICP typically leverages data structures such as k-d trees. These trees recursively partition the 3D space and enable efficient nearest-neighbor search, greatly reducing computational overhead.

Once the correspondences are established, the algorithm computes the optimal rigid transformation by minimizing the Euclidean distances between corresponding pairs. This transformation is then applied to the source point cloud, and the process is repeated: new correspondences are found, and a new transformation is estimated. With each iteration, the source point cloud gradually aligns more closely with the target, until the change in transformation parameters falls below a predefined threshold. This iterative process is illustrated in Fig. 2.5.

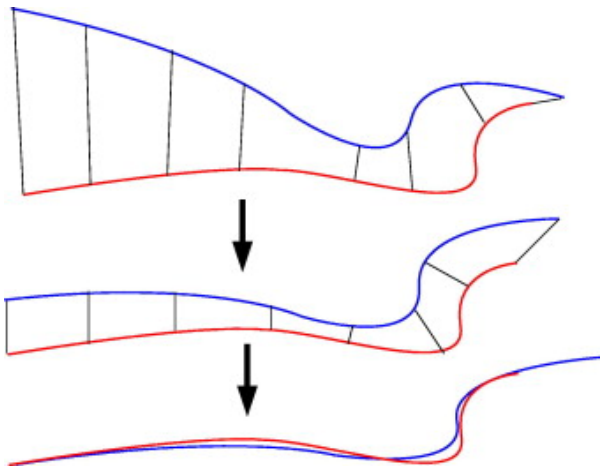


Figure 2.5: ICP algorithm's iterative alignment process[54]

The advantages of ICP lie in its simplicity and efficiency, making it capable of achieving high-accuracy alignment. However, a notable limitation is its sensitivity to the initial pose. As a result, a reliable initial alignment is often required to ensure convergence to a correct solution. This is precisely why, in the registration workflow of this study, the SAC-IA algorithm is introduced prior to ICP to provide an effective coarse alignment.

2.8. JARKUS

In the Netherlands, long-term measurements of the coastline are collected by the JARKUS program (JAarlijkse KUStmetingen), which has been active since 1965. The program is managed by Rijkswaterstaat, a part of the Dutch Ministry of Infrastructure and Water Management. Every year, JARKUS collects cross-shore elevation profiles at regular intervals of approximately 250 meters along the entire Dutch coast[46]. These profiles record the shape of the beach, dunes, and nearshore seabed, and are aligned to the Normaal Amsterdams Peil (NAP), which is the national vertical datum.

Each profile consists of a series of elevation points, ordered by chainage, which represents the distance from a fixed base point on the coast. Because the profiles are collected at the same locations every year, they allow for direct comparison over time. This makes the JARKUS dataset highly valuable for studying coastal changes, such as shoreline retreat, dune erosion, sediment transport, or beach nourishment effects.

Researchers use JARKUS data in a wide range of coastal studies. [63] used JARKUS profiles and aerial images to investigate shoreline changes across different time scales, from single seasons to multiple decades. In many cases, JARKUS serves as a reference dataset or validation source for remote sensing products such as LiDAR or photogrammetry-based DEMs. Its high temporal resolution, wide coverage, and consistent measurement standard make it one of the most reliable sources for detecting true coastal elevation change in the Netherlands.

2.9. Orthophoto

An orthophoto is a type of aerial image that has been corrected so that all objects appear in their true positions on a flat surface, as if viewed directly from above [13]. Original aerial photos often contain distortions caused by things like sensor tilt, movement of the aircraft, or changes in terrain height. Orthorectification removes these distortions, so that every point in the image looks as if it was taken from directly overhead (nadir). Because of this, orthophotos can be directly used to determine geographic locations, measure distances, compute areas, and extract other spatial information about the surveyed region [16]. In this study, orthophotos are used to observe and analyze the spatiotemporal evolution of coastal landscapes. High-resolution orthophotos provide a clear depiction of shoreline position, beach morphology, intertidal zone distribution, and vegetation cover, enabling a direct visual assessment of coastal change over time.

The standard orthophoto generation workflow generally consists of three main steps: image orientation, reprojection using a digital terrain or surface model, and image mosaicking [8]. Image orientation involves estimating the exterior orientation parameters (camera position and attitude) and the interior orientation parameters of each image using Structure-from-Motion (SfM) techniques, thereby determining the spatial geometry of the image set. Next, the original images are orthorectified using a Digital Terrain Model (DTM) or Digital Surface Model (DSM) to remove geometric distortions caused by terrain relief and camera tilt, ensuring that every pixel corresponds precisely to a geographic coordinate. Finally, the orthorectified images are mosaicked together, with radiometric balancing and seamline optimization applied in the overlapping areas, to produce a seamless orthophoto with a uniform scale.

The quality and accuracy of an orthophoto are influenced by multiple factors. First, the resolution and overlap of the original imagery directly affect the spatial detail and registration precision of the final orthophoto. Second, the accuracy of camera interior and exterior orientation parameter estimation impacts the positional accuracy of the orthorectified image. At last, the precision of the DTM or DSM used is critical for eliminating relief displacement, particularly in areas with significant terrain variation [20].

3

Data Description

3.1. Benchmark Dataset of Toronto Downtown

The first benchmark dataset covers the downtown area of Toronto, Canada. This dataset originates from the ISPRS Test Project on Urban Classification, 3D Building Reconstruction and Semantic Labeling, and includes two typical high-density urban blocks within Toronto's financial district, each measuring approximately $530\,\mathrm{m} \times 600\,\mathrm{m}$[47]. The data used in this study include both aerial imagery and airborne LiDAR data, which were acquired between September 2011 and the end of 2012. The short temporal gap between the two datasets, combined with the fact that the streetscape of downtown Toronto remained largely unchanged during this period, allows for a reliable validation of the registration workflow by minimizing the impact of temporal differences.

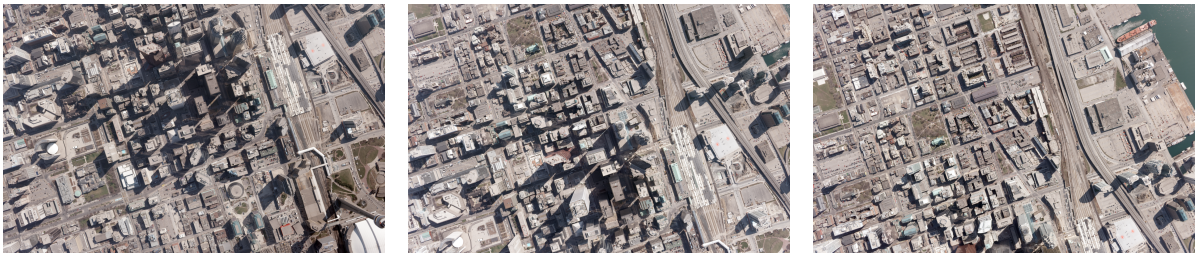
3.1.1. Downtown Toronto Aerial Imagery

The aerial photographs over downtown Toronto were captured using Microsoft's UltraCam-D digital aerial camera. As illustrated in Figure 3.1, the area was covered by three flight strips, with a side overlap of 0.3 and a forward overlap of 0.6. A total of 13 images were captured, and the complete exterior orientation parameters are available. Each image frame has a resolution of 7,500 × 11,500 pixels, with a physical pixel size of 9 µm [47].

Figure 3.1: Aerial Image Strip[47]

The aerial images within 3 strips coverage provided are in .tif format.Based on a visual inspection of

the overlaps and image quality, three representative images were selected as inputs for the photogrammetric 3D reconstruction: `03751.tif`, `03753.tif`, and `03755.tif`. These images belong to a sub-area within flight strip 2 and were chosen to ensure sufficient overlap for stereo reconstruction as well as to capture both high buildings and surrounding ground features. The selected images are shown in Fig. 3.2, where their coverage illustrates the dense urban environment of downtown Toronto, including tall buildings casting long shadows, transportation infrastructure, and open spaces.



(a) Visualization of 03751.tif **(b)** Visualization of 03753.tif **(c)** Visualization of 03755.tif

Figure 3.2: Overall caption describing all three images.

3.1.2. Airborne Laser Scanning (ALS) Data

The ALS data were acquired using an ALTM ORION sensor onboard an aircraft flying at an altitude of 650 m and a speed of 120 knots. The sensor operated at a wavelength of 1064 nm in the near-infrared (NIR) range. It scanned the ground surface with a 20° field of view at a scanning frequency of 50 Hz. The reflected laser echoes were digitized at a sampling rate of 100 kHz, enabling high-density and high-precision 3D point cloud data collection suitable for urban mapping applications[47].

According to the ISPRS benchmark dataset, six flight strips were collected over the downtown Toronto area, as is shown in the left image in Fig. 3.3. The coverage extent of these strips is indicated by the red box in the figure. After visual inspection, it was observed that the dark blue region in the strip map corresponds to the scanning area of strip 2 — the same flight strip selected in the previous subsection for photogrammetric reconstruction. Therefore, the ALS data from the dark blue strip were selected as the ground truth for this benchmark dataset, enabling direct comparison with the point cloud generated from aerial imagery.

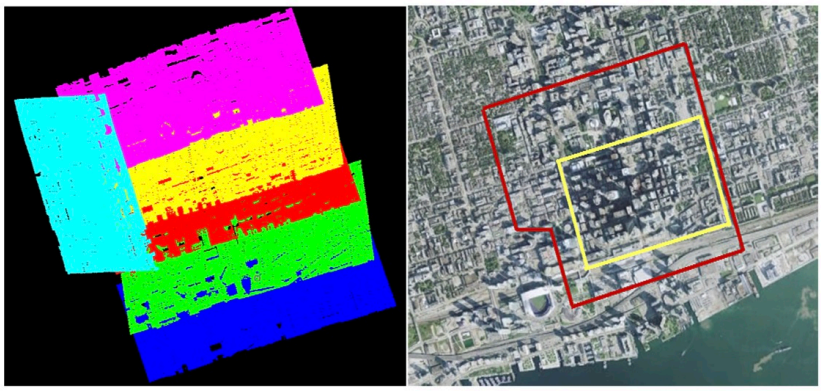


Figure 3.3: Toronto ALS strip and coverage area(Red line on the right image covers the area of ALS strips)[47]

3.2. Datasets of Westkapelle, the Netherlands

Westkapelle is a coastal town located in the province of Zeeland, in the southwestern Netherlands, as is shown in Fig. 3.4. Situated on the western tip of the Walcheren peninsula, it faces the North Sea and is known for its coastal dike system, sandy beaches, and historical sea defenses. The region has a variety of surface types—including urban buildings, dunes, grassland, and engineered sea walls, making it an ideal test site for topographic comparison studies.

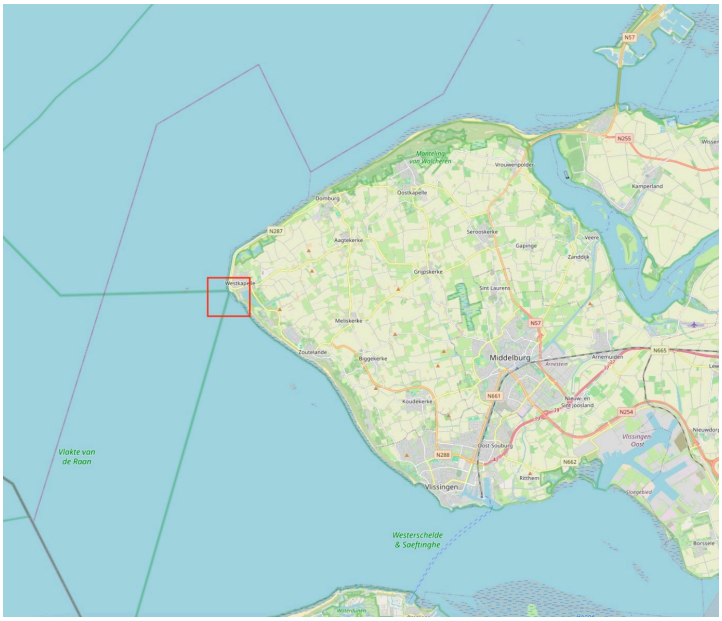


Figure 3.4: Location of Westkapelle, the Netherlands. The red rectangle part covers the research area.

Figure 3.5: Zoom in of research area of Westkapelle, the Netherlands

This section describes two datasets used in this study: a benchmark dataset for validation purposes, and a test dataset for research purpose.

3.2.1. Benchmark LiDAR Data: AHN2 and AHN4

While the Toronto dataset enabled validation of the registration workflow under the condition without time inconsistency, it did not solve another important question that brought in research statement, which may affect the accuracy of point cloud alignment—inconsistency in data acquisition methods. To evaluate the robustness of the registration pipeline under varying acquisition conditions, an additional benchmark was required: one where the datasets were collected using the same acquisition method but at

different time periods. Thus, two national elevation datasets from the Netherlands are selected, the AHN2 dataset and the AHN4 dataset.

The Actueel Hoogtebestand Nederland (AHN) is a national elevation dataset of the Netherlands, based on airborne laser scanning (LiDAR) surveys. It is managed by a consortium of Dutch water boards and governmental agencies and provides high-resolution elevation information with centimeter-level vertical accuracy[3].

The Netherlands is divided into multiple grid regions by the AHN system, with each grid consisting of 25 sub-tiles. Each sub-tile covers an area of approximately 1.25×1.25 kilometers. For each grid area, AHN provides LiDAR point cloud data, along with DSM and DTM products derived from the point cloud, available at spatial resolutions of 5 meters and 0.5 meters. In addition, orthorectified aerial imagery from the years 2016 to 2024 is available to support the DSM products.

The AHN 2 dataset was acquired between 2007 and 2012, while AHN 4 was collected between 2020 and 2022, resulting in a temporal separation of around 10 years. Specifically, the AHN 2 data for Westkapelle were collected in 2007, whereas the AHN 4 data were acquired in 2020 [2]. Both datasets were produced by official Dutch government agencies using airborne LiDAR, ensuring high spatial accuracy, vertical precision, and overall consistency. Furthermore, both are referenced to the same RD New coordinate system (EPSG:28992) and use NAP (Normaal Amsterdams Peil) as the vertical datum.

A summary of the two datasets information is shown in Table 3.1.

Table 3.1: Comparison of AHN2 and AHN4 datasets

Attribute	**AHN2**	**AHN4**
Acquisition Period	2007–2012	2020–2022
Point Number	7,731,302	4,621,620
Point Density	6–10 pts/m²	10–20 pts/m²
Coordinate System	RD New (EPSG:28992)	RD New (EPSG:28992)
Vertical Datum	NAP (Normaal Amsterdams Peil)	NAP
Data Provider	Dutch Government (RWS/Kadaster)	Dutch Government (RWS/Kadaster)
Data Format	LAS 1.2 / LAZ	LAS 1.2 / LAZ

However, running the full registration and validation workflow on the entire Westkapelle area would place heavy demands on computing resources. To shorten processing time, we first examined the complete study region and selected a representative sub-area that contains pronounced geomorphic features and avoids excessive overlap of repeat point clouds, as shown in Fig. 3.6.

Figure 3.6: The Subset Benchmark Area of Westkapelle

3.2.2. Test Dataset: Aerial Imagery and AHN4

Aerial Imagery

Before deciding Westkapelle as the research area, we conducted an extensive survey of all publicly available historical aerial imagery of the Dutch coastline, focusing on data acquired during the 1990s. The primary goal was to identify candidate regions suitable for photogrammetric reconstruction and coastal change analysis.

Most of the datasets were from the TU Delft Library's open access aerial image archives, which contain scanned analog photos from various decades. Due to the historical nature of these datasets, several limitations exist:

- The spatial resolution is generally low, especially compared to modern aerial imagery.
- Some photographs are often only available in non-digital formats (e.g., scanned prints or TIFFs rather than modern lossless formats such as PNG or JPEG2000).
- Camera metadata (e.g., focal length, flight altitude, sensor size) is often incomplete or missing.
- Radiometric quality is limited due to aging of physical prints and scanning artifacts (e.g., color fading, scratches, or noise).
- Georeferencing information is typically absent, requiring manual ground control point (GCP) collection or approximation from external sources.

A summary of the collected datasets is presented in Table 3.2. Based on visual inspection and evaluation of image overlap, the dataset named Kustfotos acquired in 1990 was ultimately selected for this study, as it provided sufficient coverage of the Westkapelle coastal area and exhibited the highest potential for successful photogrammetric reconstruction among the available candidates. The selected historical imagery is shown in Fig. 3.7

Table 3.2: Overview of Aerial Photographs of the Dutch Coastline

Data Name	Source	Author	Date
Kustfotos	TU Delft Library	Unknown	1982, 1987, 1990
Luchtfotos kust Noord Nederland	TU Delft Library	Unknown	1990
Luchtfoto's van de Nederlandse kust - Zeeland	TU Delft Library	Unknown	1994
Luchtfoto's van de Nederlandse kust - Zuid Holland	TU Delft Library	Unknown	1994
Luchtfoto's van de Nederlandse kust - Noord Holland, Texel	TU Delft Library	Unknown	1994
Luchtfoto's van de Nederlandse kust - Waddeneilanden	TU Delft Library	Unknown	1994

(a) Kustfotos of Westkapelle-1 **(b)** Kustfotos of Westkapelle-2 **(c)** Kustfotos of Westkapelle-3

Figure 3.7: Historical Images of Westkapelle

3.3. JARKUS Data

In this study, the JARKUS dataset is employed to validate coastal elevation change results derived from photogrammetric reconstructions and LiDAR analyses. The data were obtained from the public Deltares OpenEarth repository, where the transect overview is available as a KML file and can be visualized in QGIS[18]. This dataset provides cross-shore elevation transects along the entire Dutch coastline. Within the Walcheren coastal cell, the Westkapelle block contains approximately forty such

transects. Each transect originates from a fixed benchmark pole within the village, extends across the dune crest, the sea dike, the intertidal beach, and the upper shoreface, and terminates at approximately -8 m NAP water depth.

For the Westkapelle region, the profile line settings were obtained from shown in Fig. 3.8. Within the area of interest, there are three relevant transects. According to the JARKUS naming convention, their official IDs are `16002195`, `16002215`, and `16002235`, which correspond to 21950, 22150, and 22350 in Fig. 3.8. The green-yellow lines (jarkusraaien) are the JARKUS cross-shore profiles themselves, spaced every 250 m and extending from land out to approximately -200 m, and the oranges dots represent the fixed benchmark pole.

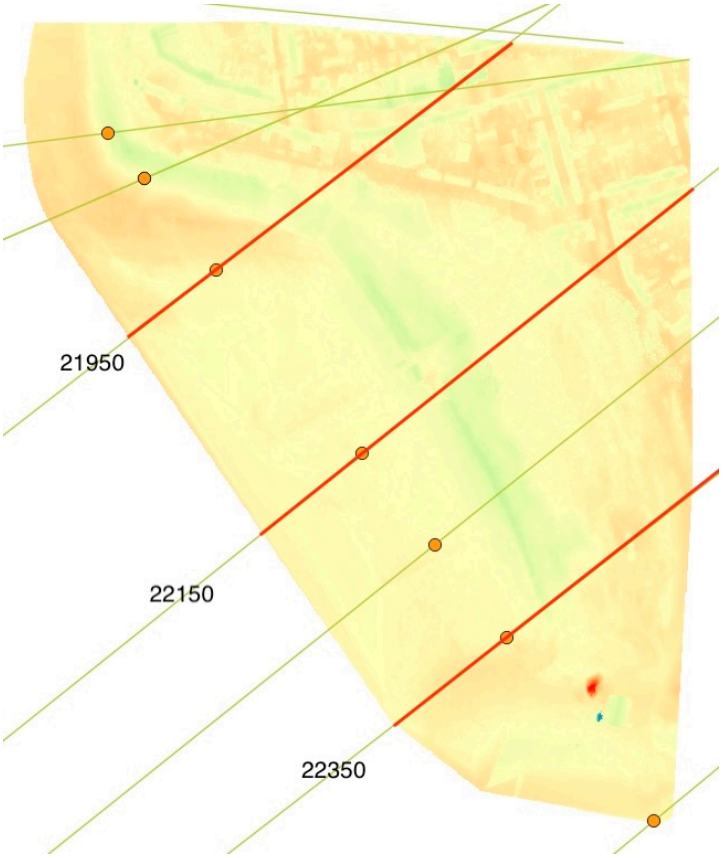


Figure 3.8: JARKUS profiles around Westkapelle

4

Methodology

In this chapter, the overall technical methodology used in this research will be described in detail. Based on the research objectives and theoretical foundations presented in the previous section, this study designs a complete point cloud generation and alignment process to ensure that the point cloud generated from historical images can be well aligned with the ground truth point cloud, so as to ensure that the comparison of the later DEMs is reasonable and credible. The methodology is illustrated using the Westkapelle test dataset, as it is the only study dataset for which the complete workflow can be fully reproduced from start to finish. In contrast, for the benchmark datasets—Toronto Downtown and the Westkapelle benchmark—the specific characteristics of the data meant that certain workflow steps were unnecessary. For example, the Toronto dataset did not require surface masking, while the Westkapelle benchmark dataset did not require point cloud generation from photogrammetry, as such data were already available.

Section 4.1 introduces the photogrammetric processing flow, Section 4.2 discusses the point cloud registration workflow, Sections 4.3 and 4.4 present the evaluation of point clouds and DEMs, and finally, Section 4.5 explains how JARKUS data are used to validate the elevation change.

4.1. Photogrammetry Workflow for Point Cloud Generation

This section presents the complete photogrammetric workflow used to generate high-density 3D point clouds from historical aerial imagery, which generates the foundational dataset for future point cloud registation. Initial step sea masking, was carried out using the Segment Anything model. The remaining photogrammetry steps including feature matching, camera pose estimation, georeferencing, orthophoto generation and dense point cloud reconstruction were performed using Agisoft. Additionally, QGIS was used to collect and manage Ground Control Points (GCPs) to support accurate georeferencing. The overall workflow is illustrated as follow:

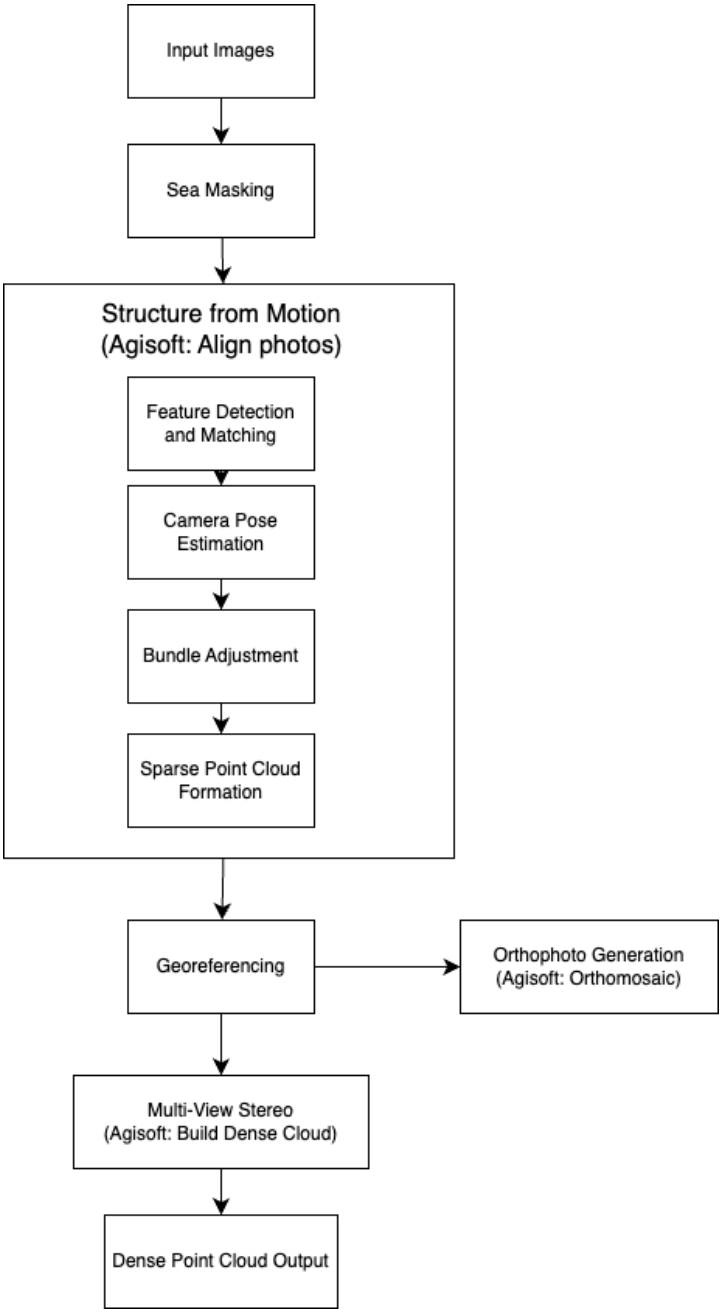


Figure 4.1: Photogrammetry Workflow

4.1.1. Sea Masking

The historical aerial imagery used in this research covers the coastal area of Westkapelle, the Netherlands, where wave patterns are typically observed. Due to temporal difference between image acquisitions and the movement of waves, there is a variation in the position of the waves between images, which violates the assumption that photogrammetry requires temporal consistency among the images used for 3D reconstruction[55, 39], and interferes with the matching of keypoints in the future processing.

In preliminary tests, the results of Metashape's feature detection and matching of historical aerial imagery showed notable mismatches in the wave regions due to the drastic changes in the waves on the water surface. For example, in Fig. 4.2, the red line represents the mismatch point in two overlap images, from where it can be told that the wave's position will lead to mismatch. Theses regions may

show similar textures, although their spatial locations do not correspond. This type of mismatch due to local texture variations in the dynamic water surface directly affects the accuracy of the camera position estimation and lowers the accuracy of the 3D reconstructed point cloud. Therefore, sea masking is introduced in the image pre-processing stage.

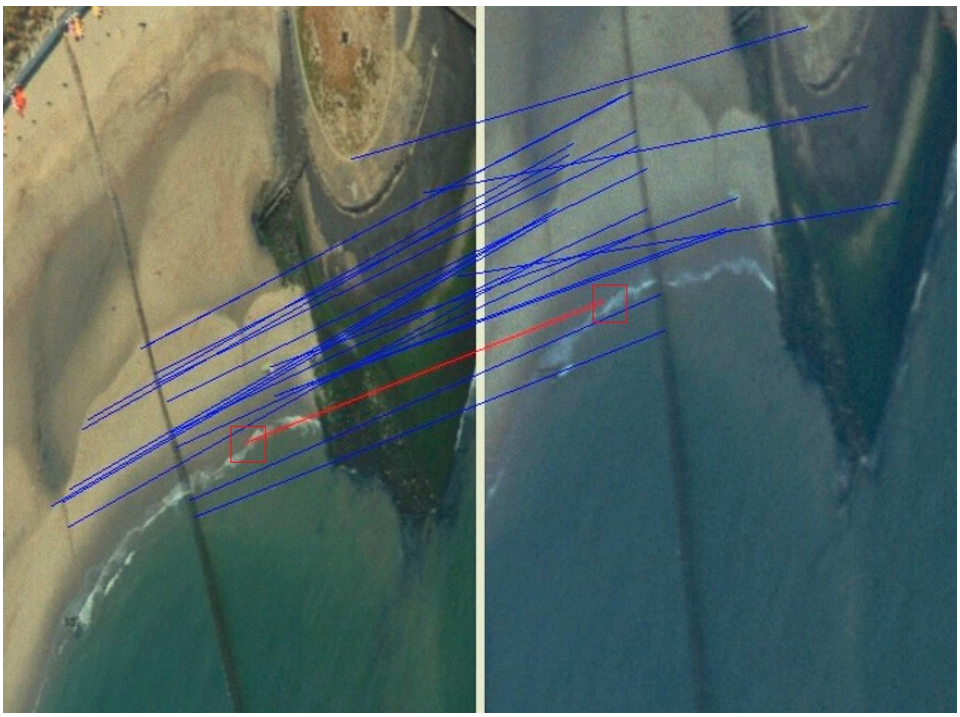


Figure 4.2: Agisoft Tie Points Result: Red Line represents potential falsely matched features, while blue lines represents matching features

The Segment Anything Model (SAM), a general-purpose semantic segmentation framework, is chosen to automatically segment the sea area in the image and generate the masked images[23]. The Segment Anything Model (SAM) is designed to segment objects of interest in images based on user-provided prompts, which can include a single point, multiple points, bounding boxes, or even textual input. The core idea behind this approach is that the model has acquired a general understanding of what constitutes an object, enabling it to segment a wide variety of objects once they are indicated. As a result, SAM demonstrates strong potential for zero-shot segmentation, that is, accurately segmenting previously unseen object categories without the need for additional training [33]. This ability of SAM can accurately separate the water surface area without the need of pixel-level annotation. However, considering that there are still a few possible regions missing from the SAM masking, manual refinement is used to ensure the complete removal of water areas. The manual refinement is particularly for wave edges and shallow water areas, ensuring that the future feature extraction and matching process focuses only on the stable land regions.

4.1.2. Feature Detection and Matching

In photogrammetry, one of the most widely used and classical methods for feature detection is SIFT. It consists of four main stages: scale-space extrema detection, keypoint localization, orientation assignment, and keypoint descriptor generation[1].

The first step, scale-space extrema detection, involves building a scale-space representation of the image using Gaussian filtering. Specifically, the input image $I(x, y)$ is convolved with a Gaussian kernel at $G(x, y, \sigma)$ different scales σ, resulting in blurred images defined as:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{4.1}$$

where the Gaussian kernel equation is:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{4.2}$$

The resulting set of blurred images forms the scale space. The Difference of Gaussians (DoG) is then computed by subtracting adjacent Gaussian-blurred images, which approximates the scale-invariant Laplacian of Gaussian. For each scale level, pixels are compared with their neighbors in the current, previous, and next scales. A pixel is identified as a candidate keypoint if it is a local extreme value.

After detecting a large number of candidate keypoints, the algorithm refines their locations for greater accuracy. This is done by fitting a 3D quadratic function to the local sample points to determine the precise position and scale of each keypoint. Low-contrast points and poorly localized edge responses are discarded to improve robustness.

Once stable keypoints are established, the next step is to assign an orientation to each keypoint to ensure rotation invariance. This is achieved by analyzing the gradient directions in the local neighborhood around each keypoint and constructing an orientation histogram. The peak of the histogram indicates the dominant direction, which is assigned to the keypoint. Finally, based on the keypoint's position, scale, and orientation, a descriptor is computed from the local image region. This descriptor aims to be distinctive and robust against changes in viewpoint and illumination.

In this research, SIFT was carried out using the *Align Photos* function in Agisoft, which can identify keypoints in an image by detecting visually stable features that remain consistent across different viewpoints and lighting conditions, and then the software calculates local descriptors based on the pixel fields surrounding each point to achieve keypoint matching for different images. According to the official documentation, this process is similar to the SIFT algorithm, but with some optimization to improve alignment accuracy. The detected keypoints are the basis for estimating the relative position and orientation of the camera, which is used to construct the initial sparse point cloud.

4.1.3. Camera Pose Estimation and Bundle Adjustment

After finishing feature detection and matching, the next step is to estimate the camera poses. In this study, both the initial Structure-from-Motion (SfM) estimation and the following optimization were performed using Agisoft. Usually, an initial SfM estimation is performed to initialize the 3D scene structure and camera parameters. SfM analyzes image correspondences and camera positions to provide an initial estimation of the 3D structure and camera poses. However, this initial reconstruction often contains errors and requires bundle adjustment to refine the estimated camera positions and orientations.

To refine the reconstruction, Agisoft applies bundle adjustment to simultaneously adjusts camera poses and 3D point positions. Bundle adjustment is an optimization process used to correct errors from the initial 3D reconstruction, including inaccuracies in camera poses, scene structure, or feature tracking. It begins with calculating the reprojection error for each image, which measures the discrepancy between observed 2D features and their current projected positions derived from the 3D structure. A loss function is then defined, typically as the sum of the reprojection errors across all images. The task is to minimize this loss function, which can be solved using algorithms such as the Gauss-Newton method or the Levenberg-Marquardt algorithm [67].

4.1.4. Georeferencing

After the reconstruction and bundle adjustment, a set of optimized camera parameters and a sparse 3D point cloud can be obtained. However, one of the typical characteristics of point clouds generated from historical imagery is that they do not contain any global geographic coordinate information.

In order to align the reconstructed point cloud through photogrammetry with the AHN data, it is important that both datasets are referenced in the same global coordinate system. Without such alignment, significant differences in position and scale may exist between the two datasets, making direct registration unfeasible. Furthermore, since the reconstructed point cloud is later used to generate a DEM and compare it with an AHN derived DEM, it must be georeferenced to a known global spatial reference system.

Acquisition of Ground Control Points

Ground Control Points (GCPs) play an important role in photogrammetry workflow, since it enables accurate georeferencing of aerial imagery to real-world coordinate systems. GCPs are physical features or locations with known coordinates used to align the reconstructed 3D model.

Under ideal conditions, images can be identified by direct georeferencing, which needs the real-time recording of the camera's spatial position and attitude while the image is being acquired. This requires a high-precision GNSS system and synchronization of the trigger moment of image acquisition with navigation system sampling [52]. Another approach is to lay high-contrast artificial markers (e.g., checkerboard targets) at the scene of the images and identify them. However, for the historical aerial imagery used in this research, neither of these two conditions are available: on the one hand, the images were acquired long ago and do not have metadata synchronized with the satellite navigation system; on the other hand, the images do not contain known control markers manually deployed. Therefore, a manual extraction by combining modern dataset and historical images is required.

In order to realize the initial rough alignment, the AHN dataset released by the Dutch government is used as a reference, which contains high-precision point cloud data obtained by laser scanning. Meanwhile, a DSM with a resolution of 1.0 m is obtained by linear interpolation based on the AHN4 point cloud data. After obtaining the DSM, combining the historical images and the interpolated modern DSM to manually select the GCPs in the QGIS software.

GCPs are selected according to the following basic rules: firstly, the points must be recognizable in both the historical images and the modern images, and secondly, the selected points should be relatively stable features, such as road intersections, corners of buildings, or fixed terrain structures, to ensure the long-term spatial consistency. The selected GCPs are shown in Fig. 4.3.

(a) Selected GCPs from AHN4 derived DSM

(b) Corresponding GCPs in Historical Images

Figure 4.3: Comparison of selected ground-control points (GCPs) in the DSM and the historical photographs

Georeferencing

After identifying the GCPs in the historical images, their corresponding real-world coordinates were obtained using QGIS. Specifically, the XY positions were extracted by querying the modern DSM base layer, while elevation values were assigned using the DSM. And it is important to ensure that the order of GCPs in QGIS matched the order of markers later defined in Metashape. These GCPs, along with their accurate real-world XYZ coordinates, were then imported into Agisoft. Each GCP was defined as a marker in the image block, and the georeferencing was completed through alignment and bundle adjustment using the imported control point data.

The quality of the alignment after importing the GCPs is summarized in Table 4.1. The accuracy refers to the manually assigned coordinate accuracy for each control point, typically set to 0.005 m in Agisoft. The error (m) represents the spatial residual error after bundle adjustment, while error (pix) indicates the image-space residual in pixels, reflecting possible inaccuracies in GCPs marking.

In this research, all GCPs were assigned an accuracy of 0.005 m with an expectation of centimeter level georeferencing precision. However, the actual spatial errors ranged from 0.084 m to 0.588 m, with a mean of 0.302 m. The fourth GCP showed the largest error, which, as shown in Fig. 4.3a, is located near a corner area of the DSM. This point was difficult to accurately identify in the historical image, and only an approximate location could be marked.

Regarding the image-space residuals, the errors ranged from 62.445 to 209.617 pixels, with an average of 130.954 pixels, which is high. This could be attributed to the low resolution of the historical images, challenges in accurately marking GCPs, or incomplete convergence of the bundle adjustment. Therefore, while this initial georeferencing provides a rough geospatial reference for the reconstructed point cloud, further refinement through precise registration is still necessary.

GCPs ID	Accuracy (m)	Error (m)	Error (pix)
1	0.005	0.084	62.445
2	0.005	0.099	137.012
3	0.005	0.205	68.025
4	0.005	0.588	209.617
5	0.005	0.169	115.074
6	0.005	0.207	130.221
7	0.005	0.405	138.368
Mean Error (m)		0.302 **m /** 130.954 **pix**	

Table 4.1: Summary of GCP projection errors in historical image alignment.

4.1.5. Dense Point Cloud Generation

After the sparse 3D reconstruction and georeferencing, dense surface reconstruction was performed using the Build Dense Cloud function in Agisoft, which is implemented under the principle of Multi-View Stereo algorithm. It computes depth maps for each calibrated image based on the previously estimated camera parameters, and merges them into a dense point cloud.

The resulting dense cloud significantly increases point density compared to the sparse model and captures finer surface features, which are essential for accurate elevation modeling. Usually, after generating the dense point cloud, a confidence filtering operation is applied to remove low-confidence points in order to avoid the influence of mismatches or poorly constrained depth estimates on the final point cloud. However, a special aspect of this study is that the historical images used for 3D reconstruction showed relatively low reconstruction quality due to image poor resolution, and insufficient image overlap. Therefore, under such conditions, if confidence filtering is still applied to assess point cloud quality, a large number of points would be removed, especially in low-texture areas such as bare soil or vegetation zones, which will reduce the coverage and density of the resulting surface model, and making the subsequent DSM generation difficult. As a result, no filtering was applied to the reconstructed dense point cloud in this study; instead, the originally generated dense cloud was directly retained.

4.1.6. Orthophoto Generation

After completing the above steps, we have obtained all the essential inputs required for orthophoto generation:

- Multi-view imagery
- Image orientation parameters, including both interior orientation and exterior orientation, which are derived from SfM
- Georeferenced dense point clouds

Although the orthophoto is generated through a single operation step Orthomosaic in Agisoft, Agisoft internally executes the following four sequential processes:

1. Extraction of a surface model from the dense point cloud.
2. Orthorectification by intersecting camera rays with the DSM, thereby eliminating relief displacement and correcting for camera tilt.
3. Seamline optimization and blending of overlapping image regions to ensure visual consistency.
4. Resampling and exporting the final image as a GeoTIFF at the target ground sampling distance (5 cm in this case), using the project's coordinate reference system (EPSG:28992).

4.2. Point Cloud Registration

This section introduces the process of aligning the reconstructed point cloud from historical aerial images with the ground truth obtained from the AHN dataset. The registration workflow consists of several steps: surface masking, persistent analysis, voxel downsampling, and registration.

The registration between these two datasets faces two main challenges, temporal inconsistency and differences in data sources and acquisition methods. To ensure the feasibility of the point cloud registration process and the reliability of the final result, the workflow is not only applied to experimental data, but validated on two benchmark datasets.

4.2.1. Surface Masking

In this research, both the reconstructed point cloud and the AHN point cloud contain several million points. Because the area of interest contains large part of land, one distinguishing characteristic of the dataset is the large quantity of ground points. These nearly planar surface regions share very similar, or even identical geometric characteristics, meaning their feature descriptors can be almost the same. Lacking uniqueness, they make the later registration algorithm prone to the sliding effect of point-cloud registration[48]. To mitigate this problem, a two-stage preprocessing workflow is adopted. The first stage performs a coarse segmentation and removal of large, flat surfaces, in particular, the ground.

Flat surface removal was carried out by using CloudCompare and Python. Firstly a visual inspection of the point cloud is performed in CloudCompare. This step is to identify the general elevation range corresponding to ground-level points. Based on this inspection, an approximate elevation threshold is defined. All points falling below this threshold representing low-lying or flat areas with limited topographic variation, thus, will be removed from the dataset. It is worth noting that this visual inspection only provides a rough approximation of the majority of ground points, which means it cannot eliminate the point with similar features perfectly. Thus, besides serving as a coarse solution to avoid the sliding problem, this step is more of a method to reduce the point cloud size for the later persistence analysis, which can accelerate the algorithm.

4.2.2. Persistent Analysis

After removing the ground points from the point cloud, the next step is to perform persistence analysis. Although the preceding surface-filtering stage effectively deletes ground data, eliminating the ground alone does not fully prevent the sliding issue discussed earlier. Hence, a second stage based on persistence analysis is introduced.

Persistence analysis identifies geometrically salient key-points across multiple spatial scales. Its objective is to suppress the many repetitive or common features in the dataset—such as extensive flat ground—that appear frequently but contribute little to robust matching [48].

The persistent analysis algorithm comprises these steps:

Normal Estimation

Before computing descriptors, the normal vector for each point must be estimated. The core idea is to use Principal Component Analysis (PCA) on a point's local neighborhood, defined by a fixed search radius r_{norm}, to determine the local surface orientation.

Let $\mathcal{N}_i$ denote the set of neighboring points of the query point $\mathbf{p}_i$, and let $|\mathcal{N}_i|$ be the number of points in this neighborhood. The centroid of the neighborhood $\bar{\mathbf{p}}_i$ is defined as:

$$\bar{\mathbf{p}}_i = \frac{1}{|\mathcal{N}_i|} \sum_{\mathbf{p}_j \in \mathcal{N}_i} \mathbf{p}_j. \tag{4.3}$$

The 3×3 covariance matrix describing the local surface variation is:

$$\mathbf{C}_i = \frac{1}{|\mathcal{N}_i|} \sum_{\mathbf{p}_j \in \mathcal{N}_i} (\mathbf{p}_j - \bar{\mathbf{p}}_i)(\mathbf{p}_j - \bar{\mathbf{p}}_i)^\top. \tag{4.4}$$

The eigenvector $\mathbf{v}_{\min}$ corresponding to the smallest eigenvalue $\lambda_{\min}$ of $\mathbf{C}_i$ is taken as the estimated normal direction:

$$\mathbf{C}_i \mathbf{v}_{\min} = \lambda_{\min} \mathbf{v}_{\min}, \quad \mathbf{n}_i = \frac{\mathbf{v}_{\min}}{\|\mathbf{v}_{\min}\|}, \tag{4.5}$$

where $\mathbf{n}_i$ denotes the unit surface normal vector at point $\mathbf{p}_i$.

At last, because surface normals have a direction ambiguity, a consistent orientation (e.g., aligned with the positive z-axis or oriented towards a viewpoint) must be enforced across the entire point cloud. Without this step, inconsistent normal directions can lead to unstable feature descriptors and incorrect matches.

Fast Point Feature Histogram

With the oriented normals estimated in the previous step, we compute the FPFH to describe the local surface geometry around each point. In our implementation, FPFH is computed using the normals from radius r_{norm} and neighborhoods defined by a fixed search radius r_{feature}. For each point, we loop over all neighbors within r_{feature} and choose an ordered source–target pair based on the smaller viewing angle to ensure a stable orientation. A local coordinate frame is then built from the source normal and the vector connecting the two points, from which the three angular features (α, ϕ, θ) are calculated.

All angular feature triples are binned into a 3D histogram with `div` bins per dimension over the valid ranges, producing a histogram vector for each point. Finally, the fast aggregation step adds the point's own histogram to a weighted sum of its neighbors' histograms, where weights are the inverse of the point-to-neighbor distances.

Single-scale Persistent Analysis

Once the FPFH descriptors have been computed for all points, they are used to identify rare points through single-scale persistence analysis. For every selected radius r_i, we perform a neighbourhood search and keep only those points whose neighbour counts lie between $n_{\min} = 5$ and $n_{\max} = 200$, thereby excluding regions that are too sparse or overly dense. Surface normals are estimated for each valid point using r_{norm} nearest neighbours, and FPFH descriptors are computed using r_{feature} nearest neighbours. Each FPFH is a 33 dimensional vector.

By arranging the FPFH descriptors of all valid points as rows, we obtain a feature matrix $\mathbf{F} \in \mathbb{R}^{N \times 33}$, where N is the number of points and each row represents the descriptor of a single point. The mean histogram is then calculated as:

$$\boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{f}_i, \tag{4.6}$$

where $\mathbf{f}_i$ denotes the FPFH of the i-th point.

For each point, the Euclidean distance $d_i = \|\mathbf{f}_i - \boldsymbol{\mu}\|_2$ is computed. A threshold equal to the 90^{th} percentile of $\{d_i\}$ is applied; points with d_i exceeding this threshold are marked as outliers and classified as rare points, which is the so called key points as well.

Multi-scale Persistent Analysis

While the single-scale analysis identifies rare points at a fixed radius, it may be sensitive to the specific scale chosen. To improve the robustness of key point detection, persistence analysis is performed at multiple spatial scales. In this study, three radius are selected:

$$r_i = \left[2\, d_{\text{mean}},\ 2\sqrt{2}\, d_{\text{mean}},\ 4\, d_{\text{mean}}\right],$$

where d_{mean} denotes the average nearest-neighbor distance within the point cloud. The first radius $2\, d_{\text{mean}}$ captures fine scale geometric variations, the second radius $2\sqrt{2}\, d_{\text{mean}}$ corresponds to an intermediate scale that is larger by a factor of $\sqrt{2}$ in area coverage, and the third radius $4\, d_{\text{mean}}$ captures coarser structures while preserving significant topographic features.

For each radius r_i, a single-scale persistence analysis is performed, yielding a set $R(r_i)$ of points identified as rare (i.e., geometrically salient) at that scale. The final set of persistent key points is obtained by intersecting the results across all k scales:

$$K = \bigcap_{i=1}^{k} R(r_i), \tag{4.7}$$

where K contains only those points that are consistently identified as rare across multiple scales. The whole algorithm to realize multi-scale persistent analysis is shown in below:

Algorithm 1 Multi-Scale Persistence Analysis

Input: point cloud $P = \{\mathbf{p}_i\}_{i=1}^{N}$; radius set $\mathcal{S} = \{r_1, r_2, r_3\}$; neighbor count bounds $[n_{\min}, n_{\max}]$; r_{norm}; r_{feature}
Output: persistent key-points P_f

Normal estimation. Estimate surface normals via PCA based on radius r_{norm}, and enforce consistent orientation
for all $r \in \mathcal{S}$ **do** // Single-scale persistence at radius r
  For each $\mathbf{p}_i$, build the neighborhood $\mathcal{N}_i(r)$
  Form the valid set $V(r) = \{\, i \mid n_{\min} \le |\mathcal{N}_i(r)| \le n_{\max} \,\}$.
  Compute FPFH for each valid point.
  Stack row-wise to obtain feature matrix $\mathbf{F}(r)$; compute the mean feature vector $\boldsymbol{\mu}(r)$
  for all For each $i \in V(r)$ **do**
    Compute the distance $d_i(r)$.
    Compute the threshold $\tau(r)$ as the 90^{th} percentile of $\{d_i(r) \mid i \in V(r)\}$.
    Define the rare set at scale r:

$$R(r) = \{\, i \in V(r) \mid d_i(r) > \tau(r) \,\}.$$

  end for
  if $R(r) = \varnothing$ for all $r \in \mathcal{S}$ **then**
    $P_f \leftarrow \varnothing$
  else
    Multi-scale persistence (intersection):

$$K = \bigcap_{r \in \mathcal{S}} R(r), \qquad P_f \leftarrow \{\, \mathbf{p}_i \in P \mid i \in K \,\}.$$

  end if
end for
return P_f

4.2.3. Voxel Downsampling

After applying sea masking and persistence analysis to both the reconstructed and the AHN point clouds, it was observed that the two datasets have different point density and spatial distribution. To address these two problems, voxel downsampling was applied to the reconstructed point cloud and AHN point cloud.

The first issue concerns the variation in point density between the two datasets. Point clouds generated from historical aerial images tend to be much denser in areas with clear textures and many details, but are sparse in blurry or occluded areas. In contrast, AHN usually have a more uniform distribution across the entire scene. This mismatch in density brings great challenges for registration, because many algorithms, such as FPFH or ICP, assume that the point clouds from different datasets have roughly similar sampling density. If this assumption is violated, it may lead to inaccurate alignment.

The second issue relates to the uneven spatial distribution of points after masking. Surface masking, particularly when removing large planar ground surfaces, can completely eliminate certain areas of the point cloud. This uneven spatial distribution can cause problems in the later registration process. For example, when calculating geometric features for each point, we usually search for neighboring points within a certain radius. But if the spatial distribution is inconsistent, some points might have many neighbors, while others have very few. This difference in neighborhood size affects the stability of feature calculation and can make the matching results unreliable.

To solve these issues, voxel downsampling was applied to the point cloud. The principle of voxel downsampling is firstly, dividing the input point cloud into a 3D grid of small cubes called voxels. For each voxel, all the points inside it are replaced by their centroid. This reduces the total number of points while still keeping the overall shape and structure of the point cloud. The method is efficient and produces a more evenly distributed point cloud.

4.2.4. Registration

After sea masking, persistence analysis and voxel down-sampling, both datasets yield point clouds with approximately the same point density and a more uniform spatial distribution. The persistent keypoints extracted from the reconstructed photogrammetric cloud form the source keypoint cloud, denoted P_k^s, while the persistent keypoints from the AHN data form the target keypoint cloud, denoted P_k^t, which serves as the reference. The goal of registration is to estimate a rigid transformation that aligns P_k^s to P_k^t in the same coordinate frame, enabling direct terrain comparison.

To prevent the initial pose of the source cloud from overly influencing fine registration, a two-stage strategy is adopted in this research:

- Coarse alignment: using Sample Consensus Initial Alignment (SAC–IA) to roughly match the position and orientation of the two clouds.
- Refine registration: using Iterative Closest Point (ICP) to refine the result, with the goal of achieving centimetre-level accuracy.

Coarse alignment

SAC–IA performs feature-based registration using FPFH. Before computing FPFH, surface normals must be estimated. Two parameters are critical:

- r_{normal}: neighbourhood radius for normal estimation.
- r_{feature}: neighbourhood radius for FPFH computation.

In practice, r_{normal} is typically $2 \sim 3$ times the average point spacing, while r_{feature} is $3 \sim 5$ times. To identify optimal values, SAC–IA is run iteratively with different parameter combinations.

Once normals are computed, an FPFH descriptor is built for each point in P_k^s and P_k^t. A KD-tree is constructed for the target descriptors, and for each source descriptor the $k=5$ nearest neighbours are retrieved, giving each source point five candidate matches.

The SAC–IA stage then enters a RANSAC loop:

1. Randomly select four keypoints from P_k^s.
2. For each keypoint, randomly choose one of its five candidate matches in P_k^t to form four correspondence pairs.
3. Estimate a rigid transform T_0 from these pairs using the singular value decomposition (SVD) method, where SVD factorises a matrix into orthogonal rotation components and scaling factors,

enabling the optimal rotation and translation that minimise the squared alignment error to be found.

4. Apply T_0 to P_k^s and count inliers using a KD-tree built on the target cloud's XYZ coordinates. Inliers mean points whose matched distance in P_k^t is less than $2\,r_{\text{feature}}$.

The loop stops when the inlier count exceeds a threshold or 1000 iterations are reached. The transform with the highest inlier count is retained and applied to the full-resolution reconstructed cloud, yielding the coarsely aligned cloud P^*.

As shown in Algorithm 2, the alignment procedure first applies SAC–IA for coarse registration:

Algorithm 2 Coarse Registration via SAC–IA

Input: Source cloud P_k^s, target cloud P_k^t ; feature radius range $\mathcal{R}_{\text{feature}}$; normal radius range $\mathcal{R}_{\text{normal}}$; neighbour count bounds
Output: Coarse rigid transform $\mathbf{T}_{\text{best}}$ aligning P_k^s to P_k^t

Assume *all_inliers* $= 0$, $\mathbf{T}_{\text{best}} = \mathbf{I}$
for $r_{\text{feature}} \in \mathcal{R}_{\text{feature}}$ **do** // outer loop: FPFH radius
  for $r_{\text{normal}} \in \mathcal{R}_{\text{normal}}$ **do** // inner loop: normal radius
    Calculate Normals & features on P_k^s: estimate normals based on r_{normal} and neighbour bounds; build KD-tree and collect neighbours within r_{feature}; then compute FPFH $\mathbf{F}_S$
    Calculate Normals & features on P_k^t: same steps, obtain $\mathbf{F}_T$

    Build KD-tree on $\mathbf{F}_T$ and on P_k^t // SAC–IA Algorithm
    Pre-compute 5-NN feature correspondences for every source point
    Assume *best_inliers* $= 0$, $\mathbf{T}_{\text{loc}} = \mathbf{I}$
    for *iter* $= 1$ **to** 1000 **do**
      Randomly pick a minimal set of source indices; choose one feature-space match for each
      Compute provisional rigid transform $\mathbf{T}$
      Apply $\mathbf{T}$ to P_k^s; count inliers wrt. P_k^t within $\varepsilon = 2r_{\text{feature}}$
      if current inliers $>$ *best_inliers* **then**
        $\mathbf{T}_{\text{loc}} = \mathbf{T}$, *best_inliers*$=$ current inliers
      end if
    end for
    if *best_inliers* $>$ *all_inliers* **then**
      $\mathbf{T}_{\text{best}} = \mathbf{T}_{\text{loc}}$, *all_inliers*$=$ *best_inliers*
    end if
  end for
end for
return $\mathbf{T}_{\text{best}}$

Refine registration

The ICP algorithm refines the coarse alignment by iteratively finding the nearest neighbour in full target cloud $\mathcal{T}$ for each point in P^*, then solving for the rigid transformation that minimises the mean squared error between all matched pairs via SVD. This process repeats until convergence or until the maximum number of iterations is reached.

Algorithm 3 Refinement via ICP

Input: Coarse transform $\mathbf{T}_{\text{best}}$ from Alg. 2; full source cloud $\mathcal{S}$, target cloud $\mathcal{T}$; rejection threshold τ
Output: Refined rigid transform $\mathbf{T}_{\text{final}}$ (coarse $\rightarrow$ fine)

```
Apply T_best to the full source cloud get P*
Assume  T_icp = I
for k = 1 to 1000 do
    For each point in P*, find closest point in T via KD-tree
    Reject pairs with distance > 3τ
    Compute incremental transform ΔT from the accepted pairs
    Update pose: T_icp = ΔT T_icp
    Compute mean registration error ε_k;   if |ε_k − ε_{k−1}| < τ then break
end for
Compose transforms: T_final = T_icp T_best
return T_final
```

4.3. Point Cloud Registration Evaluation

The accuracy of the point cloud registration was evaluated using both qualitative and quantitative approaches. The source point cloud (photogrammetric reconstruction) and the target point cloud (reference AHN/ALS dataset) were compared after the final rigid transformation was applied.

For qualitative assessment, the aligned source and target point clouds were visually inspected in CloudCompare by overlaying them in a common reference frame. For quantitative evaluation, cloud-to-cloud (C2C) distances were computed between the registered source and target using the nearest-neighbour search in Euclidean space. These distances $\{d_i\}_{i=1}^{N}$ were summarised as a histogram, where bin centres represent distance intervals (in metres) and counts indicate the number of points within each range.

From the histogram data, the following statistical measures were calculated:

$$M_{\text{med}} = \text{median}\{d_i\}, \tag{4.8}$$

$$M_{\text{mean}} = \frac{1}{N}\sum_{i=1}^{N} d_i, \tag{4.9}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - M_{\text{mean}})^2}, \tag{4.10}$$

$$d_{\min} = \min\{d_i\}, \tag{4.11}$$

$$d_{\max} = \max\{d_i\}, \tag{4.12}$$

$$\tag{4.13}$$

4.4. Digital Elevation Model Formation and Evaluation

This section outlines the procedures for generating Digital Elevation Models (DEMs) from the aligned point clouds and assessing their accuracy. DEM formation involves interpolating the registered 3D point cloud data onto a regular grid to produce a continuous elevation surface. The evaluation step then compares the photogrammetric DEMs against reference DEMs derived from AHN or ALS data, quantifying vertical deviations through statistical measures and spatial difference maps.

4.4.1. Derivation of DEMs

To enable direct elevation comparison and quantitative analysis, it was necessary to generate DEMs from both the photogrammetrically reconstructed point clouds and the reference AHN datasets. The DEM generation procedure was identical for both the reconstructed and reference datasets to ensure methodological consistency.

In this study, two types of LiDAR datasets were used as reference elevation data:

- Actueel Hoogtebestand Nederland (AHN) — a national airborne laser scanning dataset covering the entire Netherlands.

- Airborne Laser Scanning (ALS) survey — an independent LiDAR dataset covering the Toronto study site.

Both datasets are based on airborne LiDAR measurements and provide high-resolution elevation information with centimetre-level vertical accuracy. Here, we describe the DEM generation procedure using AHN as an example; the same steps were applied to the ALS dataset and the reconstructed point cloud dataset.

The AHN is a national elevation dataset of the Netherlands, and divide the Netherlands into multiple grid regions, with each grid consisting of 25 sub-tiles. Each sub-tile covers an area of approximately 1.25×1.25 kilometers. For each grid area, AHN provides LiDAR point cloud data, along with DSM and DTM products derived from the point cloud, available at spatial resolutions of 5 meters and 0.5 meters.

However, the study area in this research is relatively small, covering approximately the extent of a single sub-tile. For individual sub-tiles, AHN only provides raw point cloud data and does not offer a pre-generated DEM for the area. Therefore, the ground truth DEM required for comparison had to be generated manually from the point cloud data.

Interpolation and rasterization were carried out using a custom Python script. The point cloud data, provided in .las format, was first imported and the X,Y,Z coordinates of each point were extracted. A regular spatial grid with a resolution of 1 meter was then constructed. Linear interpolation was applied to project the discrete elevation values onto the grid, resulting in a continuous DEM. Throughout the processing workflow, the coordinate reference system was assumed to be the Dutch national projection RD New (EPSG:28992), in order to ensure consistency with AHN standard data.

The DEM from the reconstructed point cloud was generated using the same Python script. Specifically, the X,Y,Z coordinates of the point cloud were extracted, a regular spatial grid was constructed, and linear interpolation was applied. The resulting DSM was assigned the RD New projection (EPSG:28992), consistent with the AHN data, to enable subsequent elevation comparison and error evaluation.

4.4.2. DEMs Evaluation

After DEMs derivation, two DEMs were generated: a photogrammetric DEM reconstructed from historical aerial imagery, and a AHN derived DEM. Both DEMs were co-registered in RD New and referenced to NAP.

DEM Qualitative Analysis

A visual inspection of the two DEMs was performed to assess their morphological similarity along the coastal zone. While qualitative analysis cannot yield quantitative accuracy measures, it can reveal large-scale discrepancies, shoreline shifts, and morphological changes in dune and intertidal zones. When both DEMs are of sufficiently high quality and spatial resolution, visible differences in elevation and shoreline position can indicate measurable coastal change.

DEM Difference Qualitative and Quantitative Analysis

To evaluate coastal morphological change, a DEM difference map was computed:

$$\Delta h \;=\; \mathrm{DEM}_{\mathrm{gt}} \;-\; \mathrm{DEM}_{\mathrm{photo}}, \tag{4.14}$$

where $\mathrm{DEM}_{\mathrm{photo}}$ represents the DEM derived from photogrammetric point cloud and $\mathrm{DEM}_{\mathrm{gt}}$ represents DEM derived from AHN point cloud. Positive values indicate elevation gain compared to the past, and negative values indicate erosion.

Qualitative analysis involved visually inspecting the Δh map to identify spatial patterns of erosion and accretion, particularly along the shoreline, dune crests, and intertidal flats. Distinct clusters of positive or negative values were interpreted as zones of morphological change.

To quantify vertical accuracy we computed the following metrics for the full set of residuals:

$$\text{Median} \quad \tilde{h} = \text{median}(\Delta h_i), \tag{4.15}$$

$$\text{Mean} \quad \bar{h} = \frac{1}{N}\sum_{i=1}^{N}\Delta h_i, \tag{4.16}$$

$$\text{Standard deviation} \quad \sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\Delta h_i - \bar{h}\right)^2}. \tag{4.17}$$

To reduce the influence of outliers, the extreme values were characterised using the 5% lower and upper quantiles of the distance distribution, hereafter referred to as the robust minimum and robust maximum, respectively. Unlike the absolute minimum and maximum, which can be dominated by a few erroneous or noise-induced points, the robust minimum and maximum provide a more stable representation of the typical range of deviations by ignoring the most extreme 5% of values at both ends of the distribution.

$$\text{Robust minimum} \quad \Delta h^*_{\text{min}} = Q_{0.05}(\{\Delta h_i\}), \tag{4.18}$$

$$\text{Robust maximum} \quad \Delta h^*_{\text{max}} = Q_{0.95}(\{\Delta h_i\}), \tag{4.19}$$

where Q_p denotes the p-th percentile. Under an approximately normal error model, about 95% of residuals are expected within $\pm 2\sigma$, while the lowest/highest 5% represent the tails by design.

4.5. JARKUS Processsing and Combining with DEMs

All profile handling was automated with the open-source JARKUS Analysis Toolbox (JAT)[59]. The toolbox encapsulates download, datum conversion, resampling and uncertainty analysis in a set of MATLAB/octave functions.

4.5.1. Elevation Change Profile Supported by JARKUS

JARKUS (*JAarlijkse KUStmetingen*) provides annual cross-shore transect data along the Dutch coastline, including profile elevation values, shoreline (RSP) coordinates, and other derivatives. In the study area, three transects were selected: ID16002195, 16002215, and 16002235. These IDs correspond to fixed survey lines from the Rijkswaterstaat database and were chosen because they intersect the main coastal section of interest and have long, consistent time series.

Each profile represents the cross-shore distance from the Dutch Rijks Strandpalen (RSP) line—a baseline defined by numbered beach poles spaced roughly 100 m apart along the coast. The pole located at the coastline is defined as $x = 0$ m. Distances are positive ($x > 0$) landward of the RSP line and negative ($x < 0$) seaward (on dunes or dikes).

To facilitate inter-annual comparisons, we computed elevation-change profiles by differencing yearly elevation profiles along the same transect (aligned by cross-shore distance). This can be expressed as:

$$\Delta z_{y_a,y_b}(x) = z_{y_a}(x) - z_{y_b}(x), \tag{4.20}$$

where $\Delta z_{y_a,y_b}(x)$ is the elevation change at cross-shore coordinate x between survey years y_a and y_b, and $z_y(x)$ is the elevation profile from year y referenced to the NAP vertical datum. In this study, for the JARKUS-based profiles y_a corresponds to the year of the historical aerial imagery (1990) and y_b corresponds to the year of the AHN4 survey (2020). A concise Python script was used to convert raw yearly profiles into single difference curves for each transect and data source.

4.5.2. Elevation Change Profile Supported by DEMs

To obtain the correct corresponding elevation change profile supported by DEMs, it is necessary to accurately locate the spatial positions of these three transects in the raster and extract the corresponding raster values. In this study, the following procedures were accomplished with the help of QGIS 3.9.1:

- **Load transect vectors**
  Download `JarKus_transects_overview.kml` from the JAT data repository and load it into QGIS via *Layer* → *Add Layer* → *Add Vector Layer*. The file contains roughly 14 000 JARKUS transect polylines.
- **Select transects of interest**
  Open the attribute table, click *Select by expression*, and enter `"ID" IN (16002195,16002215,16002235)`. Save the highlighted features to a new layer named `selected_transects`.
- **Extract elevation profiles**
  In the QGIS *Processing Toolbox* run "*Elevation profile*" Use `selected_transects` as the input lines, the DEM Difference as the raster, and export the results as a `CSV`.

After the QGIS workflow we extracted elevation–change profiles (DEM_{2014} – DEM_{1990}) for JARKUS transects 16002195, 16002215, and 16002315 (Fig. **??**). The only outstanding preprocessing step is to align the cross-shore axes of the two survey epochs. Using the RSP line ($x = 0$ m) as the common origin, we clip each profile to the landward segment extending 200 m inland ($-200 \leq x \leq 0$ m). All subsequent analyses are performed on this 200 m-long elevation-change profile.

4.5.3. Quantitative Metrics of Elevation Change Profiles

To quantitatively compare the results from the two data sources, four statistical metrics were computed for each elevation change profile. The calculations were performed along the cross-shore axis of each transect as follows:

$$\text{Mean} = \frac{1}{N}\sum_{i=1}^{N}\Delta z(x_i), \tag{4.21}$$

$$\text{Min} = \min_{1\leq i\leq N}\Delta z(x_i), \tag{4.22}$$

$$\text{Max} = \max_{1\leq i\leq N}\Delta z(x_i), \tag{4.23}$$

$$\text{Std} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left[\Delta z(x_i) - \text{Mean}\right]^2}, \tag{4.24}$$

where N is the number of cross-shore sampling points along the transect, x_i is the i-th cross-shore coordinate referenced to the RSP baseline ($x = 0$ m), and $\Delta z(x_i)$ is the elevation change at x_i between two survey years.

These metrics provide a concise yet comprehensive description of profile differences: the mean indicates net accretion (> 0) or erosion (< 0), the minimum and maximum capture the extremes of change, and the standard deviation quantifies spatial variability along the profile. Applying the same set of metrics to both JARKUS- and DEM-based profiles ensures direct and consistent quantitative comparison between the two methods.

5

Results

5.1. Registration Settings and Parameters

In this section, the parameter settings used for point cloud registration across three datasets(Toronto, Westkapelle, and the Test dataset) are described. Two key parameters-the surface normal estimation radius r_n and the FPFH feature radius r_f-were systematically tuned. Parameter ranges were selected based on empirical guidelines and dataset specific characteristics, such as mean point spacing and the presence of repetitive or irregular structures. The final parameter choices were determined through multiple trial runs, with the optimal settings selected according to the highest inlier counts and lowest alignment errors.

5.1.1. Toronto Parameters

This section presents the parameter settings and results for the Toronto benchmark dataset.

Firstly, in surface masking step, the threshold for retaining only building surfaces was set at an elevation of $z > 60\,\mathrm{m}$. Then, persistence analysis is performed to extract topologically salient points. This process yielded $39{,}477$ keypoints in the source (photogrammetric) point cloud and $8{,}850$ keypoints in the target (ALS) point cloud. After this step, the average point cloud spacings were 0.142 m for the source cloud and 0.431 m for the target cloud. Finally, voxel downsampling was applied to harmonize the two datasets, resulting in a average point cloud spacing of 1.872 m for both.

Coarse alignment was then conducted using SAC-IA. As discussed in Section 4.2.4, SAC-IA has two key parameters: the normal estimation radius r_{normal} and the FPFH feature radius $r_{feature}$. In practice, r_{normal} is typically chosen to be $2\sim3$ times the average point spacing, while $r_{feature}$ is $3\sim5$ times. A three-round parameter sweep over r_{normal} and $r_{feature}$ was performed to determine the optimal values for this dataset, as shown in Table 5.1.

Table 5.1: 3 Trial Rounds for settting SAC-IA parameters for the Toronto Dataset

Trial	r_n range	footstep	r_f range	footstep	optimal parameters	inliers
1	3.5–5.5	0.1	6.5–7.5	0.1	$r_n = 4.0$, $r_f = 7.6$	4681
2	3.5–5.5	0.1	7.5–8.5	0.1	$r_n = 5.2$, $r_f = 7.9$	4804
3	3.5–5.5	0.1	8.5–9.0	0.1	$r_n = 5.2$, $r_f = 8.8$	4783

After the coarse alignment, the transformation matrix corresponding to the best parameter settings obtained from each set of loops is applied to the whole complete reconstructed point cloud, and then this updated reconstructed point cloud is used to perform the fine alignment with the original complete target(ALS) point cloud, which is realized by ICP. The experimental results show that three different sets of initial poses, after different numbers of iterations, yield convergent fine alignment results.

5.1.2. Westkappelle Benchmark Dataset

This section presents the parameter settings and results for the Westkapelle benchmark dataset.

First, in the surface masking step, only building surfaces above an elevation threshold of $z > 1\,\mathrm{m}$ were retained to remove ground and low-relief features. Then, persistence analysis was performed to extract topologically salient points. This process yielded $6{,}870$ keypoints in the source point cloud and $2{,}360$ keypoints in the target point cloud. After this step, the average point cloud spacing was 0.225 m for the source cloud and 0.315 m for the target cloud. Finally, voxel downsampling was applied to harmonize the two datasets, resulting in a final average point cloud spacing of 2.118 m for both.

Coarse alignment was then conducted using SAC-IA. A three-round parameter sweep over r_{normal} and $r_{feature}$ was performed to determine the optimal values for this dataset, as shown in Table 5.2.

Table 5.2: 3 Trial rounds for settting SAC-IA parameters for the Westkappelle benchmark dataset

Trial	r_n range	Step	r_f range	Optimal parameters	Inliers
1	4–7	0.1	6–8	$r_n = 5.2$, $r_f = 7.0$	1413
2	4–7	0.1	8–10	$r_n = 5.6$, $r_f = 9.1$	1531
3	4–7	0.1	10–11	$r_n = 6.4$, $r_f = 11.0$	1446

5.1.3. Test Dataset Parameters

This section presents the parameter settings and results for the Westkapelle test dataset, which required several adjustments compared to the benchmark cases due to its unique characteristics.

An initial inspection in CloudCompare revealed that a large number of repetitive planar surfaces were concentrated at heights between 1–2 m. Therefore, in the surface masking step, only points with elevations greater than $z > 1\,\mathrm{m}$ were retained to suppress these low-level repetitive structures. Subsequently, persistence analysis was performed to extract topologically salient points using three neighbourhood radii. This process yielded $6{,}618$ keypoints in the source cloud and $3{,}000$ in the target cloud, with corresponding mean spacings of 0.113 m and 0.161 m, respectively. After voxel downsampling, both datasets reached a final average spacing of 6.500 m.

Another notable difference from the benchmark datasets is the criterion used to select optimal parameters. For the Toronto and Westkapelle benchmark case, the optimal settings were chosen based on the highest inlier counts from SAC-IA. However, this approach proved unreliable for the test dataset because the reconstructed photogrammetric point cloud was derived from poor quality historical imagery, resulting in weak and incomplete representation of ground features. In practice, we observed cases where SAC-IA produced high inlier counts but severe misalignments, such as flipped orientations (top/bottom or left/right inversion). To address this, we saved the alignment results for all parameter settings, visualized them in CloudCompare, and selected the setting that produced the closest orientation match to the AHN reference cloud in each loop.

Coarse alignment was then conducted using SAC-IA. However, due to the exceptionally large final average spacing of this dataset, the required parameter ranges became significantly larger than in previous cases. Using the standard step size of 0.1 for parameter sweeps would result in prohibitive computational costs. Moreover, our prior ICP experiments on the Toronto benchmark datasets indicated that, provided the initial pose difference is not excessive, ICP will converge to a similar final alignment regardless of small variations in the coarse alignment parameters.

Based on these considerations, we adopted a modified parameter search strategy:

- First loop: step size set to 1.0 for both r_{normal} and $r_{feature}$ over their full ranges.
- Second loop: same r_{normal} as in the first loop, but reduced step size and a narrower range for $r_{feature}$.
- Third loop: same $r_{feature}$ range as in the first loop, but reduced step size and a narrower range for r_{normal}.

The parameter ranges and optimal results for each loop are listed in Table 5.3.

Table 5.3: 3 Trial Rounds for settting SAC-IA parameters for the Westkappelle Test Dataset

Trial	r_n range	footstep	r_f range	optimal parameters
1	20–32	1	11–18	$r_n = 17$, $r_f = 25$
2	20–26	0.5	11–18	$r_n = 12.5$, $r_f = 22$
3	20–32	0.5	14–18	$r_n = 5.2$, $r_f = 7.9$

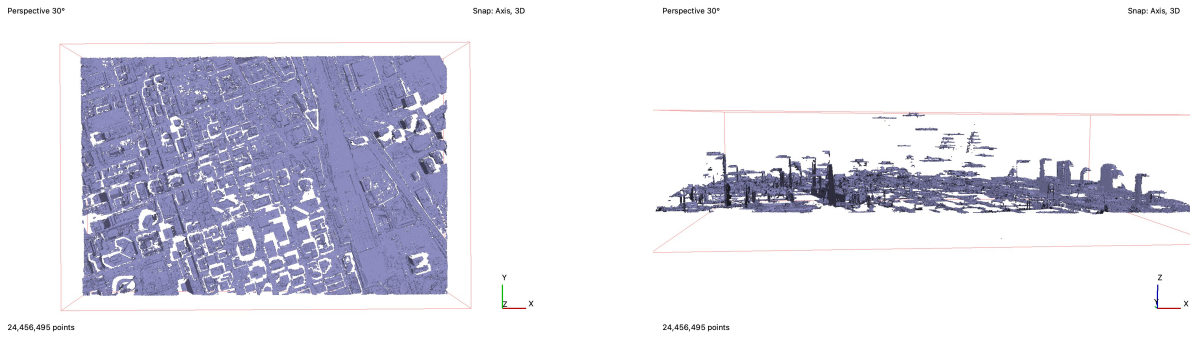
Finally, each of the three optimal coarse alignment parameter settings was used as the initial transformation for fine alignment via ICP. In all cases, ICP converged to the same final rigid-body transformation matrix.

5.2. Result of Benchmark Toronto Downtown Dataset

This section presents the evaluation results for the benchmark Toronto Downtown dataset, which was selected to assess the workflow under differing data acquisition conditions. The analysis covers two complementary perspectives: (i) point cloud registration, evaluated using cloud-to-cloud (C2C) distance metrics, histograms, and spatial visualizations before and after alignment; and (ii) DEM comparison, in which vertical differences between the photogrammetric DEM and ALS DEM are analysed to quantify residual elevation deviations after registration.

5.2.1. Reconstructed Point Cloud

Based on the three .tif images described in 3.1.1, the reconstructed result is shown in Fig. 5.1.



(a) Top View **(b)** Side View

Figure 5.1: Photogrammetric point cloud of Toronto Downtown (overview)

The point cloud features uniform density and rich detail. However, due to the strictly nadir imaging geometry, voids of varying size appear where tall buildings obstruct the vertical line of sight. These gaps are particularly evident at the base of high-rise structures and within narrow street corridors, where the ground and lower facades are entirely occluded from the overhead perspective. As illustrated in Fig. 5.2, such occlusions manifest as white patches adjacent to buildings and in shadowed alleyways. In addition, vertical surfaces such as walls and facades are either poorly reconstructed or completely absent, as they are not directly visible in nadir imagery.

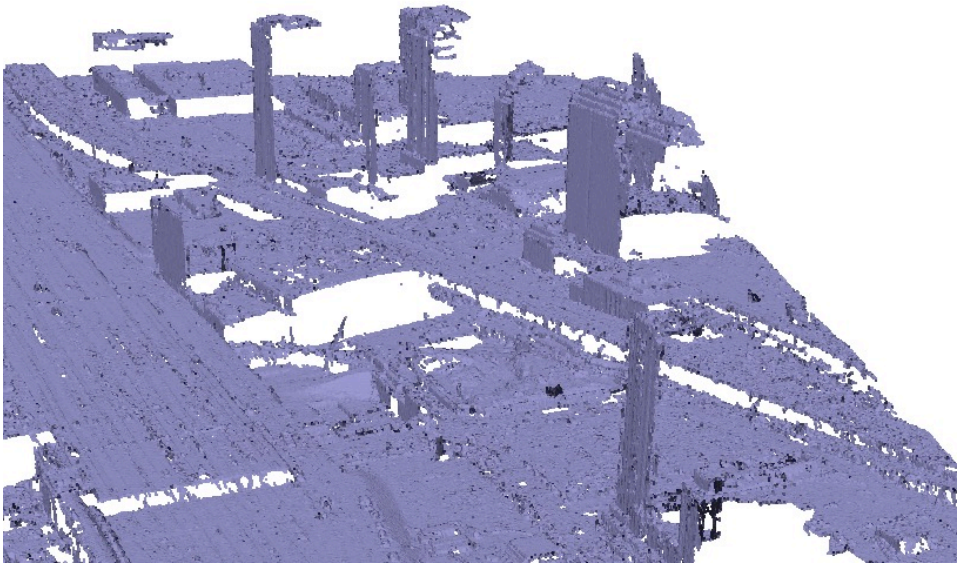


Figure 5.2: Zoom-in view of a section of the photogrammetric Toronto point cloud. The white patches are the result of occlusions caused by large buildings.

Notably, its spatial coverage exceeds that of the blue ALS strip. To reduce computational load and improve subsequent registration accuracy, the point cloud was cropped to approximately match the ALS strip's extent; the cropped result is presented in Fig. 5.3.

Fig. 5.3 overlays the two clipped point clouds (red: photogrammetric; blue: ALS). Most structural outlines align well, but three main discrepancy types are highlighted. In Area A, only the ALS point cloud is available, because the nadir only photogrammetric imaging cannot capture vertical facades or surfaces hidden beneath building eaves. In Area B, shown in the zoom-in figure in Fig. 5.4a, the photogrammetric and ALS building point clouds differ markedly in both height and density. Subsequent inspection revealed that this discrepancy is likely due to the ALS scan having been collected before the building was fully constructed. In Area C, shown in the zoom-in figure in Fig. 5.4b, the photogrammetric and ALS point clouds show structural similarity, but with clear misalignments. These registration errors are also evident in the ground level point clouds.

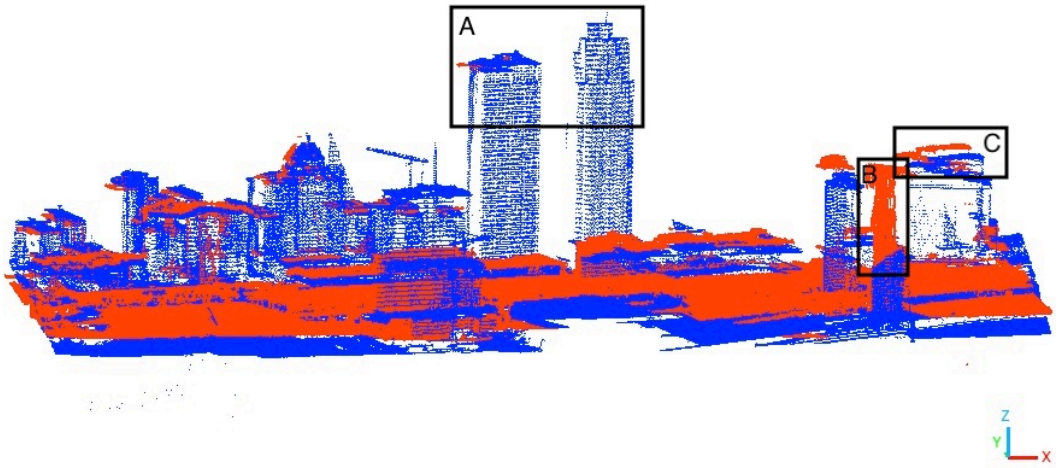
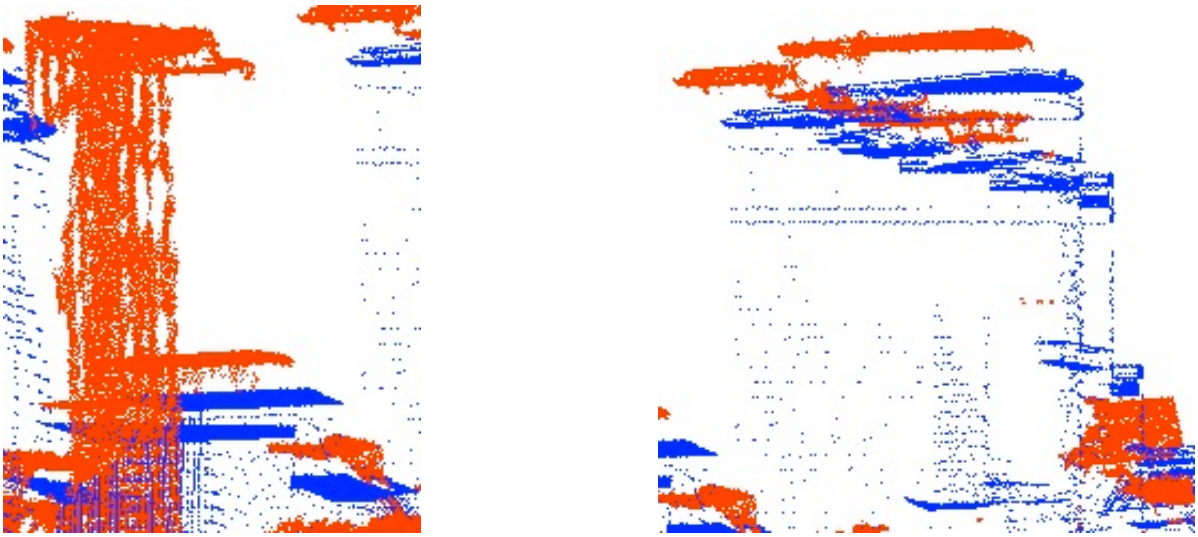


Figure 5.3: Clipped Point Clouds(Blue: ALS point cloud; Red: Photogrammetric point cloud)



(a) Zoom in Visualization of Section B

(b) Zoom in Visualization of Section C

Figure 5.4: Visulization Toronto Point Cloud C2C Difference

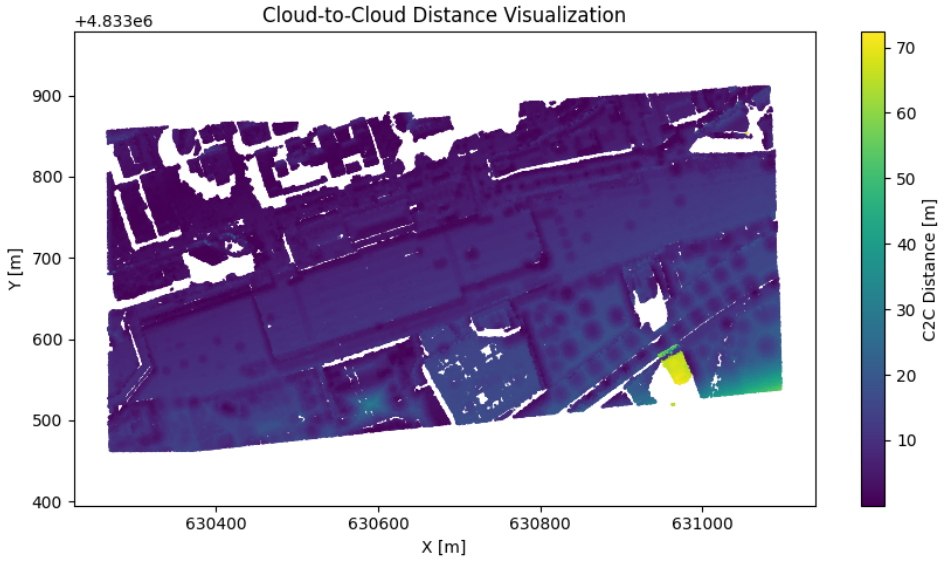
5.2.2. Point Cloud Registration Result and Evaluation

Fig. 5.5 and Fig. 5.6 use the cloud-to-cloud distance (C2C). The C2C procedure takes every point in the photogrammetric cloud as a pin, find the nearest point in the ALS cloud, and measure the straight line distance between the two. Doing this for all points, a distance is obtained.

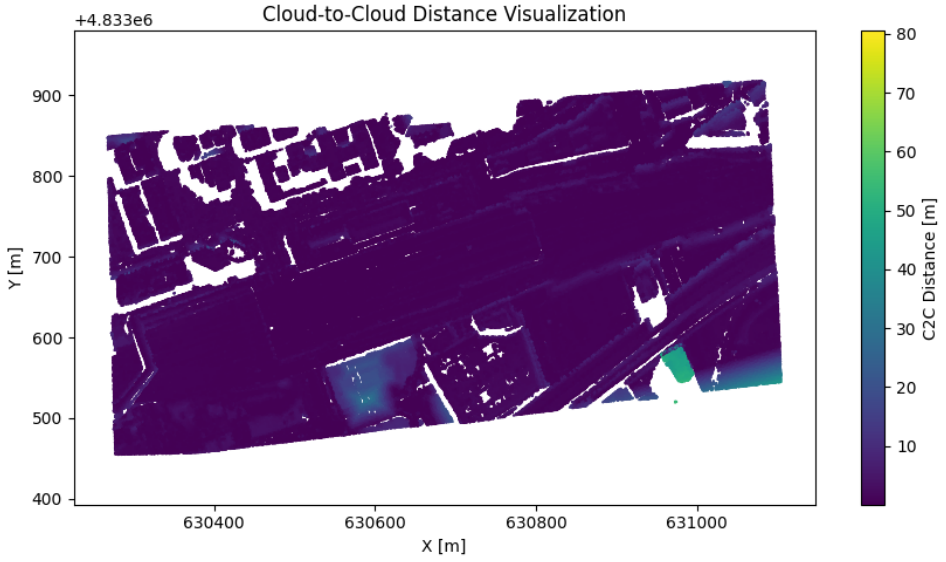
Before registration, roofs and roads are mostly purple and blue. Building facades, ridges, and tower tops show green and yellow patches, with a maximum difference of about 72 m. Some of the largest local errors can be explained by differences in the datasets themselves rather than misregistration. For example, in Fig. 5.5a, the tallest building appears in green and yellow because the ALS cloud only captures roughly half its height, while the photogrammetric cloud represents the full structure. This discrepancy is likely due to acquisition differences, either the ALS survey missed the upper storeys due to occlusion/incidence angle, or the building was incomplete at the time of the ALS flight. Another

prominent yellow patch in the lower right corner of the C2C map results from imperfect cropping of the datasets to a shared footprint, leaving points in one cloud with no counterpart in the other.

After applying the final registration transformation, the C2C colour map (Fig. 5.5b) is almost entirely dark blue, with only a few light-green patches along roof eaves and in narrow shadowed corners, indicating that most corresponding points now differ by only a few metres.



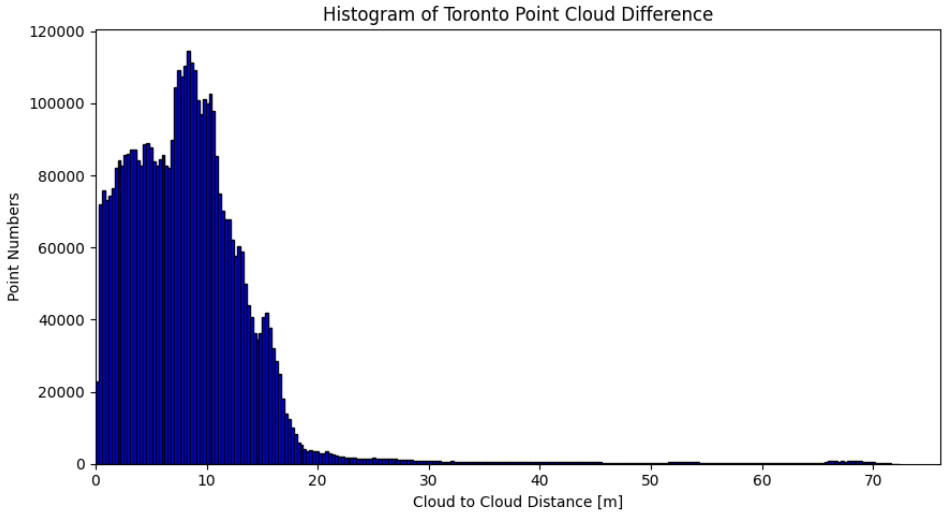
(a) Before registration



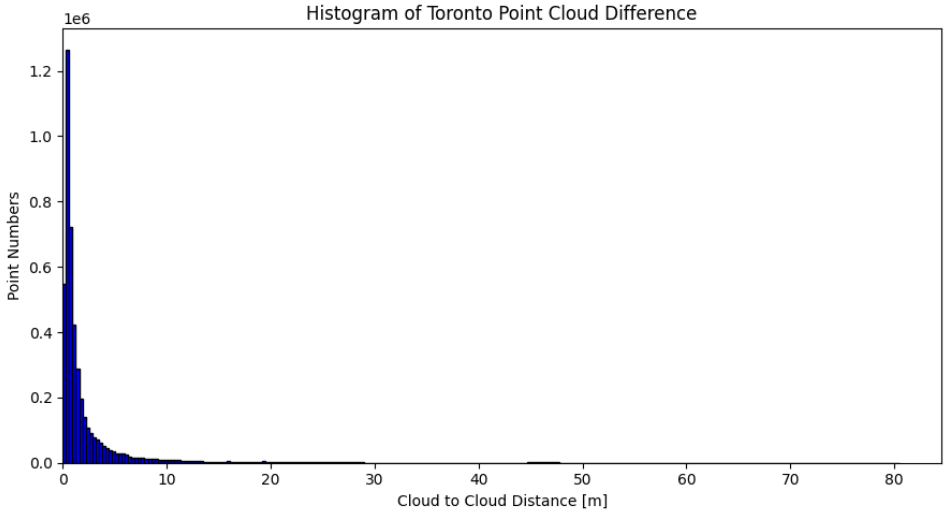
(b) After registration

Figure 5.5: Visulization Toronto Point Cloud C2C Difference

The histogram in Fig. 5.6a peaks in the $5 - 10$ m range and shows a long tail, consistent with the presence of systematic translation and rotation errors. The C2C statistics (Table 5.4) confirm this: the mean distance is 8.518 m, the median is 7.790 m, and the standard deviation is 6.783 m. After registration, the histogram in Fig. 5.6b shows a sharp peak in the $0 - 5$ m range, and a maximum error increased to around 80 m



(a) Before registration



(b) After registration

Figure 5.6: Hisotgram of Toronto Point Cloud C2C Difference

Quantitatively, the mean C2C distance decreased from 8.518 m before registration to 2.368 m afterwards, representing a reduction of approximately 72%. The median showed an even greater relative improvement, dropping from 7.790 m to 0.790 m. The standard deviation decreased from 6.783 m to 4.984 m, indicating a tighter overall distribution after alignment. Despite this improvement, the maximum distance increased slightly (from 72.291 m to 80.340 m), which is attributable to a small number of unmatched or non-overlapping areas (e.g., incomplete building heights or footprint mismatch). These outliers produce a long but low-amplitude tail in the histogram, while the bulk of points concentrate in the $0 - 2$ m bin.

Table 5.4: Statistical results of Cloud-to-Cloud distance (in metres) before and after registration.

Case	Median	Mean	Std. dev.	Min	Max
After registration	0.790	2.368	4.984	0.161	80.339
Before registration	7.790	8.518	6.783	0.151	72.291

5.2.3. DEMs Results and Evaluation

Figures 5.7 and 5.8 present a comparison of the DEMs generated from the photogrammetric point cloud and the ALS point cloud, both rasterized at a spatial resolution of $1\ \mathrm{m}$. Overall, the large-scale terrain morphology and the general building footprint distribution are consistent between the two datasets. However, distinct local discrepancies are observed in the highlighted areas:

- Area A: In the photogrammetric DEM, this building is entirely absent, whereas it is clearly represented in the ALS-derived DEM. This omission in the photogrammetric result is likely due to the building not being successfully reconstructed, a consequence of occlusion from surrounding structures and suboptimal image coverage that prevented the generation of sufficient tie points for dense matching.
- Area B: This structure is visible in the photogrammetric DEM but does not appear in the ALS-derived DEM. A plausible explanation is that the building was incomplete, under construction, or absent at the time of the ALS survey, resulting in its non-representation in the ALS data.
- Area C: The photogrammetric DEM shows distorted and irregular elevation patterns over this rooftop, in contrast to the more uniform and planar representation in the ALS DEM. This deformation is indicative of local reconstruction inaccuracies, potentially caused by poor texture, repetitive patterns, or limited viewing angles that degraded the quality of dense matching.

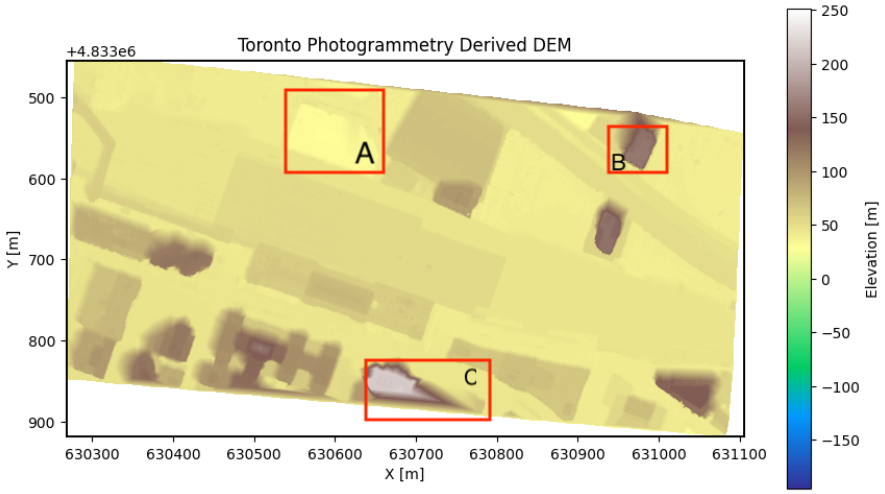


Figure 5.7: The Photogrammetry Derived DEM for Toronto

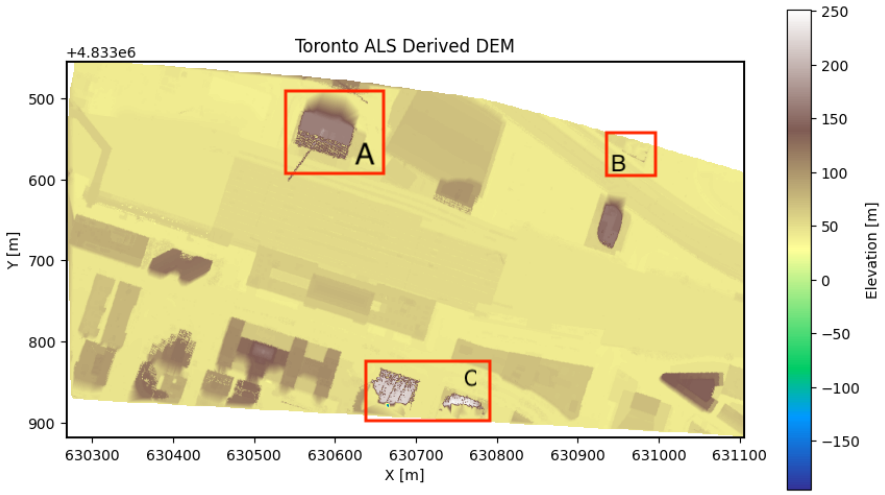


Figure 5.8: The ALS Derived DEM for Toronto

To evaluate the DEMs differences, Fig. 5.9 visualizes the elevation difference between the two point clouds after removing areas with no extreme large negative value, which is due to errors. The color bar represents elevation difference in meters. While most of the differences fall within a relatively narrow range, several prominent outliers remain, especially around building edges.

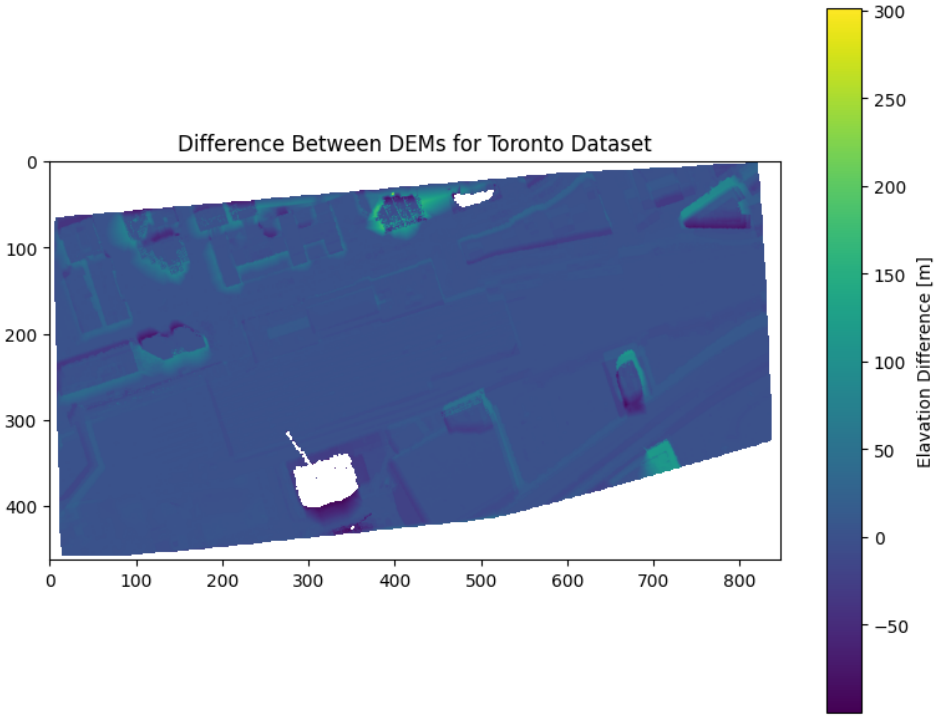


Figure 5.9: The difference between the photogrammetry and ALS derived DEMs for Toronto Dataset

For DEMs difference, the computed statistics show a mean of 1.8586 m, a median of -0.0402 m, and a standard deviation of 17.3580 m. The robust minimum and maximum values, defined as the mean of the lowest and highest 5% of observations, are -33.6416 m and 56.8380 m, respectively. The near zero median suggests that the majority of the elevation differences are symmetrically distributed around zero, indicating good overall vertical alignment between the two datasets.

The relatively high mean compared with the median implies that a small number of large positive deviations are skewing the distribution. Visual inspection confirms that these large residuals are predominantly located at building boundaries, where photogrammetric reconstructions tend to produce blurred, smeared, or over-inflated geometries due to occlusions and poor feature matching. These artifacts result in erroneously high elevation values not present in the ALS-derived DEM. Consequently, although most regions align well, the remaining outliers significantly affect summary metrics like the mean and standard deviation.

Table 5.5: Statistical results of elevation differences between the two DEMs (in metres)

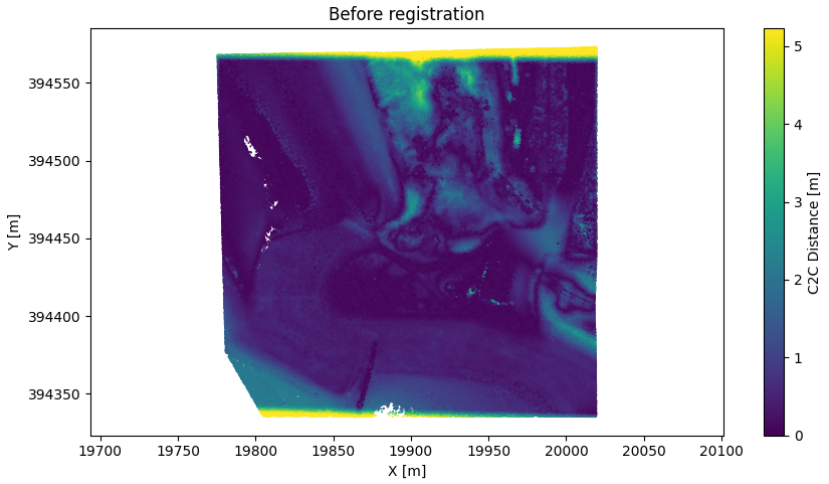
	Mean	Median	Std. dev.	Robust min	Robust max
Value (m)	1.859	-0.040	17.358	-33.642	56.838

5.3. Result of Benchmark Westkapelle, Netherland Dataset

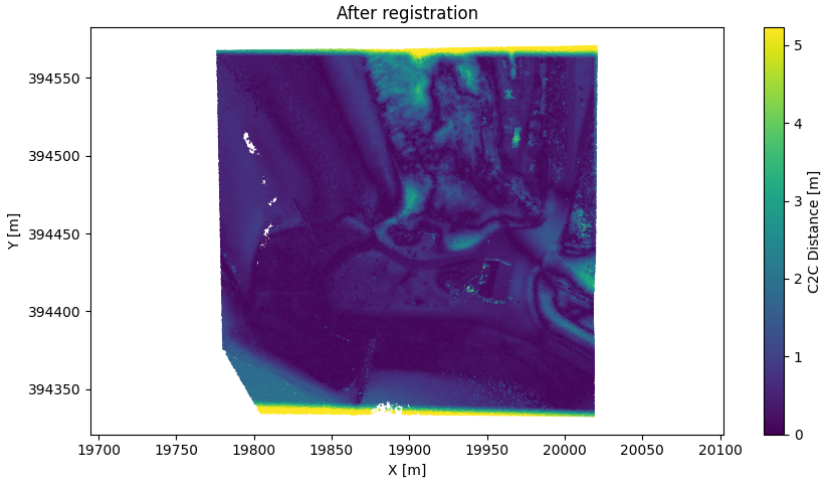
For this benchmark dataset, no DEM generation workflow was carried out in order to simplify the processing. Instead, the data were used solely to validate the point cloud registration workflow.

As shown in Fig. 5.10 AHN2 and AHN4 demonstrate high similarity and overlap, resulting in relatively small errors even before registration. The C2C distance is predominantly blue, while distinct green and

yellow bands clearly visible at the top and bottom edges do not reflect actual terrain changes. Instead, these colored bands result from incomplete spatial overlap between the two datasets. After registration, the overall distribution of differences becomes less uniform: large-scale systematic offsets are removed, and localized higher C2C values emerge, particularly along the coastline, dike structures, and some built-up areas. Although these brighter patches may give the visual impression of greater differences, the accompanying statistics (Table 5.6) show that the mean and standard deviation both decrease after registration, confirming a genuine improvement in overall alignment accuracy. The greener area reflects potential true changes in dynamic coastal zones between acquisition epochs. In particular, the distinct green patches visible on the post-registration map along the dune crest and in certain intertidal areas are consistent with known morphological changes such as dune reshaping, sand nourishment, and shoreline retreat or advance. These areas correspond to parts of the coastline that are naturally dynamic and are periodically altered by both natural processes and human interventions.



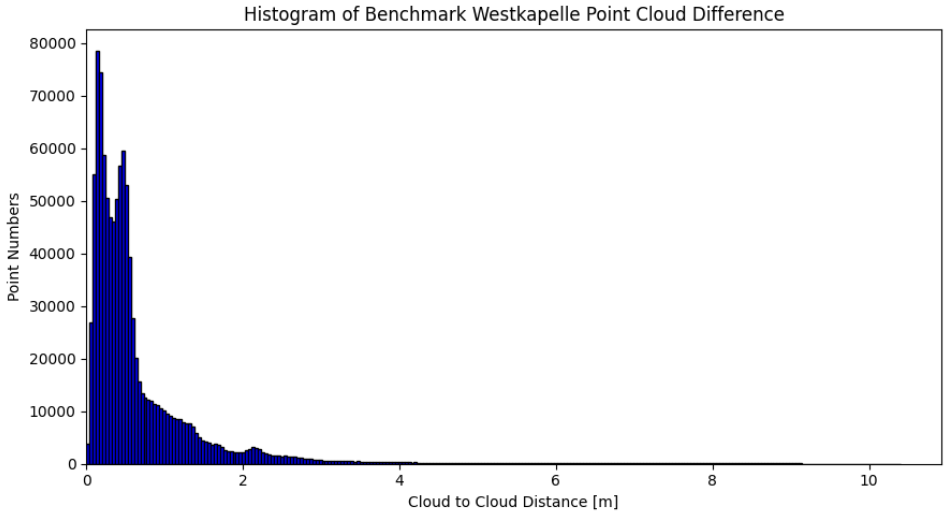
(a) Before registration



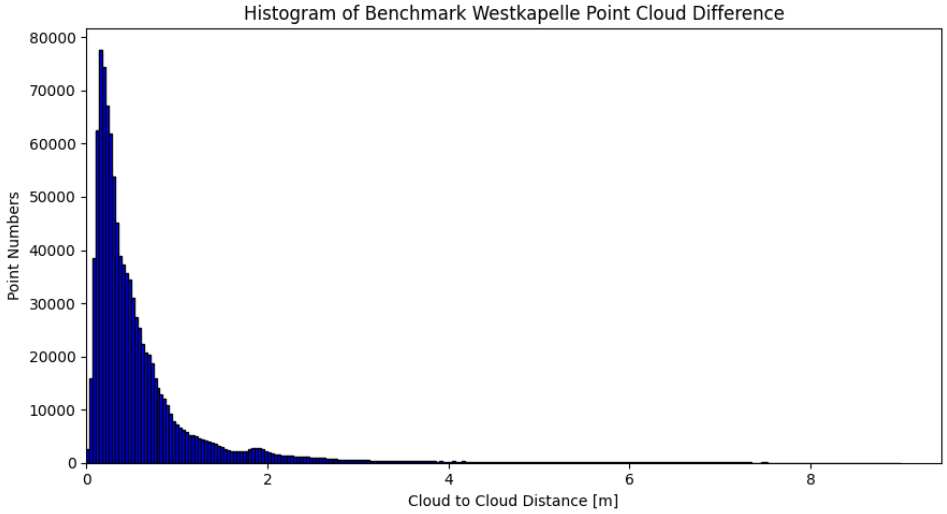
(b) After registration

Figure 5.10: Visulization of Benchmark Westkapelle Point Cloud C2C Difference

The histograms in Fig. 5.11 show that before registration, the C2C distance distribution had two distinct peaks, whereas after registration, the distribution converged into a single dominant peak. The secondary peak in the pre-registration histogram likely corresponds to a systematic offset between the datasets, suggesting that the error was not purely random. Registration effectively eliminated this systematic component, leaving only residual local differences.



(a) Before registration



(b) After registration

Figure 5.11: Histogram of Benchmark Westkapelle Point Cloud C2C Difference

Before registration, the mean C2C distance was 0.702 m, with a standard deviation of 0.934 m, indicating already high accuracy. After registration, as shown in Table.5.7, the mean C2C distance decreased to 0.626 m, and the standard deviation dropped to 0.806 m, reducing the overall error and further improving the accuracy.

Table 5.6: Statistical results of Benchmark Westkapelle C2C distance (in metres) before and after registration.

	Median	**Mean**	**Std. dev.**	**Min**	**Max**
Before registration	0.427	0.702	0.934	0.020	10.383
After registration	0.371	0.626	0.806	0.020	8.978

5.4. Results of Westkapelle, Netherland Dataset

This section presents the evaluation results for the Westkapelle dataset in the Netherlands, which serves as the primary study area in this research. The analysis focuses on two aspects: (i) point cloud registration, evaluated using C2C distance metrics, histograms, and spatial visualizations before and after alignment; and (ii) DEMs comparison, which examines vertical differences between the photogrammetric DEM and the ALS DEM to quantify residual elevation deviations following registration.

5.4.1. Reconstrcuted Point Cloud

In Fig. 5.12, the blue points represent the point cloud generated from photogrammetry, while the RGB-coloured points correspond to the AHN4 reference point cloud. The photogrammetric point cloud shows clear variations in density across different areas. Two representative zones are highlighted with red rectangles in the overview and enlarged on the right for closer inspection.

In the upper highlighted area, located within the built-up zone, the reconstructed points are sparse and unevenly distributed. This is mainly caused by lower image quality and less favourable viewing angles in this part of the dataset. Oblique or partially occluded views of buildings reduce the number of reliable tie points between images, leading to incomplete reconstruction. Consequently, building roofs and façades are only partially captured, and many gaps appear in the point cloud.

In contrast, the lower highlighted area, situated around the curved section of the dike and the adjacent sandy foreshore, contains a much denser and more continuous point cloud. Here, the images have higher clarity and more favourable viewing angles, providing rich texture for feature matching. The open terrain and gentle slopes further facilitate reconstruction, resulting in a more complete and geometrically accurate point cloud in this region. In the visualization, the blue points in this area are mostly obscured by the overlying RGB points from the AHN4 dataset, which explains why they are less visible despite being present.

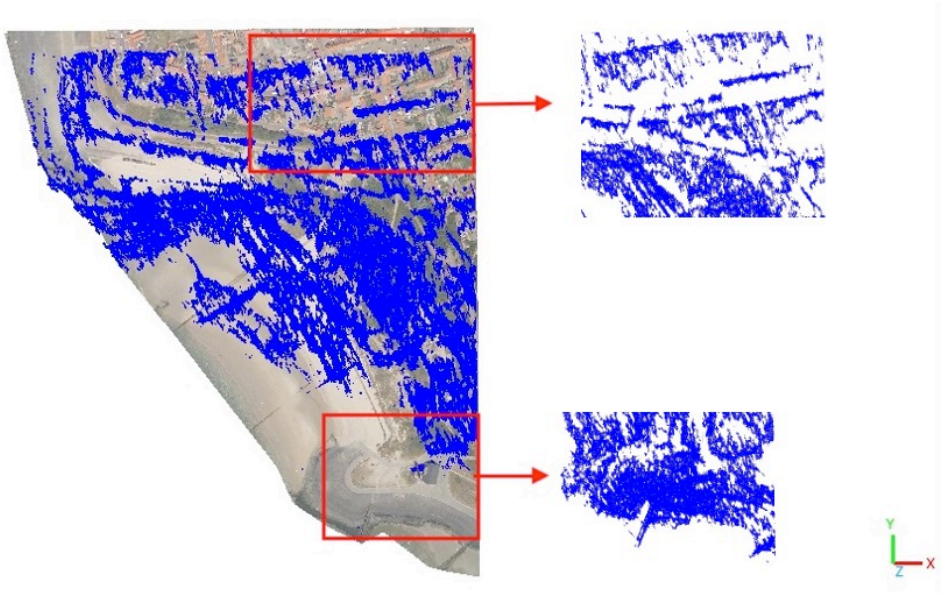
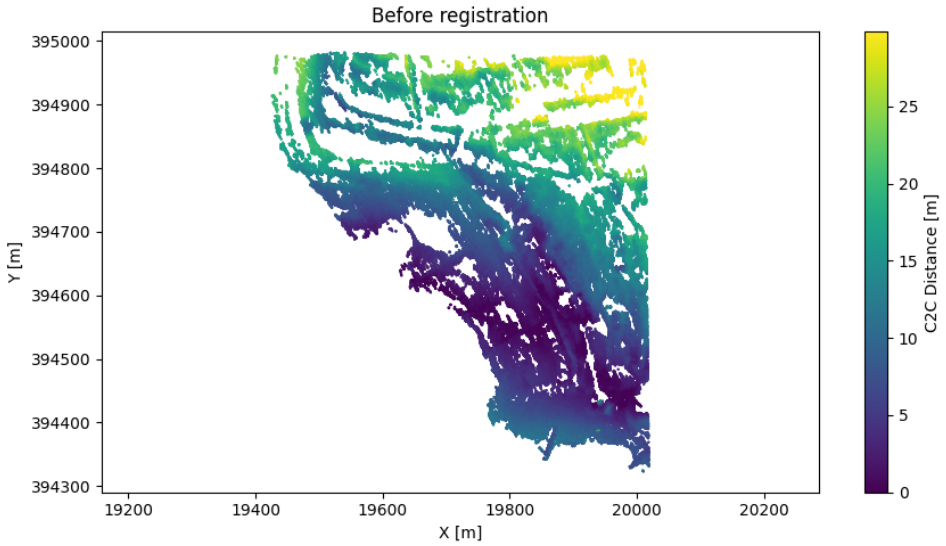


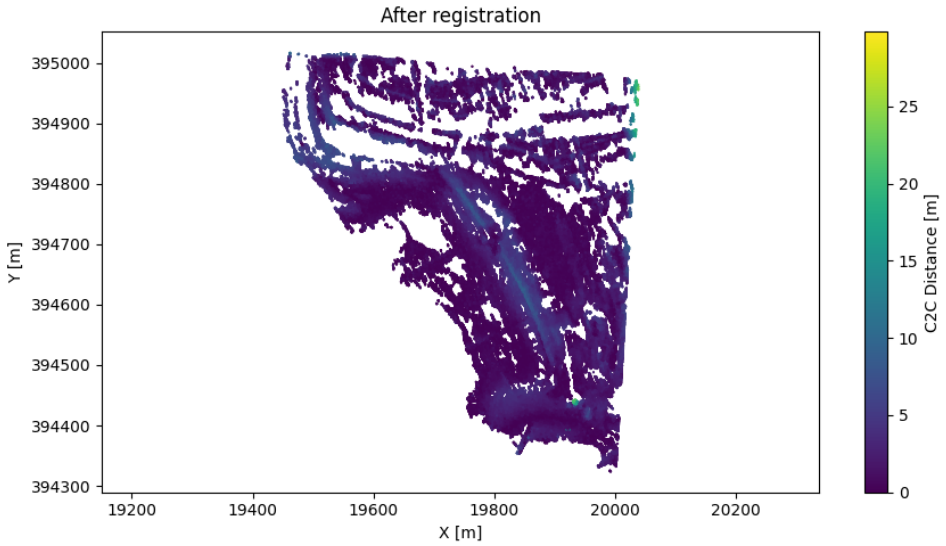
Figure 5.12: The Reconstrucetd Point Cloud and AHN4 of Westkapelle(Blue: Reconstruced point cloud; RGB point cloud: AHN4)

5.4.2. Point Cloud Registration Result and Evaluation

The two images in Fig. 5.13 show the C2C distance distribution for Westkapelle before and after registration. Before registration, large green and yellow area appear across the northern parts and along curved interior belts, indicating severe mismatch of location between two point clouds in those parts. After registration, most points shift to dark purple ($< 5\,\mathrm{m}$), confirming that the systematic offset has been removed. The remaining light blue/green signals are confined to a few coherent strips that coincide with dune/dike slopes and narrow built–up features. These residuals likely reflect a combination of (i) perspective/geometry mismatches on sloped or vertical surfaces and (ii) real coastal morphology changes. Localised bright spots near the dataset boundary are attributed to limited overlap rather than terrain change.



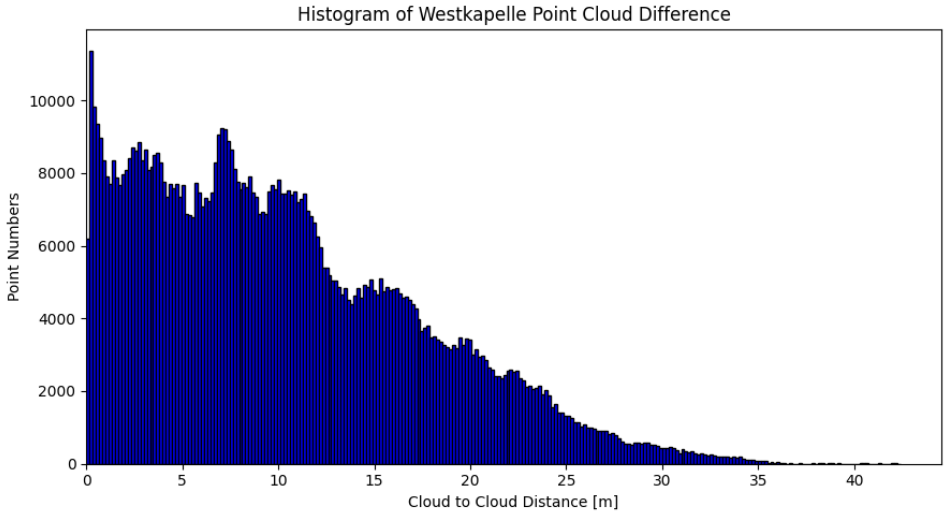
(a) Before registration



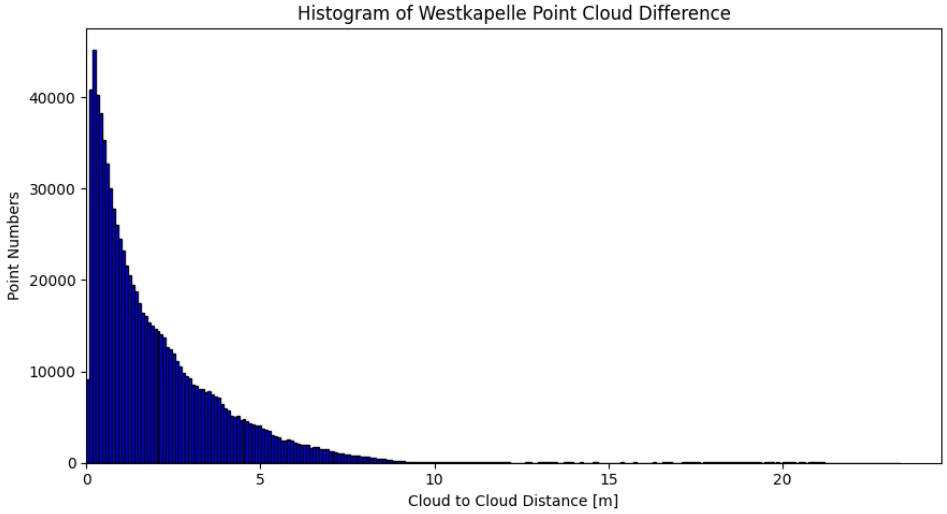
(b) After registration

Figure 5.13: Visualization of Westkapelle Point Cloud C2C Difference

Figure 5.14 shows the corresponding histogram and Table. 5.7 shows the statistic results of C2C distance. Before registration, the C2C distances show a broad and irregular distribution, with a mean of 10.287 m and a standard deviation of 7.236 m. The histogram reveals a peak in the 0-$2\,\mathrm{m}$ bin, followed by several smaller peaks around $4\,\mathrm{m}$, $8\,\mathrm{m}$, and $11\,\mathrm{m}$, and a long tail extending beyond $40\,\mathrm{m}$, indicating large systematic shifts and large errors caused by low overlap, viewing-angle differences, and occlusion in poorly reconstructed areas. After registration, he mean C2C distance drops to $2.025\,\mathrm{m}$ with a standard deviation of $2.037\,\mathrm{m}$, and the histogram (Fig. 5.14b) becomes concentrated: approximately over half of points fall within 0-$1\,\mathrm{m}$, over 90% lie within $6\,\mathrm{m}$, and only a few outliers exceed $10\,\mathrm{m}$. This pronounced narrowing of the distribution confirms that the alignment procedure effectively removed most systematic offsets and reduced local mismatches.



(a) Before registration



(b) After registration

Figure 5.14: Histogram of Westkapelle Point Cloud C2C Difference

Table 5.7: Statistical results of Westkapelle C2C distance (in metres) before and after registration

	Median	Mean	Std. dev.	Min	Max
Before registration	9.035	10.287	7.236	0.088	42.337
After registration	1.420	2.025	2.037	0.050	23.340

5.4.3. DEMs Results and Evaluation

Figures 5.15 and 5.16 compare the DEMs generated from the photogrammetric point cloud and the AHN4 dataset for the Westkapelle area, both rasterized at $1\ \mathrm{m}$. Overall, the photogrammetric DEM is markedly smoother and blurrier than the AHN4 DEM: sharp breaks in slope are broadened into wide bands, small□scale relief is infilled by interpolation, and building and road edges are softened. By contrast, the AHN4 DEM preserves crisp anthropogenic edges and distinct crest–toe transitions along the dike. Local discrepancies in the annotated regions are as follows:

- Area A: The characteristic dike structure visible in AHN4 (a narrow, sharp crest with a clear break in slope) is *not reconstructed* in the photogrammetric DEM. Instead, this section collapses into a diffuse, nearly planar surface with no distinct crest or toe. This failure is consistent with limited

seaward viewpoints, low texture on the grassed slope, and consequent gaps that were bridged by interpolation.

- Area B: The photogrammetric DEM exhibits blurred and incomplete representation of residential blocks and adjacent streets, whereas AHN4 delineates rooftops and road corridors clearly. The degradation likely stems from partial occlusions in the aerial photographs and insufficient parallax for narrow urban features.
- Area C: A spurious high patch exceeding $40\ \mathrm{m}$ appears in the photogrammetric DEM but corresponds to a smaller, correctly shaped structure in AHN4. This artefact arises from reconstruction failure around the tall tower, where occlusions and limited viewing angles produced void filling and inflated elevations.

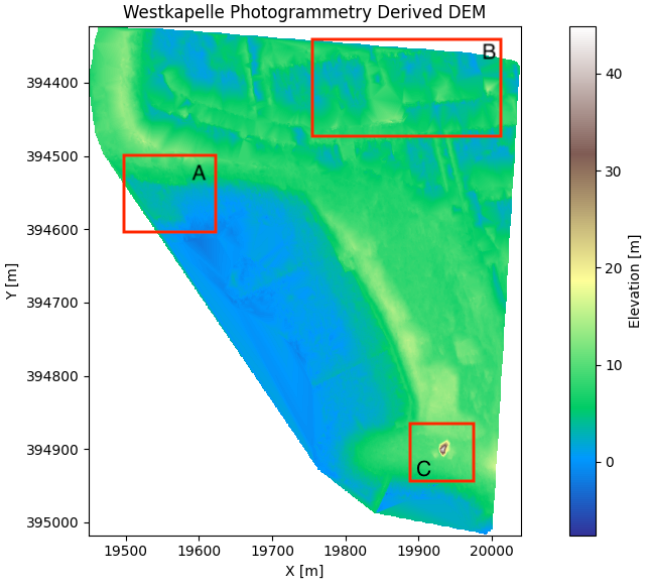


Figure 5.15: The Photogrammetry Derived DEM for Westkapelle

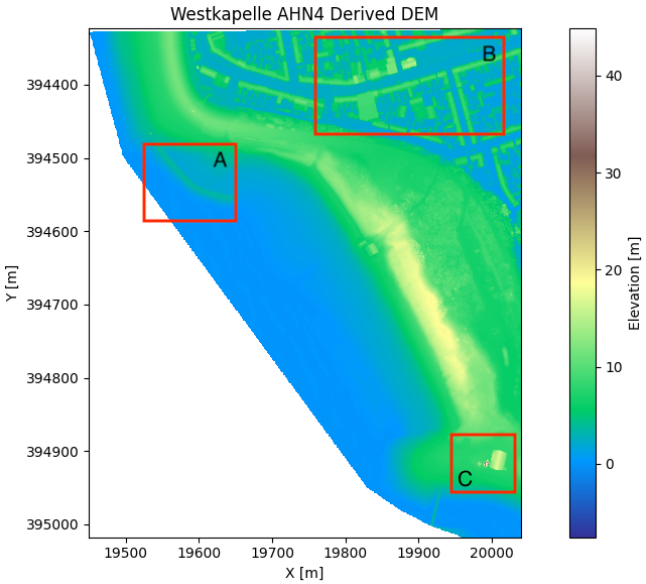


Figure 5.16: The AHN4 Derived DEM for Westkapelle

To quantify the DEMs differences, Fig. 5.17 shows the height difference between the two DEMs. Blue to

purple tones indicate areas where the photogrammetric DEM is higher, yellow tones mark areas where it is lower, and green represents near-zero differences. Distinct yellow patches are visible over building roofs, where differences often exceed $+10\,\mathrm{m}$; these are likely artefacts from stereo-matching errors, occlusions, or residual misalignment. Along the dune crest, however, a continuous yellow band also reaches around $+10\,\mathrm{m}$, which may partly reflect genuine elevation changes since the acquisition dates of the two datasets, in addition to measurement noise. Conversely, the slope base and several shaded surfaces show negative differences of -2 to $-5\,\mathrm{m}$, likely due to underestimation in the photogrammetric model caused by poor image texture and lighting conditions. In the lower part of the scene, two conspicuous anomalies—a purple patch and an adjacent yellow–green patch—correspond to the location of a tall tower. The reconstructed photogrammetric point cloud failed to capture the full tower geometry, and even after registration, the DEM derived from it remains inaccurate in this area. The statistics in Table 5.8 confirm that most differences remain small, with a mean of $-0.441\,\mathrm{m}$, a median of $-0.343\,\mathrm{m}$, and 95% of values within $-6.358\,\mathrm{m}$ to $+6.345\,\mathrm{m}$.

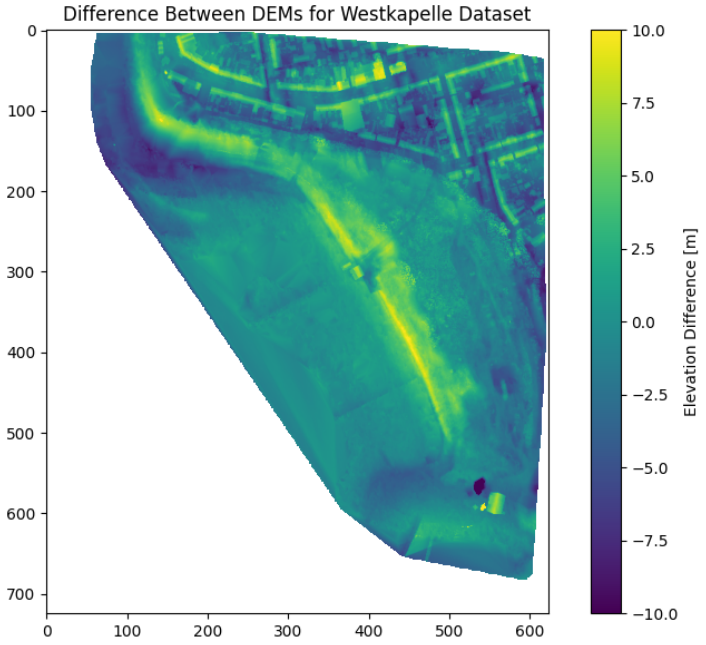


Figure 5.17: The difference between the photogrammetry and ALS derived DEMs for the Westkapelle test dataset

Table 5.8: Statistical results of elevation differences (in metres)

	Mean	**Median**	**Std. dev.**	**Robust min**[1]	**Robust max**[2]
Value (m)	-0.441	-0.343	2.981	-6.358	6.345

5.5. JARKUS

Section 5.4 presents a comparison between the DEM generated from experimental data through photogrammetric reconstruction and the reference ground-truth DEM. However, due to the poor quality of the historical images used, the reconstructed DEM still exhibits significant blurring and uncertainty in several areas. In addition, during the registration process, the two point clouds could not be perfectly aligned, and some residual alignment errors remain. To assess the reliability of the DEM differencing results and to further distinguish whether the observed differences arise from actual elevation changes in coastal dunes or from systematic errors, we introduce the more accurate JARKUS survey data as a reference for validation.

5.5.1. JARKUS Supported Elevation Change

Table 5.9 presents the elevation change statistics from 1990 to 2020 for three JARKUS transects: 16002195, 16002215, and 16002235. It is evident that transect 16002215 exhibits the most significant changes, with the highest mean elevation change, maximum elevation change, and standard deviation. This can be attributed to the specific location of the transect, which lies in the central part of the dune area. Unlike the other two transects, which are partially influenced by relatively stable man-made structures, the elevation changes in this transect mainly result from natural dune accretion or erosion.

Table 5.9: Elevation Change between 1990 and 2020 statistics for JARKUS transects.

Transect	**Mean** (m)	**Min** (m)	**Max** (m)	**SD** (m)
16002195	1.264	-0.430	3.930	1.271
16002215	2.648	0.060	10.240	2.601
16002235	1.264	-0.420	6.250	1.608

Fig. 5.18 shows more detailed elevation change curves for the three transects at different positions. The horizontal axis represents the landward distance from the coastal baseline, while the vertical axis indicates elevation changes between 1990 and 2020. The blue line corresponds to transect 16002195, which shows relatively stable changes with only slight fluctuations of elevation change between 1 and 5 meters. The orange curve represents transect 16002215, which shows a significant elevation increase of up to 10 meters around 85–90 meters. This indicates a concentrated deposition zone, likely formed by wind blown sand transport or scouring events caused by storm surges. The green curve represents transect 16002235, where a notable elevation rise occurs near the front section, around 35 meters. In summary, we suspect that the area where transect 16002215 is located is more strongly affected by aeolian processes. These central dune zones are more susceptible to wind-driven sand accumulation, resulting in more pronounced elevation changes. Unlike the other two transects, this area lacks coastal defense structures or vegetation, which further amplifies the natural variations observed along this profile.

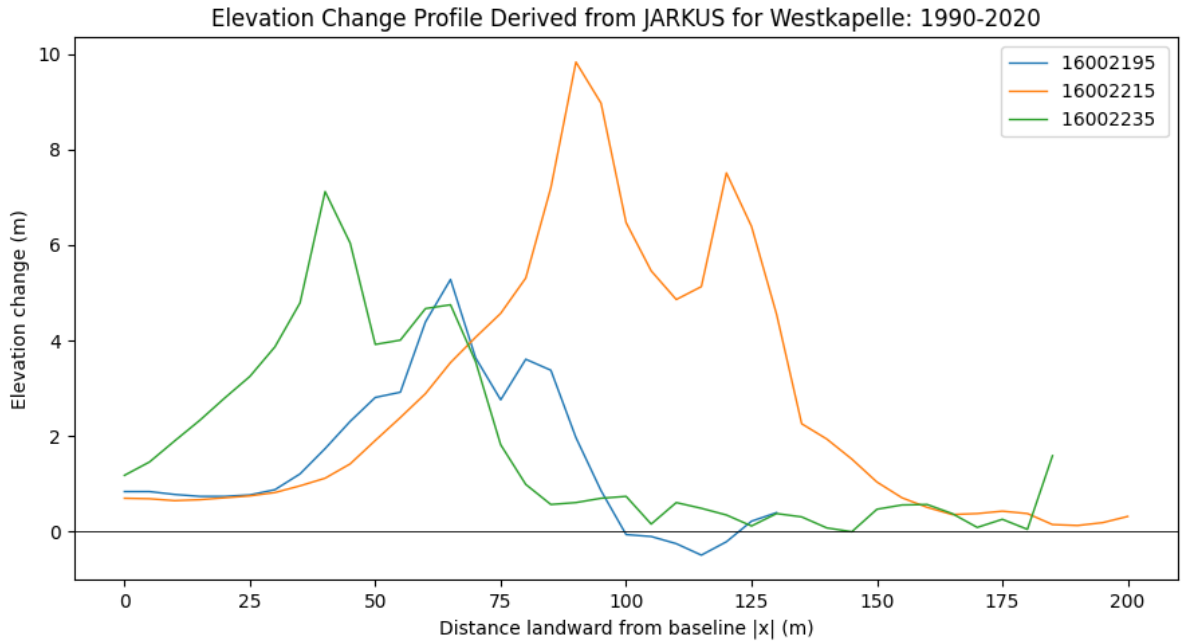


Figure 5.18: Elevation Change Profile Supported by JARKUS

5.5.2. DEMs Supported Elevation Change

Table 5.10 presents the elevation change metrics derived from DEMs (1990–2020). First, unlike the JARKUS data, the DEMs show very large negative elevation changes—especially in transects 16002195 and 16002235—causing the means of these two transects to drop substantially. Transect 16002215 still has the largest mean change, but its maximum value is about 3 m lower than in the JARKUS data, whereas the maximum for 16002195 increases by nearly 4 m. At the same time, the standard deviation of 16002195 rises markedly.

Table 5.10: Elevation Change derived from DEMsbetween 1990 and 2020 statistics for JARKUS transects.

Transect	Mean (m)	Min (m)	Max (m)	SD (m)
16002195	−0.411	−6.510	7.880	3.210
16002215	1.816	−1.126	7.202	1.848
16002235	−0.907	−5.849	5.680	2.400

Fig. 5.19 provides a detailed visualization of elevation change as a function of distance. From this figure—and in conjunction with the table above—we can better explain the observed phenomena. First, DEM differencing for transect 16002195 yields many negative values concentrated beyond approximately 130 m, and shows a very large positive spike around 180–190 m that is highly anomalous compared to the JARKUS data. This likely stems from poor DEM reconstruction quality in this area. Transect 16002195 lies in the upper part of the study area, extending inland for 200 m to the vicinity of built structures. Due to scarce historical imagery and limited acquisition angles, the reconstructed DEM here exhibits severe registration errors and noise, producing spurious elevation fluctuations near buildings; the abnormal drop between 25–60 m also results from this issue. Transect 16002215's overall trend closely overlaps with the JARKUS data—particularly at the sedimentation peaks—indicating that DEM differencing captures the main depositional processes, though the limited reconstruction quality reduces precision. Transect 16002235 matches the JARKUS trend well over the first 100 m, but again shows unrealistic negative values further along, also attributable to insufficient DEM reconstruction quality.

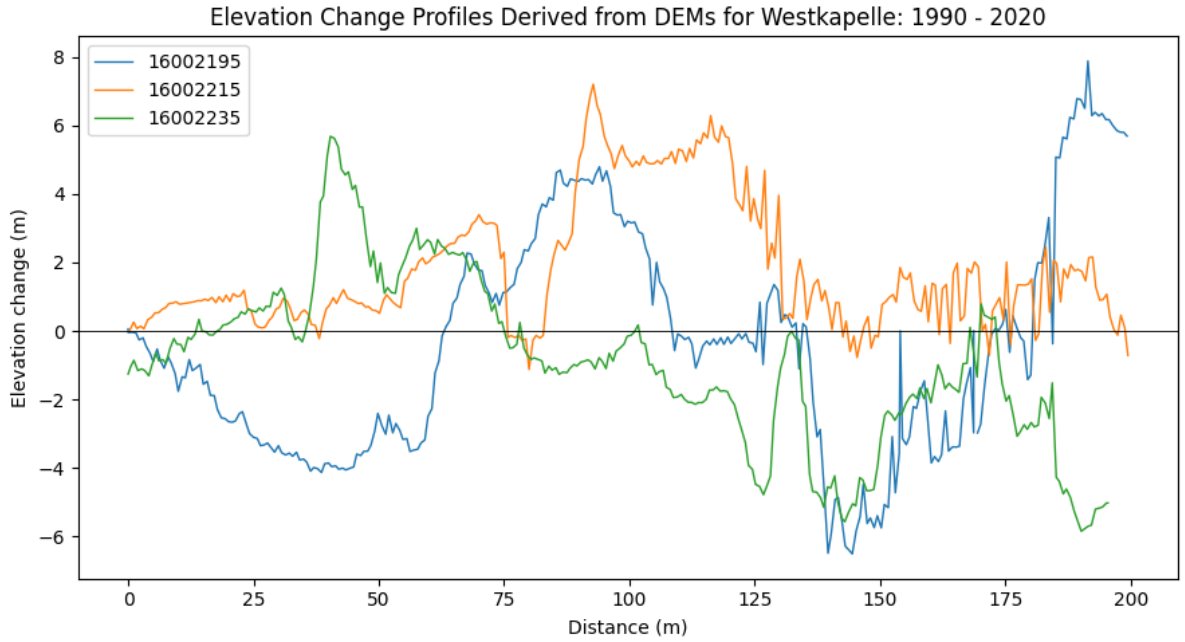


Figure 5.19: Elevation Change Profile Supported by DEMs

In summary, the elevation changes revealed by DEM differencing are able, at a broad scale, to capture

the main erosion and deposition patterns in the dune area—for example, the depositional peak of transect 16002215 and the erosion trends of 16002195/16002235 both generally coincide with the measured JARKUS data—but errors become particularly pronounced in areas with dense buildings or scarce historical imagery, manifesting as locally extreme negative values or unnatural uplift (such as the dramatic fluctuations of 16002195 between 130–200 m).

5.6. Reconstructed Orthophoto and AHN Orthophoto

The orthophoto generated through Metashape is shown in Fig. 5.20. Its overall usability is severely limited due to several significant issues that undermine both its geometric fidelity and radiometric completeness.

- Extensive missing data (NoData holes): The image exhibits large torn-like gaps along the edges, while within the urban core, numerous white voids are scattered throughout. These gaps indicate failure in dense matching or textureless regions during the orthorectification process. Such voids severely compromise the completeness of the coastal landscape and hinder any follow-up spatial comparison or classification.
- Low geometric accuracy and distortions: Key features such as rooftops, ridgelines, and roads appear blurred or spatially deformed. Several buildings exhibit “pulled” or warped outlines, suggesting that the DEM or DSM used in the orthorectification step lacks sufficient vertical precision, or that the raw image resolution was too coarse to retain detailed urban geometry.
- Insufficient ground resolution: Based on the 190-meter scale bar, the ground sampling distance (GSD) of the orthophoto is estimated to be approximately 1 m/pixel. For urban-scale mapping, typical standards recommend a GSD of 10–30 cm. At this resolution, the orthophoto is suitable only as a coarse background layer or for general overview, not for fine-scale measurements or analysis.

These limitations are consistent with the nature of the source material. The historical imagery set comprised only a small number of oblique photographs, taken from limited angles and without true vertical coverage, inherently restricting reconstruction quality in high-relief areas and along vertical façades. Bundle adjustment was likely under-constrained by sparse ground control points (GCPs) and weak camera geometry, amplifying geometric distortions. Furthermore, steep oblique acquisition angles caused perspective deformation in features located farther from the camera, leading to spatial warping and misalignment in overlapping strips.

Despite these constraints, the reconstructed orthophoto retains value for qualitative change detection. As highlighted in Fig. 5.20 and compared with the AHN orthophoto in Fig. 5.21, several buildings remain in the same locations but display noticeable structural modifications. The surrounding vegetation cover has also changed markedly, indicating both natural succession processes and possible anthropogenic alterations over the period between 1990 and 2020.

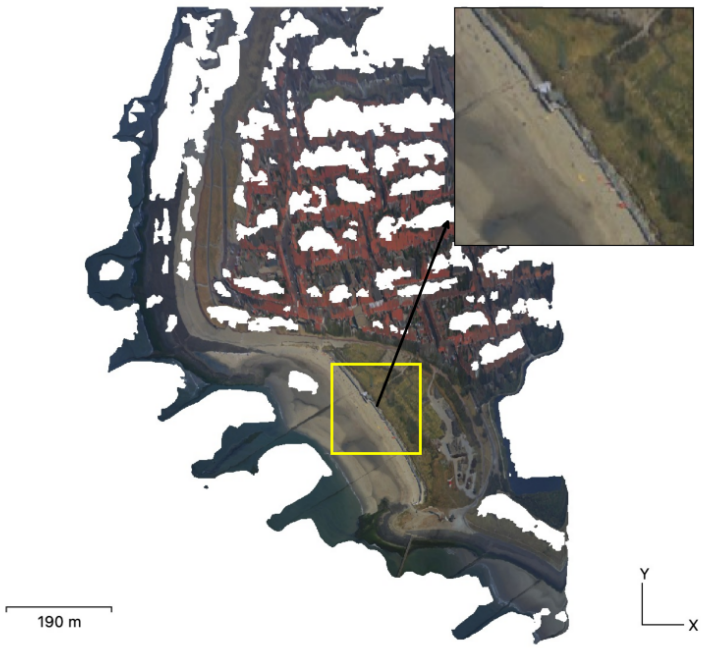


Figure 5.20: Reconstructed Orthophotos of Westkappelle

Figure 5.21: AHN provided Orthophotos of Westkappelle

6

Discussion

This section discusses the results in the context of the main research question and its sub-questions, focusing on how photogrammetry applied to historical images can contribute to the understanding of coastal changes, what changes can be detected, and how the availability and suitability of historical imagery influence the outcomes.

6.1. The Availability and Suitability of Historical Imagery

The photogrammetric workflow combined standard SfM-MVS steps with targeted pre-processing, such as masking dynamic water surfaces and low quality regions. Despite these measures, the overall quality of the reconstruction remained strongly influenced by the properties of the historical imagery. In the Westkapelle dataset, only a small number of oblique photographs were available, providing limited angular coverage. This constraint led to a generally smoother and blurrier reconstrcuted DEM compared to AHN4 derived DEM, with softened edges, interpolated small scale relief, and partial or missing reconstruction of certain structures. In Area A, the characteristic dike crest and sharp break in slope present in AHN4 are absent in the photogrammetric DEM, indicating that image geometry and overlap were the primary limiting factors. Consequently, the Westkapelle case offers limited insight into the availability annd suitability of imagery, making it more suitable to consider the Toronto benchmark as a reference for defining archival image selection criteria.

The Toronto benchmark dataset demonstrates the impact of both image resolution and acquisition geometry. Its exceptionally high scanning resolution produces strong tie point quality and detailed reconstructions, but the near-exclusive use of nadir imagery limits the recovery of vertical structures such as building facades. This geometry is effective for roofs and open ground but results in incomplete models where walls and occluded areas remain poorly defined. A mix of nadir and moderately oblique images would improve intersection geometry, extend coverage to vertical features, and reduce occlusion effects, often providing greater benefit than extreme pixel resolution alone.

In most coastal monitoring applications, a ground sampling distance of 10 cm/pixel is sufficient to delineate dune crests, shoreline positions, and major surface breaks with adequate precision[15]. Image sets should also maintain at least 60% forward and 30% side overlap to ensure stable bundle adjustment and dense matching, especially in low-texture areas such as grassed dikes or sandy beaches[5]. Applying these criteria during archival screening can prioritise datasets most likely to yield reconstructions of acceptable geometric quality. For Westkapelle, additional oblique coverage of the seaward face of the dike would likely have preserved the crest and slope break missing from the photogrammetric DEM and reduced the excessive smoothing observed across the scene.

In many coastal and built environments, moderate camera tilt is essential for capturing vertical and sloping features effectively. Such oblique imagery provides improved visibility of building facades, dike slopes, and dune sides, which are often poorly represented by strictly nadir views[37]. The Westkapelle case in this study illustrates this limitation clearly, as the absence of suitable oblique coverage

contributed to the incomplete reconstruction of the dike crest and seaward slope observed in the photogrammetric DEM.

6.2. Point Cloud Registration Reliability and Accuracy

The three datasets in this study□Benchmark Toronto, Benchmark Westkapelle, and Test Westkapelle, offer complementary perspectives on the reliability of the registration workflow under different acquisition conditions.

For the Benchmark Toronto dataset, the combination of persistence keypoint selection, SAC–IA coarse alignment, and ICP refinement achieved sub-metre resolution between the photogrammetric and ALS point clouds. The mean C2C distance after registration was $0.79\,\mathrm{m}$, with residuals concentrated in a narrow range. The high accuracy is attributed to the dense and evenly distributed features across the urban scene, high image quality, and extensive surface overlap between datasets.

For the Benchmark Westkapelle dataset, the registration shows a mean C2C distance of $0.63\,\mathrm{m}$ after alignment. This performance reflects the dataset's favourable conditions: both source and target point clouds were derived from high-quality aerial acquisitions with consistent scale and minimal occlusion. The dense and homogeneous point coverage across coastal and built-up areas supported stable keypoint matching and robust transformation estimation.

In contrast, the Test Westkapelle dataset presented a far more challenging scenario. The initial mean C2C distance exceeded $10\,\mathrm{m}$ due to mismatches in scale, occlusions, and irregular point densities. After registration, the mean distance decreased to $2.03\,\mathrm{m}$, with 95% of points within $6.0\,\mathrm{m}$. While sufficient for identifying broad erosion and accretion patterns in DEM differencing, this accuracy is inadequate for detecting fine-scale morphological changes at sub-metre resolution. The main limitations stem from sparse and uneven keypoint coverage in low-texture regions such as sand flats, vegetation, and shadowed building façades, which reduced the stability of the coarse alignment and led to larger local residuals after ICP refinement.

Overall, these results highlight three dominant factors controlling registration reliability: the geometric and radiometric quality of both source and target point clouds; the density and spatial distribution of matched keypoints; and the degree of surface overlap and common visibility. Under well-controlled conditions, as in the Benchmark Toronto and Benchmark Westkapelle datasets, high-precision alignment with minimal bias is achievable. In more heterogeneous, occluded datasets such as Test Westkapelle, local deviations should be interpreted cautiously, particularly when analysing small-scale morphological change.

6.3. DEMs Reliability and Accuracy

The DEM difference results from the Benchmark Toronto dataset primarily serve to validate the workflow under near-ideal conditions. After registration, the mean elevation difference between the photogrammetric and ALS DEMs was within the sub-metre range, with residual deviations largely confined to building edges. These localised discrepancies are consistent with the limitations of SfM-derived surfaces in reconstructing fine vertical structures when there is no good angle to reveal the vertical structure, but have negligible impact on the overall elevation pattern.

In the Test Westkapelle dataset, the DEM comparison revealed a complex spatial pattern in which true morphological change and residual registration errors are intertwined. After registration, the mean elevation difference was $-0.44\,\mathrm{m}$, with most deviations falling within $\pm 2\,\mathrm{m}$, but isolated larger differences occurred in low-texture or poorly reconstructed areas. The most prominent and coherent signal is a broad belt of positive elevation change corresponding to the foredune–beach ridge complex, which aligns closely with the accretionary trend observed in the long-term JARKUS profiles. This consistency suggests that the signal captures genuine sediment accumulation over the 30-year interval, driven largely by repeated sand nourishments implemented under Rijkswaterstaat's *Kustlijnzorg* programme, where foreshore or nearshore sediment is gradually transported landward by hydrodynamic and aeolian processes.

At the same time, certain localised deviations, such as thin linear bands along shadowed dune slopes, vegetation edges, or building façades, reflect the residual effects of imperfect registration and mis-

matches in surface representation between the photogrammetric DEM and the LiDAR DEM. These artefacts, together with occlusion-induced reconstruction gaps, contribute to the overall DEM difference pattern and must be considered when interpreting smaller-scale changes. As such, the Test Westkapelle DEM differences should be understood as a combination of genuine elevation change and residual misalignment, with the JARKUS results providing an essential reference for distinguishing between the two.

6.4. Coastal Changes

This section discusses the analysis of coastal changes in the Westkapelle study area by integrating evidence from multiple sources: DEM differences between the 1990 photogrammetry-reconstructed dataset and the 2020 AHN4 reference, long-term shoreline and profile trends derived from the JARKUS coastal monitoring dataset, and visual interpretation of the generated orthophotos. Together, these datasets provide a complementary basis for identifying and interpreting both morphological change patterns and their potential driving processes.

6.4.1. Elevation Change

Differencing the DEM reconstructed from 1990 historical imagery with the 2020 AHN4 DEM highlights a broad swath of net uplift 100–125 m landward of the present shoreline. On the elevation-change map this zone forms a continuous beach-ridge/foredune complex. Incorporating the long-term JARKUS cross-shore profiles (raai 2195, 2215 and 2235) confirms the pattern: although the photogrammetric DEM is less precise, it captures the same upward trend found in the annual survey data.

The accretion visible in this sector is best explained by a combination of drivers, foremost among them the repeated sand nourishments executed under Rijkswaterstaat's Programma Kustlijnzorg. In this programme Rijkswaterstaat periodically deposits North Sea sand on the beach, foreshore, or channel margins wherever erosion or projected relative sea-level rise produces a sand deficit, thereby both maintaining the Basiskustlijn (1990 reference coastline) and allowing the sandy foundation of the coast to 'keep up' with rising sea level [45].

For the three transects that frame our study site—raai 2195, 2215, and 2235—part of the major nourishment interventions between 1990 and 2014 are summarised in Table 6.1.

Table 6.1: Part of major nourishment projects influencing *raai* 2195–2235 (1990–2020)

Year	Nourishment type	*raai* range	Volume (10^3 m^3)
2006	Beach nourishment	2180–3458	1 439
2009	Channel wall	2180–2500	6 254
2011	Beach nourishment	2195–2660	702

After a beach nourishment, the newly placed sand is initially concentrated on the foreshore or in shallow nearshore waters. During winter, strong south-westerly storms push this subtidal sand—including the nourished material—landward and upward, often past the high-tide line. As a result, sediment that was formerly underwater or in the intertidal zone is spread across the upper beach, naturally raising the beach surface.

Westkapelle lies on a pronounced headland at the western tip of Walcheren, where the shoreline bends clockwise by about 25 ° and merges with the deep Oostgat tidal channel. When waves enter this curved section, the alongshore flow slows down; the moving sand therefore "hits the brakes" and is more likely to settle. This makes the 2195–2235 transect a sediment-accumulation point where sand builds up over time.

Once the beach widens, wind during low tide blows the exposed sand landward. Pioneer plants trap and fix this wind-blown sand, gradually forming new embryo dunes and adding another layer of elevation to the beach surface.

6.4.2. Landscape Change

A side-by-side inspection of the 1990 and 2020 orthophotos shows a clear trend: the closer the terrain lies to the Westkappelle sea-dike and the surrounding housing, the sparser the vegetation cover becomes. Small green patches that were still visible in 1990 have largely vanished by 2014. Published studies indicate that this “thinning” of vegetation is most likely driven by three mutually reinforcing processes:

- Human-induced compaction: Heavy foot traffic during the tourist season and repeated passes by beach-cleaning machinery compress the originally loose sand. Bulk density rises, infiltration capacity drops, and both seeds and young roots struggle to penetrate the hardened surface [25, 6].
- Moisture and salt stress: Because compaction reduces drainage, rainwater and wave splash linger in the uppermost few centimetres, producing a “wet-and-salty” top layer over a dry subsurface. Persistent south-westerly salt spray further suppresses the germination of pioneer species [32].
- Lowered groundwater table: Long-term drainage behind the dike lowers the water table beneath the back-dune depression of Het Vroon Westkapelle. The regional water board classifies this zone as a *TOP-verdrogingsgebied* (priority area for ecological drought), meaning shallow-rooting plants have limited access to deep soil moisture [64].

6.4.3. Coastline Change

In the experiments of this study, we attempted to use the 1990 westkappelle for image modeling (SfM-MVS) and differential analysis; however, there are limitations in the quality and spatial-temporal information of the historical images themselves, and ultimately, it was not possible to extract a DEM with sufficiently clear margins (especially at the water-land interface) and an accuracy of 1 m level to meet the requirements of reliable long-time coastline migration trajectories, and thus a reliable coastline migration trajectory can not be derived. The main difficulties that led to this result are as follows:

- Image resolution: The resolution of the early film scans was very poor, and the lens radial aberrations and track offsets were difficult to fully correct without calibrating the original film. These issues led to deformations in the SfM model, undermining the geometric accuracy of the reconstruction.
- Lack of high-precision ground control points (GCPs): No pre-laid GCPs were available in the historical imagery. As a result, only a small number of visually selected stable structures could be used as alternative control points, limiting both the positional accuracy and spatial distribution of georeferencing.
- Registration error: Even after post-registration using rigid-body transformation, the residual planimetric error remains on the meter scale. On a beach with an average slope of 2°, this horizontal error amplifies to tens of centimeters of vertical inaccuracy, causing the “edges” of the differential DEM to deviate from the true shoreline and generating pseudo-elevation steps.

Based on the above limitations, we choose not to rely on this set of differential DEMs to derive accurate shoreline changes in this study.

6.5. Limitations

Although the combined photogrammetric and point cloud registration workflow in this study has shown its potential for reconstructing historical coastal topography, several limitations have constrained the precision and applicability of the results.

First, the accessibility and quality of the historical aerial imagery imposed significant constraints. In the Westkapelle test dataset, the number of usable photographs was limited and most were oblique, with missing metadata such as flight logs, acquisition times, and camera calibration certificates. The limited view geometry produced shallow stereo baselines, which reduced 3D reconstruction quality, particularly around vertical features such as building facades, steep dune slopes, and the high foreshore. In addition, poor resolution, strong radial distortions, and uncorrected lens parameters introduced geometric deformation that could not be fully corrected during bundle adjustment.

Second, the absence of reliable, high-resolution coastline change records derived from the photogrammetric DEMs restricted the scope of quantitative shoreline analysis. Near the water–land interface, the reconstructed surfaces lacked sharp elevation boundaries and exhibited vertical uncertainty of up to several decimetres, making it impossible to extract precise shoreline migration trajectories. For example, in the Westkapelle orthophoto, the transition from wet intertidal sand to dry upper beach was blurred, and in the DEM this zone showed irregular undulations inconsistent with the expected tidal slope.

Third, the two stage registration (SAC-IA + ICP) reduced the C2C distances, but some clear errors remain and they are not uniform across the scene. In the built-up area, most errors appear around buildings. Because many photos are oblique and the image overlap is low, facades that are far from the camera are reconstructed poorly and look blurry. In the DEM difference this shows up as alternating positive/negative bands with large values around building edges and tall structures. Over the dunes, we see narrow belts of high positive elevation. Part of this signal reflects real height gain (accretion), but there are still errors mixed in. Low texture, shadows, and small residual misalignment can increase the values in these belts, so their magnitude should be interpreted with care.

Finally, the orthophoto generated from the historical Westkapelle imagery was of limited practical use. Large NoData gaps, geometric distortions, and a coarse ground sampling distance restricted its application to broad contextual interpretation rather than detailed mapping, classification, or shoreline extraction.

7

Conclusion

This chapter summarizes the main findings of the study and give recommendations for future work. This study focused on reconstruct historical coastal topography from aerial imagery, align it to modern LiDAR point clouds, and evaluate whether the resulting DEM differences can reveal meaningful coastal change. The workflow was tested on three datasets, benchmark Toronto, benchmark Westkapelle, and test Westkapelle. These datasets were chosen to represent a range of acquisition conditions and data qualities, thereby enabling a comprehensive assessment of performance across different scenarios.

7.1. Benchmark Toronto

For this dataset, after using persistence analysis in combination with SAC-IA coarse alignment and ICP refinement, the post registration mean C2C distance achieved a sub-metre level of accuracy ($\approx 0.790\,\mathrm{m}$). Residuals were primarily shown aroud building edges. These locations correspond to known limitations of SfM reconstructions, including weak intersection geometry for vertical facades in predominantly nadir images, insufficient or repetitive texture for reliable feature matching. Such factors reduce the stability of dense matching and increase local geometric noise, but the overall transformation is accurate at the scene scale.

7.2. Benchmark Westkapelle

Prior to registration, the photogrammetric and AHN4 point clouds for this dataset were already in close geometric agreement, with a mean C2C distance of approximately $0.702\,\mathrm{m}$. Applying the registration workflow further improved the alignment, reducing the mean distance to $0.626\,\mathrm{m}$ and narrowing the residual range to within $\pm 0.5\,\mathrm{m}$ across most of the scene. The remaining differences are spatially limited and can be interpreted primarily as genuine surface changes between acquisition dates, such as beach nourishment carried out by the government, rather than artefacts introduced by the registration process.

7.3. Test Westkapelle Dataset

For the Test Westkapelle dataset, the initial mean C2C distance was approximately $10.287\,\mathrm{m}$, reflecting a combination of scale mismatch, occlusions, and uneven point density between the historical and modern datasets. Applying SAC-IA coarse alignment followed by ICP refinement reduced the mean C2C distance to $2.025\,\mathrm{m}$, with 95% of points falling within $\sim 6\,\mathrm{m}$. Larger local residuals were concentrated in architecture areas, where the absence of suitable oblique coverage and low image overlap produced blurry facades far from the camera, resulting in alternating positive and negative deviations of relatively large magnitude, and along dune slopes with low texture and shadows, where weak feature definition reduced the reliability of point matching.

DEM differencing between the photogrammetric surface and the AHN4 reference revealed a broad, coherent belt of positive elevation change along the foredune–beach ridge, consistent with long term

accretion trends recorded by JARKUS profiles. However, narrow high positive bands on steep dune faces and pronounced positive and negative patches around tall structures are more likely attributable to residual registration errors or local reconstruction inaccuracies. The Test Westkapelle DEM difference therefore reflects a combination of true elevation change and residual misalignment, and elevation magnitudes near sharp breaks in slope should be interpreted with caution.

7.4. Answers to the Research Questions

RQ1: Can we observe meaningful coastal change?
Yes, at a spatial scale of several meters to tens of meters. In the Test Westkapelle case, the DEM difference reveals a broad, coherent belt of positive elevation along the foredune–beach ridge, consistent with the long term accretion trend recorded in the JARKUS profiles, and indicating genuine sediment gain. However, sub-meter shoreline migration cannot be detected reliably due to residual misalignment and weak reconstruction near the waterline.

RQ2: Under what conditions can historical images be used in similar projects?
Historical aerial imagery can be highly valuable for coastal change studies when the acquisition provides favourable image geometry, including high forward and side overlap, multiple view angles, and a reasonable stereo baseline, and when basic camera information or self calibration is available. The suitability also increases when dynamic water surfaces and strong shadows are either absent or masked during processing. Conversely, the potential of historical imagery is limited when photographs are strongly oblique with low overlap, metadata are missing, large portions of the scene have low texture, or vertical facades dominate with minimal coverage.

RQ3: What is the point cloud registration accuracy?
Across the three datasets, registration accuracy varied according to acquisition quality and scene characteristics. For Benchmark Toronto, the mean C2C distance after registration was approximately $0.790\,\mathrm{m}$, with residuals concentrated along building edges. Benchmark Westkapelle achieved an even lower post registration mean of $0.626\,\mathrm{m}$, reflecting the favourable starting alignment and high-quality source data. In the more challenging Test Westkapelle case, the mean C2C distance improved from $10.287\,\mathrm{m}$ before registration to $2.025\,\mathrm{m}$ afterwards, with 95% of points lying within about $6.0\,\mathrm{m}$. However, meter scale local residuals persisted, especially around buildings, where oblique views and low overlap produced alternating positive and negative deviations, and along shadowed or low-texture dune slopes where point matching was less reliable.

RQ4: What is the DEM differencing accuracy from this workflow?
In the Benchmark Toronto dataset, small residuals remained a, primarily at building edges, but the overall elevation pattern was preserved with sub-metre accuracy at the scene scale. For Test Westkapelle, DEM differencing between the photogrammetric surface and the AHN4 reference shows a mean elevation change of $-0.440\,\mathrm{m}$, a median of $-0.343\,\mathrm{m}$, and a standard deviation of $2.981\,\mathrm{m}$, with robust extremes ranging from -6.358 to $6.345\,\mathrm{m}$. While the broad belt of positive change along the foredune–beach ridge corresponds to genuine accretion, narrow high positive bands on steep dune faces and strong alternating patches near tall structures likely reflect residual registration or reconstruction errors.

7.5. Future Work

First, the quality of historical image reconstructions could be enhanced by optimising historical image selection criteria, particularly prioritising image sets with higher forward and side overlap, moderate oblique coverage, and sufficient stereo baseline. In the Westkapelle case, the absence of suitable oblique imagery limited the recovery of vertical features such as the seaward dike face. Targeted inclusion of such views in future acquisitions would likely improve completeness and reduce interpolation artefacts.

Second, the registration process could benefit from incorporating additional feature matching strategies to improve robustness in low texture regions. This may include testing advanced keypoint descriptors, integrating intensity or multispectral data from LiDAR, or applying machine learning feature detection methods to enhance correspondence quality.

Third, DEM differencing accuracy could be improved by using stable areas to model and remove sys-

tematic errors. For example, in the Test Westkapelle dataset, some elevation differences, such as foredune accretion, represent real change, while others, especially on steep slopes or near buildings, are caused by registration or reconstruction errors. By identifying features that have not changed between epochs and using them to generate a bias correction surface, these systematic errors can be estimated and subtracted from the DEM difference. This would make the remaining elevation changes more representative of true morphological change rather than artefacts.

Finally, extending the approach to a broader range of coastal area and time periods would allow evaluation of its generalisability and long term monitoring potential. This includes applying the workflow to multi decadal image archives, combining with other remote sensing data such as SAR or bathymetric LiDAR, and integrating with morphodynamic models to interpret observed changes within a process based framework.

References

[1] Abbas Abedinia, Michael Hahnb, and Farhad Samadzadegana. “An investigation into the registration of LIDAR intensity data and aerial images using the SIFT approach”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 2 (2008), p. 6.

[2] Actueel Hoogtebestand Nederland. *AHN 4 — Planning and Data Overview*. Accessed: 11 August 2025. 2025. url: https://www.ahn.nl/dataroom.

[3] Actueel Hoogtebestand Nederland (AHN). *Over het programma AHN*. Programme overview (in Dutch). AHN Consortium. url: https://www.ahn.nl/over-het-programma-ahn.

[4] Sheng Ao et al. “SpinNet: Learning a general surface descriptor for 3D point cloud registration”. In: *CVPR* (2021).

[5] Alba Nely Arévalo-Verjel et al. “Assessment of Photogrammetric Performance Test on Large Areas by Using a Rolling Shutter Camera Equipped in a Multi-Rotor UAV”. In: *Applied Sciences* 15.9 (2025), p. 5035.

[6] M. Ballantyne and C. M. Pickering. “The impacts of trail infrastructure on vegetation and soils: Current literature and future directions”. In: *Journal of Environmental Management* 164 (2015), pp. 53–64. doi: 10.1016/j.jenvman.2015.09.032. url: https://doi.org/10.1016/j.jenvman.2015.09.032.

[7] DM Bappy and MD HAMIDUR RAHMAN. *A Study in 3D Structure Detection Implementing Forward Camera Motion*. 2012.

[8] L Barazzetti et al. “True-orthophoto generation from UAV images: Implementation of a combined photogrammetric and computer vision approach”. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2 (2014), pp. 57–63.

[9] Paul J Besl and Neil D McKay. “A method for registration of 3-D shapes”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.2 (1992), pp. 239–256.

[10] Filippo Brandolini and Giacomo Patrucco. “Structure-from-motion (SFM) photogrammetry as a non-invasive methodology to digitalize historical documents: A highly flexible and low-cost approach?” In: *Heritage* 2.3 (2019), pp. 2124–2136.

[11] Elisa Casella et al. “Accuracy of sand beach topography surveying by drones and photogrammetry”. In: *Geo-Marine Letters* 40 (2020), pp. 255–268.

[12] Jenifer E Dugan et al. “Give beach ecosystems their day in the sun”. In: *Science* 329.5996 (2010), pp. 1146–1146.

[13] James I Ebert. “Photogrammetry, photointerpretation, and digital imaging and mapping in environmental forensics”. In: *Introduction to environmental forensics*. Elsevier, 2015, pp. 39–64.

[14] Massimo Fabris. “Coastline evolution of the Po River Delta (Italy) by archival multi-temporal digital photogrammetry”. In: *Geomatics, Natural Hazards and Risk* 10.1 (2019), pp. 1007–1027.

[15] José A Gonçalves and Renato Henriques. “UAV photogrammetry for topographic monitoring of coastal areas”. In: *ISPRS journal of Photogrammetry and Remote Sensing* 104 (2015), pp. 101–111.

[16] Ayman F Habib, Eui-Myoung Kim, and Chang-Jae Kim. “New methodologies for true orthophoto generation”. In: *Photogrammetric Engineering & Remote Sensing* 73.1 (2007), pp. 25–36.

[17] Richard Hartley. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[18] Gerrit Hendriksen. *Dataset documentation Kustlijn*. Deltares OpenEarth public wiki. Based on data received from Rijkswaterstaat. Contains Basiskustlijn (BKL), Toetsen KustLijn (TKL), Momentane Kustlijn (MKL) from 1990 until today. Aug. 2024. url: https://publicwiki.deltares.nl/spaces/OET/pages/308413068/Dataset+documentation+Kustlijn.

[19] Pavel Kabat et al. “Dutch coasts in transition”. In: *Nature Geoscience* 2.7 (2009), pp. 450–452.

[20] Dimitrios Kapnias, Pavel Milenov, Simon Kay, et al. “Guidelines for best practice and quality checking of ortho imagery”. In: *Joint Research Centre* 3 (2008).

[21] C Kasse and JW Griede. “Holocene development of the Dutch coastal plain: natural processes and human activity”. In: *Excursion Guide for 8th International Conference on Fluvial Sedimentology*. Delft, the Netherlands, 2005.

[22] Ahmed El Khazari et al. “Deep global features for point cloud alignment”. In: *Sensors* 20.14 (2020), p. 4032.

[23] Alexander Kirillov et al. “Segment anything”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2023, pp. 4015–4026.

[24] JA Klijn. “The younger dunes in the Netherlands| chronology and causation in Dunes of the European Coasts. Geomorphology-Hydrology-Soils”. In: *Catena. Supplement (Giessen)* 18 (1990), pp. 89–100.

[25] P. Kutiel, H. Zhevelev, and R. Harrison. “The effect of recreational impacts on soil and vegetation of stabilised coastal dunes in the Sharon Park, Israel”. In: *Ocean & Coastal Management* 42 (1999), pp. 1041–1060. doi: 10.1016/S0964-5691(99)00060-5. url: https://doi.org/10.1016/S0964-5691(99)00060-5.

[26] Subbu Esakkipandian Lakshmi and Kiran Yarrakula. “Review and critical analysis on digital elevation models”. In: *G eofizika* 35.2 (2018), pp. 129–157.

[27] Lies Leewis et al. “Does beach nourishment have long-term effects on intertidal macroinvertebrate species abundance?” In: *Estuarine, Coastal and Shelf Science* 113 (2012), pp. 172–181.

[28] Wei Li, Hongtai Cheng, and Xiaohua Zhang. “Efficient 3D object recognition from cluttered point cloud”. In: *Sensors* 21.17 (2021), p. 5850.

[29] Tony Lindeberg. “Scale Invariant Feature Transform (SIFT)”. In: *Scholarpedia* 7.5 (2012). Scholarpedia article, p. 10491. doi: 10.4249/scholarpedia.10491. url: http://www.scholarpedia.org/article/Scale_Invariant_Feature_Transform.

[30] Charles Loop and Zhengyou Zhang. “Computing rectifying homographies for stereo vision”. In: *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*. Vol. 1. IEEE. 1999, pp. 125–131.

[31] Qinjie Lu et al. “HRegNet: A hierarchical network for large-scale outdoor point cloud registration”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 175 (2021), pp. 448–460.

[32] M. L. Martínez and N. P. Psuty. “Coastal dunes: ecology and conservation”. In: *Ecological Studies* 171 (2004). Section on salt-spray stress, pp. 347–362. doi: 10.1007/978-3-540-74002-5_16.

[33] Maciej A Mazurowski et al. “Segment anything model for medical image analysis: an experimental study”. In: *Medical Image Analysis* 89 (2023), p. 102918.

[34] Iaroslav Melekhov et al. “Relative camera pose estimation using convolutional neural networks”. In: *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer. 2017, pp. 675–687.

[35] Edward M Mikhail, James S Bethel, and J Chris McGlone. *Introduction to modern photogrammetry*. John Wiley & Sons, 2001.

[36] Yoshikatsu Nakajima and Hideo Saito. “Robust camera pose estimation by viewpoint classification using deep learning”. In: *Computational Visual Media* 3.2 (2017), pp. 189–198.

[37] Paul Ryan Nesbit and Christopher H Hugenholtz. “Enhancing UAV–SFM 3D model accuracy in high-relief landscapes by incorporating oblique images”. In: *Remote Sensing* 11.3 (2019), p. 239.

[38] RJ Nicholls et al. “Planning for long-term coastal change: Experiences from England and Wales”. In: *Ocean Engineering* 71 (2013), pp. 3–16.

[39] Jin-Si R Over et al. *Processing coastal imagery with Agisoft Metashape Professional Edition, version 1.6—Structure from motion workflow documentation*. Tech. rep. US Geological Survey, 2021.

[40] Yeping Peng et al. “Point cloud registration based on fast point feature histogram descriptors for 3d reconstruction of trees”. In: *Remote Sensing* 15.15 (2023), p. 3775.

[41] Point Cloud Library (PCL) developers. *Fast Point Feature Histograms (FPFH) descriptors — PCL tutorial*. url: https://pcl.readthedocs.io/projects/tutorials/en/master/fpfh_estimation.html.

[42] Point Cloud Library (PCL) developers. *Point Feature Histograms (PFH) descriptors — PCL tutorial*. url: https://pcl.readthedocs.io/projects/tutorials/en/master/pfh_estimation.html.

[43] Laurent Polidori and Mhamad El Hage. “Digital elevation model quality assessment methods: A critical review”. In: *Remote sensing* 12.21 (2020), p. 3522.

[44] Sam Provoost, M Laurence M Jones, and Sally E Edmondson. “Changes in landscape and vegetation of coastal dunes in northwest Europe: a review”. In: *Journal of Coastal Conservation* 15.1 (2011), pp. 207–226.

[45] Rijkswaterstaat. *Factsheet: Uitgangspunten suppletieprogramma Kustlijnzorg*. Factsheet. The Hague: Rijkswaterstaat, Ministry of Infrastructure and Water Management, Oct. 2024. url: https://www.noordzeeloket.nl/publish/pages/238106/bijlage-3-factsheet-uitgangspunten_2024-v2.pdf.

[46] Rijkswaterstaat. *Jarkusraaien (2024)*. https://data.overheid.nl/en/dataset/46994-jarkusraaien. Dataset, CC-0 licence, updated 20 June 2025. 2025.

[47] Franz Rottensteiner et al. “The ISPRS benchmark on urban object classification and 3D building reconstruction”. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences; I-3* 1.1 (2012), pp. 293–298.

[48] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. “Fast point feature histograms (FPFH) for 3D registration”. In: *2009 IEEE international conference on robotics and automation*. IEEE. 2009, pp. 3212–3217.

[49] Radu Bogdan Rusu et al. “Persistent point feature histograms for 3D point clouds”. In: *Intelligent Autonomous Systems 10*. IOS Press, 2008, pp. 119–128.

[50] Olimpiya Saha and Prithviraj Dasgupta. “Experience learning from basic patterns for efficient robot navigation in indoor environments”. In: *Journal of Intelligent & Robotic Systems* 92.3 (2018), pp. 545–564.

[51] A. H. Jr. Sallenger et al. “Evaluation of airborne topographic lidar for quantifying beach changes”. In: *Journal of Coastal Research* (2003), pp. 125–133.

[52] Enoc Sanz-Ablanedo et al. “Accuracy of unmanned aerial vehicle (UAV) and SfM photogrammetry survey as a function of the number and location of ground control points used”. In: *Remote Sensing* 10.10 (2018), p. 1606.

[53] Yoli Shavit and Ron Ferens. “Introduction to camera pose estimation with deep learning”. In: *arXiv preprint arXiv:1907.05272* (2019).

[54] Erik Smistad et al. “Medical image segmentation on GPUs–A comprehensive review”. In: *Medical image analysis* 20.1 (2015), pp. 1–18.

[55] Mark William Smith, Jonathan L Carrivick, and Duncan J Quincey. “Structure from motion photogrammetry in physical geography”. In: *Progress in physical geography* 40.2 (2016), pp. 247–275.

[56] Weiwei Sun et al. “Coastline extraction using remote sensing: A review”. In: *GIScience & Remote Sensing* 60.1 (2023), p. 2243671.

[57] KN Tahar. “An evaluation on different number of ground control points in unmanned aerial vehicle photogrammetric block”. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 40 (2013), pp. 93–98.

[58] Bill Triggs et al. “Bundle adjustment—a modern synthesis”. In: *International workshop on vision algorithms*. Springer. 1999, pp. 298–372.

[59] Christa Van IJzendoorn. *JAT – Jarkus Analysis Toolbox, Version 2*. 4TU.ResearchData. 2021. doi: 10.4121/14199542.v2. url: https://doi.org/10.4121/14199542.v2.

[60] Mark VanKoningsveld et al. “Living with sea-level rise and climate change: a case study of the Netherlands”. In: *Journal of coastal research* 24.2 (2008), pp. 367–379.

[61] Styliani Verykokou and Charalabos Ioannidis. “Automatic rough georeferencing of multiview oblique and vertical aerial image datasets of urban scenes”. In: *The Photogrammetric Record* 31.155 (2016), pp. 281–303.

[62] Styliani Verykokou and Charalabos Ioannidis. “Oblique aerial images: a review focusing on georeferencing procedures”. In: *International journal of remote sensing* 39.11 (2018), pp. 3452–3496.

[63] K. Vos et al. “Sub-annual to multi-decadal shoreline variability from publicly available aerial imagery”. In: *Coastal Engineering* 150 (2019), pp. 160–174. doi: 10.1016/j.coastaleng.2019.04.004.

[64] Waterschap Scheldestromen. *Hydrologisch Onderzoeksrapport Walcheren*. Tech. rep. Definitieve versie 12-12-2017. Middelburg, The Netherlands: Waterschap Scheldestromen, 2017. url: https://scheldestromen.nl/sites/default/files/2023-11/Hydrologisch%20onderzoeksrapport%20Walcheren.pdf.

[65] Matthew J Westoby et al. “‘Structure-from-Motion’photogrammetry: A low-cost, effective tool for geoscience applications”. In: *Geomorphology* 179 (2012), pp. 300–314.

[66] Chris Zevenbergen et al. “Taming global flood disasters. Lessons learned from Dutch experience”. In: *Natural hazards* 65 (2013), pp. 1217–1225.

[67] Benchun Zhou, Maximilian Gilles, and Yongqi Meng. “Structure SLAM with points, planes and objects”. In: *Advanced Robotics* 36.20 (2022), pp. 1060–1075.