The Influence of the Point Cloud Comparison Methods on the Verification of Point Clouds Using the Batavia Reconstruction as a Case Study

Petra Helmholz
Spatial Sciences, Curtin University, Australia

David Belton
Spatial Sciences, Curtin University, Australia

Nick Oliver
HIVE, Curtin University, Australia

Joshua Hollick
HIVE, Curtin University, Australia

Andrew Woods
HIVE, Curtin University, Australia

Abstract

During the excavation of the *Batavia* shipwreck in the 1970s part of the wooden hull and additional artefacts were excavated and conserved. In the course of the excavation over 3500 underwater images were captured to document the expedition. These images have allowed for the digital 3D reconstruction of the *Batavia* as it was found on the sea floor. The physical shipwreck timbers were reconstructed and are now located in the Fremantle Shipwrecks Museum. In 2014, a Terrestrial Laser Scanner (Leica C10) was utilised to capture the shipwreck as displayed in the museum. This resulted in two different point clouds of the vessel: a) based on the images captured underwater and, b) based on laser scan data captured in the museum. The case study will focus on two different sections of the vessel with different complexity, and has two aims. Firstly, the goal is to compare the 3D reconstructions from the underwater archival photography with the laser scans and 3D reconstruction of the *Batavia* hull currently on display to verify if the *Batavia* as displayed today has the same configuration as the original wreck. The results show that while the difference between both datasets is often minimal, some differences are visible because the excavation took place over a long period. Secondly, the paper will evaluate the influence of different methods for point cloud comparison. The methods are namely point-to-point and point-to-model and are used to quantify the differences of the shipwreck's point clouds. Even though the same input point clouds were utilised, the paper will show that the results of the comparison were influenced by the point-cloud-comparison method used.

Keywords

Photogrammetry, laser scanning, verification, underwater archaeology, point cloud, comparison, 3D reconstruction

Introduction

Over the last few decades, photogrammetry has become a powerful tool for the recording of archaeology and significant heritage sites due to the relatively simple process of data capturing and image processing to generate 3D maps and models. While the use of underwater photography for site recordings is well established (a historical overview is given in Drap 2012), there have been recent advances in image processing technology that allow fast 3D mapping of sites utilising a high level of automation. These improved methods can be utilised for a wide range of different image data, and importantly for archaeology allows the 3D capturing of sites. This has included both land-based (Martorelli *et al.* 2014) and underwater recordings in seas (among others Balletti *et al.* 2015; Bandiera *et al.* 2015) and lakes (Bucci 2015). It has also allowed for the joint processing of in-air and underwater projects (Menna *et al.* 2015).

Nevertheless, not only has the advanced automatic processing of the images contributed to the current uptake in the use of photogrammetry for underwater archaeology, it has also seen an increase due to advances in digital cameras that are light, highly light-sensitive, wide-aperture, compact user friendly and low cost (McCarthy and Benjamin 2014). Hence, photogrammetry is used for small (McCarthy and Benjamin 2014) and medium-scale projects (Diamanti and Vlachaki 2015), as well as for large-scale projects in combination with bathymetry (Bruno *et al.* 2015). However, processing using photogrammetry is not

limited to images from current projects; it can also be performed based on archival materials of objects of interest (Metres *et al.* 2014). Nevertheless, alternative capturing methods exist in order to create 3D models of objects underwater. Some of these include structured light sensors (Bräuer-Burchardt *et al.* 2015), multibeam sonar (Bruno *et al.* 2015), airborne laser scanning (Doneus *et al.* 2015), triangulation sensors (Ekkel *et al.* 2015) and underwater laser scanning (Davidde Petriaggia and Gomez de Ayalab 2015). All of those methods create a 3D point cloud with varying accuracy, precision and point spacing.

So far, one of the large advantages of the derived 3D models is the possibility to visualise complex sites (Repola *et al.* 2015), and to create fly-through videos (McCarthy and Benjamin 2014). These methods of visualisation, as well as the production of 3D prints, help with communicating the findings to a large audience. This benefits projects, as the products can be made available within publications, exhibitions, web pages and more. On the other hand, the derived models also allow further analysis in respect to mapping of sites and creating a GIS representation (Drap 2012), performing surface analysis of artefacts (McCarthy and Benjamin 2014), providing a reference for restoration and renovation processes through the creation of 3D CAD models (Martorelli *et al.* 2014), and for the verification of theoretical models of artefacts (Drap 2012). Within the photogrammetric processing the most prominent software solutions used for underwater projects are Agisoft's PhotoScan (McCarthy and Benjamin 2014; Metres *et al.* 2014; Van Damme 2015), Autodesk's 123D Catch (D'Amelio *et al.* 2015; McCarthy and Benjamin 2014), PhotoModeler (Green *et al.* 2002) and VirtualMapper (Green *et al.* 2002). Alternative solutions also exist such as VisualSFM (Hollick *et al.* 2013; Wu 2013;), and it can be predicted that the number of software solutions will grow in the future.

However, underwater photogrammetry does face challenges. One example is the effect of moving objects such as seagrass (McCathy and Benjamin 2014). The constant movement of seagrass potentially disrupts the matching process within the software leading to an inability to create sparse and the following dense point clouds on those regions. Further challenges for photogrammetry are the reduced light at greater depths (Van Damme 2015), sufficient light creation at deep-water sites using ROVs (Drap *et al.* 2015) and reduced visibility (Van Damme 2015).

Nevertheless, when using new technologies one important aspect is to verify the accuracy of the captured and derived data. A number of different ways of assessing the accuracy were applied in the past and can be categorised as follows. The first is the use of single measurements of objects of interest (McCarthy and Benjamin 2014; Metres *et al.* 2014) or the comparison of control point observations (Green *et al.* 2002). Another is to use graphical representations of the scenes to examine the results (D'Amelio *et al.* 2015). The final group of methods is to perform point cloud comparison (D'Amelio *et al.* 2015; Martorelli *et al.* 2014; Troisi *et al.* 2015).

The main issue with evaluating the accuracy based on single measurements or control points, is that only a limited number of discrete observations are utilised while the majority of the derived 3D information is not considered. Furthermore, by comparing only discrete measurements systematic error influences, e.g. based on the location of the objects of interest (in the centre of the survey area or the outside), its orientation (pointing in a specific direction) as well as their dimension (smaller objects vs larger objects) is difficult to detect. For instance, McCarthy and Benjamin (2014) focussed on two shipwrecks—one in shallow water (3m) and one in deeper water (13–16m). The accuracy of the derived models was verified by comparing its measurements with measurements of the objects of interest taken on the sea floor. The results are summarised with a usual high level of accuracy, however, some artefacts in the dense point cloud based on miss-matches in the dense matching process are clearly visible in the derived models, but were not considered in the accuracy assessment. A similar assessment of the accuracy was done in Metres *et al.* (2014) where the derived 3D models based on archived video material are compared to the hand survey measurements with ranged errors between 18% (for smaller objects of 0.37m dimension) and 0.56% (for larger objects of 4.21m dimension). Green *et al.* (2002) used control points and grids to check the precision of photogrammetric site survey as well as tape measure trilateration. Control points and points on the grid were observed and adjusted in order to check their precision, i.e. the agreement of the observations to each other. However, a cross-comparison is missing.

For practical aspects of archaeological documentation, the accuracy of discrete points is often not relevant. Instead, in many cases the focus is to describe the geometrical aspects of objects in order to allow an identification and classification of those objects. For this purpose, D'Amelio *et al.* (2015) created 2D profiles of amphorae and inspected them visually. However, the profiles were created based on 3D point clouds. Hence, instead of comparing the whole object with all given information (points in the point cloud), a reduction of information is produced and used.

It is possible to compare actual point clouds with each other so that no reduction of information is necessary. Within the point cloud the comparison can be limited to

small objects (D'Amelio *et al.* 2015, compares amphora) or medium to large-scale projects (Martorelli *et al.* 2014, examines ship hulls). To provide a more detailed example, Martorelli *et al.* (2014) used a laser scanner to compare the results of a 3D point cloud derived based on images of a ship hull, and found that the average deviation of 1.25mm based on the Euclidian methods. However, the provided heat map shows much higher maximal values (up to 15 mm), and a clear systematic trend (larger positive values are located on the outside, and larger negative values in the inside of the compared hull face). Therefore, an important aspect when comparing point clouds is to consider how they are compared, as there are a number of different ways. In this publication, we will show that by comparing the same point clouds the results can be quite different only depending on how the point clouds are compared.

However, before introducing two methods for point cloud comparison and the analysis of their results, it is important to understand the data involved in the comparison. In the majority of publications related to archaeology where point clouds are derived from photogrammetry, a comparison is performed against point clouds captured with a laser scanner, with the laser scanner point cloud usually used as the reference.

Therefore, the structure of this paper is as follows: in the next section we will introduce our case study site, the *Batavia* shipwreck, focusing especially on the point clouds derived from underwater images and those captured using a laser scanner of the vessel's physical reconstruction in Western Australia's Shipwrecks Museum in Fremantle. Not only are the data capture and the best practice of processing the data briefly outlined in this section, but we also discuss the data sources with respect to accuracy, precision and point spacing. Afterwards, two selected methods for the point cloud comparison are introduced in detail (namely point-to-point and point-to-model) including the prediction of their results. Finally, we compare the two point clouds—photogrammetry based on underwater images and laser scanner data based on the data captured at the shipwreck 'recovered and reconstructed' at the WA Shipwrecks Museum. The evaluation will discuss two aspects. Firstly, how accurate the reconstructed shipwreck in the museum is compared to its state underwater when it was discovered. Secondly, the impact of using different point cloud comparison methods will be evaluated.

Case study—data

During the excavation of the *Batavia* in the 1970s, the remaining wooden hull was excavated and conserved, along with additional artefacts. In the course of the excavation over 3500 underwater images were captured to document the expedition. These images have allowed the digital 3D reconstruction of the *Batavia* as it was found on the sea floor (Woods *et al.* 2016). The physical shipwreck timbers were physically reconstructed and are now located in the WA Shipwrecks Museum. In 2014, the shipwreck as displayed in the museum was surveyed using a Terrestrial Laser Scanner (Leica C10), which allowed the production of two different point clouds of the vessel—a) based on images captured underwater, and b) based on laser scan data captured in the museum. This case study will focus on comparing the different point clouds with respect to two different sections of the vessel with different complexity and properties.

Laser scanning

When capturing the data of the *Batavia* shipwreck as presented in the WA Shipwrecks Museum a surveying grade Terrestrial Laser Scanner system was used, namely the Leica C10 scanner (Annesley 2014). The distance measurement accuracy of this scanner is 6mm, the angle accuracy (vertical) is 12", the angle accuracy (horizontal) is 12" and the model surface precision is 2mm (Leica Geosystems 2016). A total of four different scans were captured spaced around the shipwreck with three being captured on the ground floor and one being captured from the mezzanine walk-way. The scans of all four stations were registered together to form a point cloud with 4,370,279 points covering the inside as well as the outside of the hull.

The scanning results from the inside of the hull are shown in Figure 1, and represent the primary focus of the comparison in this paper. Due to laser scanning being an active system (emits its own light source) the scanner also captures returned intensity information, which is a function of the scanner geometry to the surface and the surface properties such as reflectance and texture in relation to the laser wavelength. Instead of a coloured point cloud, the values shown in Figure 1 represent intensity information only. Based on the metric point cloud it is easy to derive profiles and perform a range of measurements (distances, profiles, volumes, etc.) as given in Figure 1. In order to derive accurate measurements it is required that the scanner operates within the specifications, i.e. that the scanner is calibrated. A calibration of the used scanner passed the required specifications.

Laser scanner data is often used as the reference dataset, as its errors and error models for correction are well studied. Point clouds derived from laser scanners are accurate (globally) but can contain some noise/lower precision (locally) (Figure 1, bottom left). The noise is due to the integration time of the return signal and includes noise in the return signal, as well as

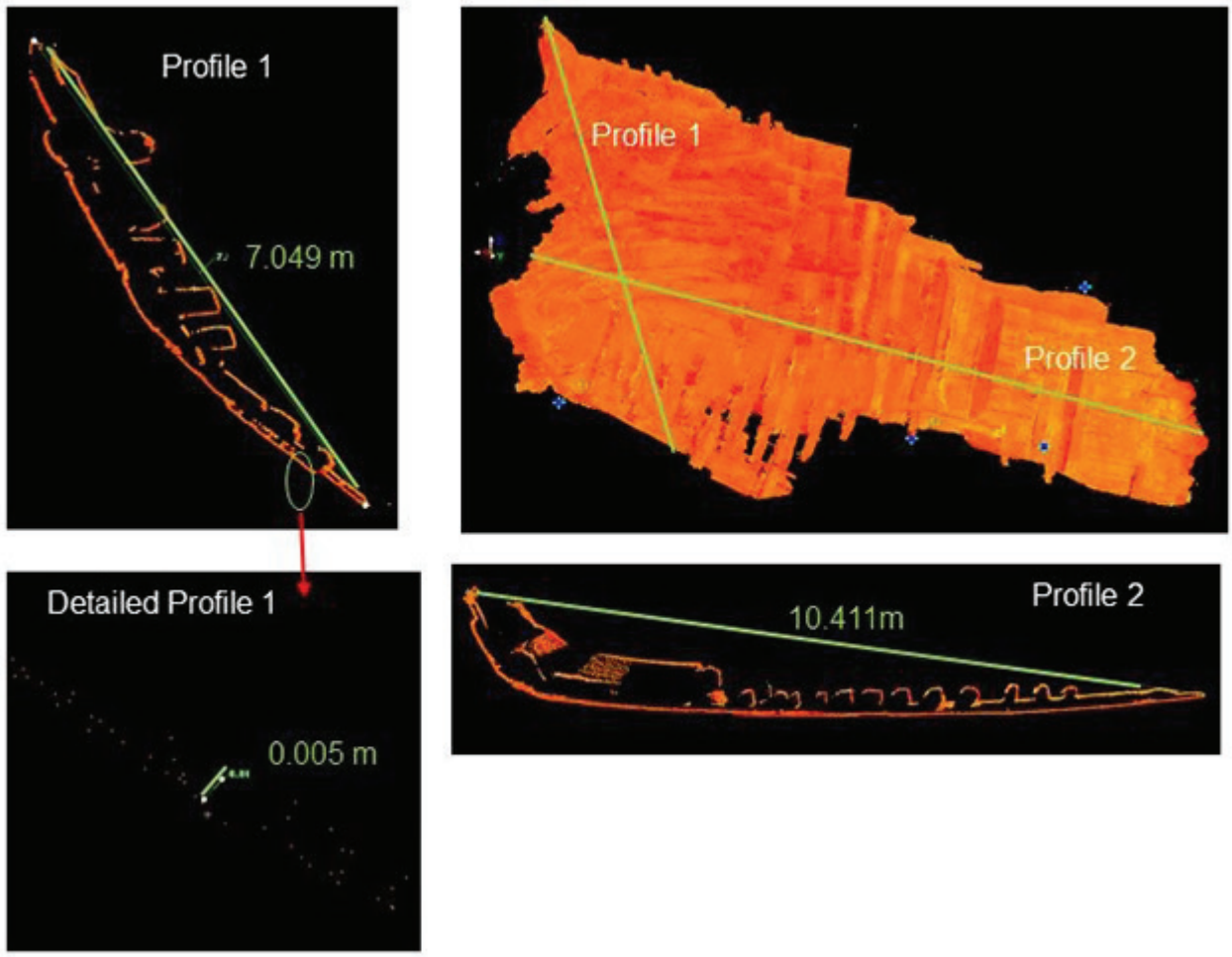

Figure 1. Point Cloud of the *Batavia* shipwreck captured with the Leica C10 laser scanner at the museum (top right) as well as two selected profiles (top left and bottom right) with a detailed section (bottom left).

noise from the beam interaction over the surface. If the distribution follows a Gaussian distribution then local interpolation should smooth out the noise and preserve the trend. The spacing of the points depends on the scanner resolution and the number of captured scans.

Photogrammetry

Based on the underwater images a number of digital 3D reconstructions of the *Batavia* have been created (Woods *et al.* 2016). To be able to extract precise measurements, the geometry in which the images are captured, the quality of the cameras and lenses used and the *in situ* calibration of the camera are all important. Sufficient and homogenous illumination was present when the underwater images were taken without any shadow areas.

From the overlapping images of the object of interest, the position of the camera can be determined. For this process Agisoft's PhotoScan was used. The metrics related to these 3D measurements were added using the known dimension of grids placed in the field of view of the camera. The generated point clouds were further processed to produce a textured meshed surface. Figure 2 shows the two selected image-based point clouds (MF1 and MF37) and their location relative to the laser scanning point cloud data.

The errors in image-derived point clouds are a function of: a) the calibration quality of the camera (interior orientation), b) the accuracy of the image alignment (exterior orientation), and c) the quality of the dense image matching procedure. The dense matching (point cloud creation) works well if the object is smooth and without sharp changes in the topography as well as having a well-textured surface. Occlusions and surfaces lacking sufficient texture often lead to mismatches (Figure 2, profile). Models derived from images usually have the characteristic of being precise but may be less accurate compared to laser scanning point clouds (Figure 2, right detailed plots), i.e. while small details

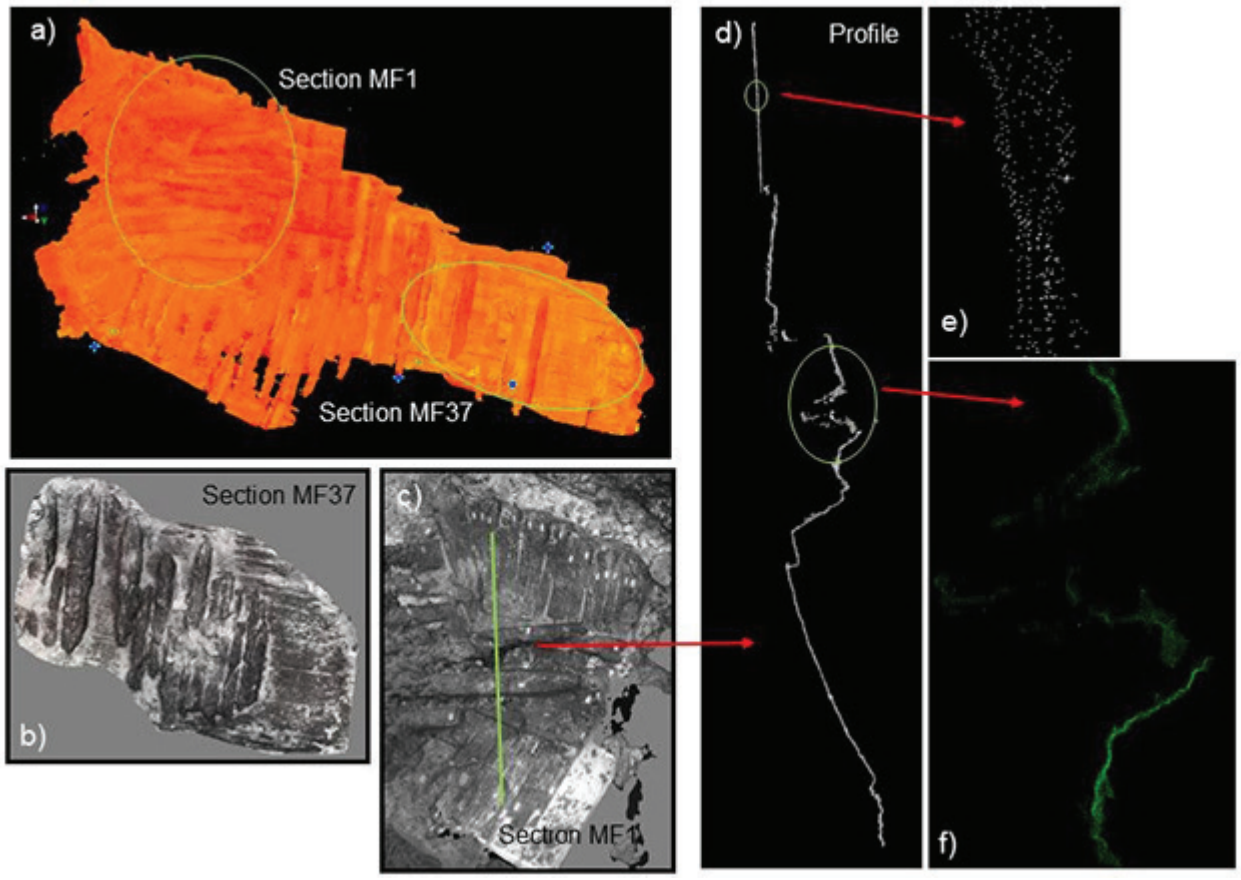

Figure 2. Intensity point cloud of the ship wreck at the museum (a), rendered 3D point clouds based on the underwater images captured of the *Batavia* – Section M37 (b) and section MF1 (c), selected profile of MF1 (d) and its detailed views 1 (e) and (f).

are usually reconstructed well, the point cloud itself can show inaccuracy globally (e.g. a surface warping effect).

Methodology

The comparison of the underwater image-derived point cloud with the laser scanning point cloud will use three different methods: visual inspection, point-to-point (P2P) comparison, and point-to-model (P2M) comparison. Through visual inspection, it should be possible to detect larger differences between the point clouds such as missing timbers. However, differences within cm or mm range will not be as easily detected using this approach. On the other hand, it should be possible using the P2P and the P2M method to detect these smaller differences. While there are a number of other methods, this paper will focus only on these two.

Point-to-Point (P2P) comparison

If it can be assumed that the density of the two point clouds is high, or the point sampling between the two are similar, then the distance between the nearest neighbouring points is close to the true distance between the point clouds (Figure 3, top). This is the basis of the point-to-point method. However, there is the issue of no 'real' correspondence between points as would be the case if control points or markers were utilised. This results in a bias ε_1 dependent on the sampling distance [Figure 3, top right]. The average distance is the maximum between the sampling distance and the point cloud separation. Hence, when using this method there will be no distance value of '0' (because 'real' correspondence do not exist), and the mean distance may be slightly larger than the true average surface separation.

Point-to-Model (P2M) comparison

Using the P2M comparison, the surface of one point cloud is locally modelled using a Least Square Plane fitting method, taking into consideration the k nearest neighbouring points (kNN). Then, the distance between this model and the closest point in the other point cloud is calculated as the projected distance between the points onto the normal vector of the modelled surface (Figure 3, bottom). While this eliminates the errors

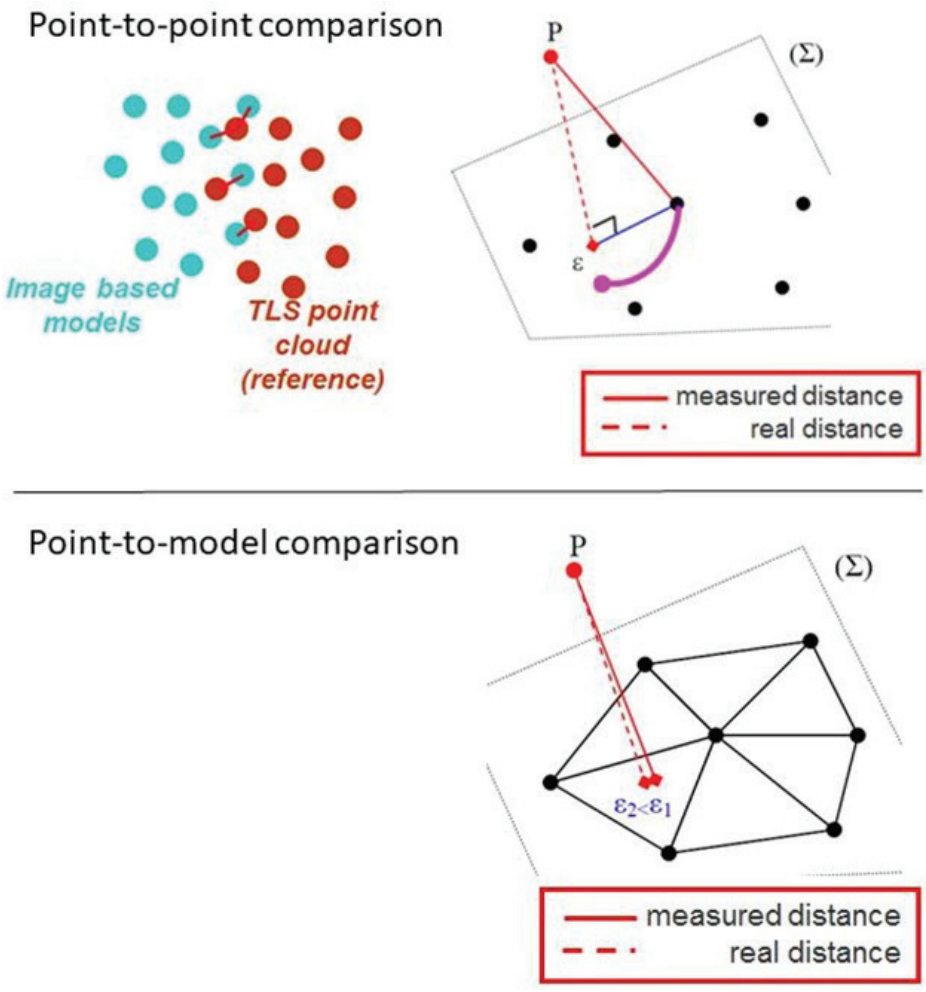

Figure 3. Point-to-point comparison (top) and point-to-model comparison (bottom) (adopted from CloudCompare, 2017.)

due to missing corresponding points, there are still the remaining issues of the model resolution, which will introduce a bias ε_2. This is due to noise in the points to the surface and slight errors in the surface normal approximation to the true surface normal. However, this bias is much smaller than those that are exhibited when using the P2P method.

Data processing

The data processing can be separated in three steps:

a. Alignment of the point clouds;
b. Application of the comparison methods; and
c. The analysis of the comparison.

In order to align the point clouds a rough alignment is the initial step. This alignment is usually done picking a number of identical points (at least four) in the two point clouds and applying a seven parameter Helmert transformation based on the corresponding points. The Helmert transformation will rotate (three degrees of freedom), translate (three degrees of freedom) and scale (usually one degree of freedom) the sample point cloud onto the position of the reference point cloud. It is recommended to apply a scaling factor. Scaling errors can have a number of different reasons such as incorrect camera calibration, layout of the control points (e.g. extrapolation effects) or simple imprecise measurements within the image processing workflow. After the two point clouds are roughly aligned, the method of 'iterative closest point' (ICP) (Besl and McKay 1992) can be applied, and both point clouds aligned precisely. The result is that the reference point cloud and the underwater-derived point cloud are comparable.

Then the three proposed comparison methods (visual inspection, P2P and P2M) are applied. Output is a) a heat map on which the distances between the point clouds are colour mapped, b) statistical information such as the 95% significance level, minimum distance,

maximum distance, distance with the most counts and c) a histogram plotting the distances. Beside a visual inspection of the histogram, it is possible to determine the Weibull distribution parameters, i.e. the shift parameter *D* (offset from 0), the shape parameter *a* (larger values means the distribution is moved away from zero), and the scale parameter *b* (assumed to be between $0 < b < 1$, larger values means the distribution is more stretched out).

Results

Visual Inspection

The hull timbers of the shipwreck were removed from the sea floor over a series of several expeditions in the 1970s (Green 1989). As this process was undertaken, underwater images were captured and a hand-drawn site diagram was developed over the period (Van Duivenvoorde 2015). The digital 3D model generated from the underwater images helps represent the state of the ship at this time; however, it was impossible to capture the detail in one period. As such, the full model had to be built from groups of images collected over successive periods as the excavation progressed. In contrast, the laser scan of the ship in the museum could be captured in a single instance. While the two models should essentially be the same, readily apparent differences can be seen. When looking at Figure 4 showing site MF37, it becomes clear that the opposite is the case, e.g. timber parts are missing in the museum model (Figure 4, bottom, yellow arrows). These are due to the condition of the timbers—not all timbers could be restored and, therefore, are not present in the museum display.

While the excavation of site MF37 was done in one excavation period (December 1972 to May 1973), the excavation period for ME1 took place over two periods within January and December 1974. More time was required to excavate the more complex structure of this site. When looking at Figure 4 it becomes clear that the matching of the three different information sets (excavation site diagram, image-based point cloud, and a photo of the *Batavia* in the museum) is also more challenging due to the complexity of the hull. However, for this site, it is true that the processed underwater model, representing just a snapshot of the nearly year-long excavation of this site, does not contain all the recovered timber pieces.

Point-to-Point comparison

The results of the point-to-point comparison of MF 1 and MF37 are presented in Figure 5. The missing timbers in MF37 are easily recognisable in the figure due to their detected differences being in the order of 0.015m and higher. The colours mapping the differences are consistent with those shown in the histogram to highlight the distribution of distances in the comparisons. Beside the missing timbers, the figure shows that the differences between the reconstruction in the museum compare well with the captured underwater model, and is within 0.01m. However, after closer examination of the histogram, a clear bias away from zero (also indicated by the Weibull shift parameter *D* of 0.000116m) is visible. This bias confirms the issue with point correspondences as discussed in the previous section. In addition, the histogram shows a strong peak in the 0.005m to 0.01m interval range. This bias, and the peak, would not be easily observable from just the raw statistical results. Ninety-five percent of all distances are within the 0.071m range, the mean difference is 0.031m, and the standard deviation 0.021m. All values are within the expected range, it can be seen that they are biased away from zero by the missing elements and the comparison method.

The results of the P2P comparison of site ME1 confirms the results of site ME37. Again, the missing timbers are easily recognisable in the figure indicated by large differences (the regions in yellow and red that are more distinct). Please note that the same colour scale for visualising the distances between the point clouds was used in Figures 9–12. This histogram also exhibits a similar response to the previous site, showing the same bias (shift parameter is 0.000115m). Again, statistical values are within the expected range but with higher values due to the higher number of missing timbers in the underwater model and the complexity of the site. 95% of all distances are within the 0.130 m range, the mean difference is 0.058 m and the standard deviation is 0.038 m.

Point-to-Model comparison

Figure 6 presents the results of the P2M comparison. When only inspecting the distances displayed in the figure, the difference to the P2P comparison result is minimal. Visually, it is difficult to identify any differences between both methods. The differences become more apparent when inspecting the statistical values. Ninety-five percent of all distances for MF37 are within the 0.075 m (instead of 0.071 m) range, the mean difference is 0.029 m (instead of 0.031 m), and the standard deviation is 0.022 m (instead of 0.021 m) and therefore apparently producing a better result. Even though the differences are small, it is recognisable that when using the same point clouds the comparison result changes due to using a different comparison method. The differences between the P2P and P2M statistical values are small due to this site not being overly complex. Therefore, the impact of the missing timbers is small. However, a significant difference

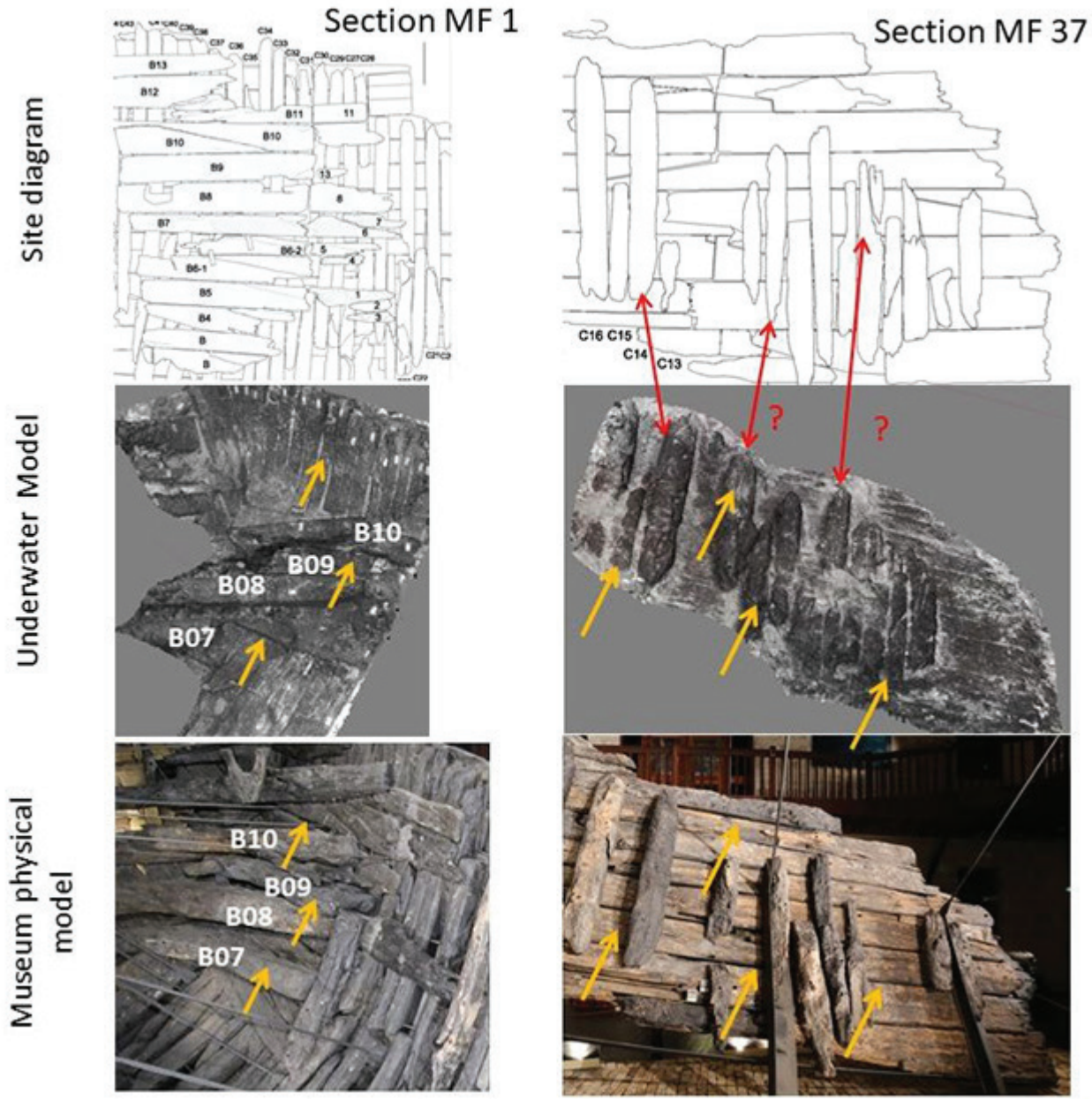

Figure 4. Results of the visual inspection of the point clouds of site MF1 (left) and MF37 (right). The top row shows the site diagrams of each site, the middle row the underwater model, and the bottom row the physical model in the museum. The red arrows indicate matching positions, yellow arrows indicating missing timbers in either the underwater model or the reconstruction at the museum. The location of the yellow arrows match. Site diagram from (W. van Duivenvoorde, 2015).

is visible when looking at the histogram. The bias away from zero has also been resolved as predicted in previous section. Furthermore, the peak of differences in the green area disappears, too. The histogram looks more balanced and closer to the expected trend.

Finally, in Figure 6 (top), the results of the more complex scene MF1 using the point-to-model comparison method are shown. Due to the complexity of this scene the differences of this heat map compared to the one presented MF37 become visible. For instance, the gaps in blue in the image point cloud due to missing timber parts (yellow and red areas) are more visible. Those areas also influence the statistical values, i.e. 95% of all distances are within the 0.119m (instead of 0.130m) range; the mean difference is 0.048m (instead of 0.058m); and, the standard deviation is 0.037m (instead of 0.038m) and again producing apparently better results. From comparing these statistical values, a clear improvement is visible when using the P2M comparison method. The improvement in the histogram of site MF1 is similar to the improvement seen in the histogram of site MF37 using the same comparison method confirming the previous findings.

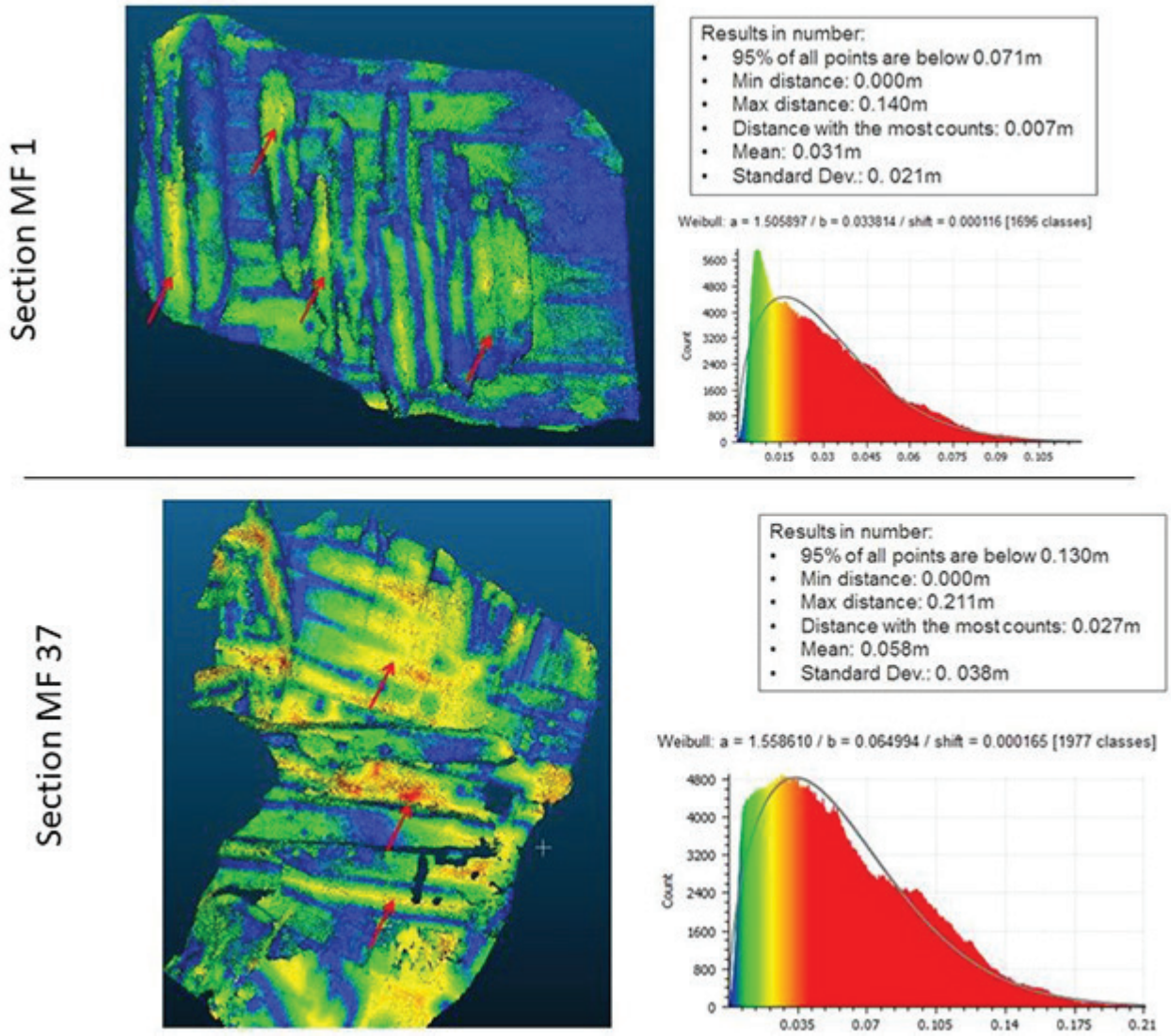

Figure 5. Results of the point-to-point comparison of the point clouds of site MF 1 (top) and MF37 (bottom). On the left: heat map, top right: summary of the statistic information, lower right: histogram with Weibull parameters. The red arrows indicate the areas with the largest differences between the point clouds.

Conclusion

The primary conclusion regarding the physical reconstruction of the *Batavia* hull in the WA Shipwrecks Museum compared to the underwater digital 3D models is that the overall reconstruction is observed to be close to the original, besides the detection of some missing timbers. Reasons for missing timber in the digital 3D underwater model are that the excavation took place over a long period, so underwater 3D models only present a snapshot in time, but not the overall extent. In order to make a more complete conclusion the different underwater models should be merged before comparing them to the point cloud of the physical model of the vessel. The reasons for the missing timbers in the physical reconstruction are due to the fact that not all timbers could be restored and, therefore, are not present in the museum display.

It needs to be highlighted that the goal of the comparison in this paper was to show that different comparison methods can create different results, even when the same point clouds are used for the comparison. In order to compare the underwater point cloud with the physical model, a scaling correction was performed to make the two point clouds comparable. Therefore, scaling differences (e.g. due to the exposure of the timber to the atmosphere in the museum) could not be detected. If scaling differences are the focus of the study, the scale must be calculated independently of the model. However, this was outside of the scope of this work.

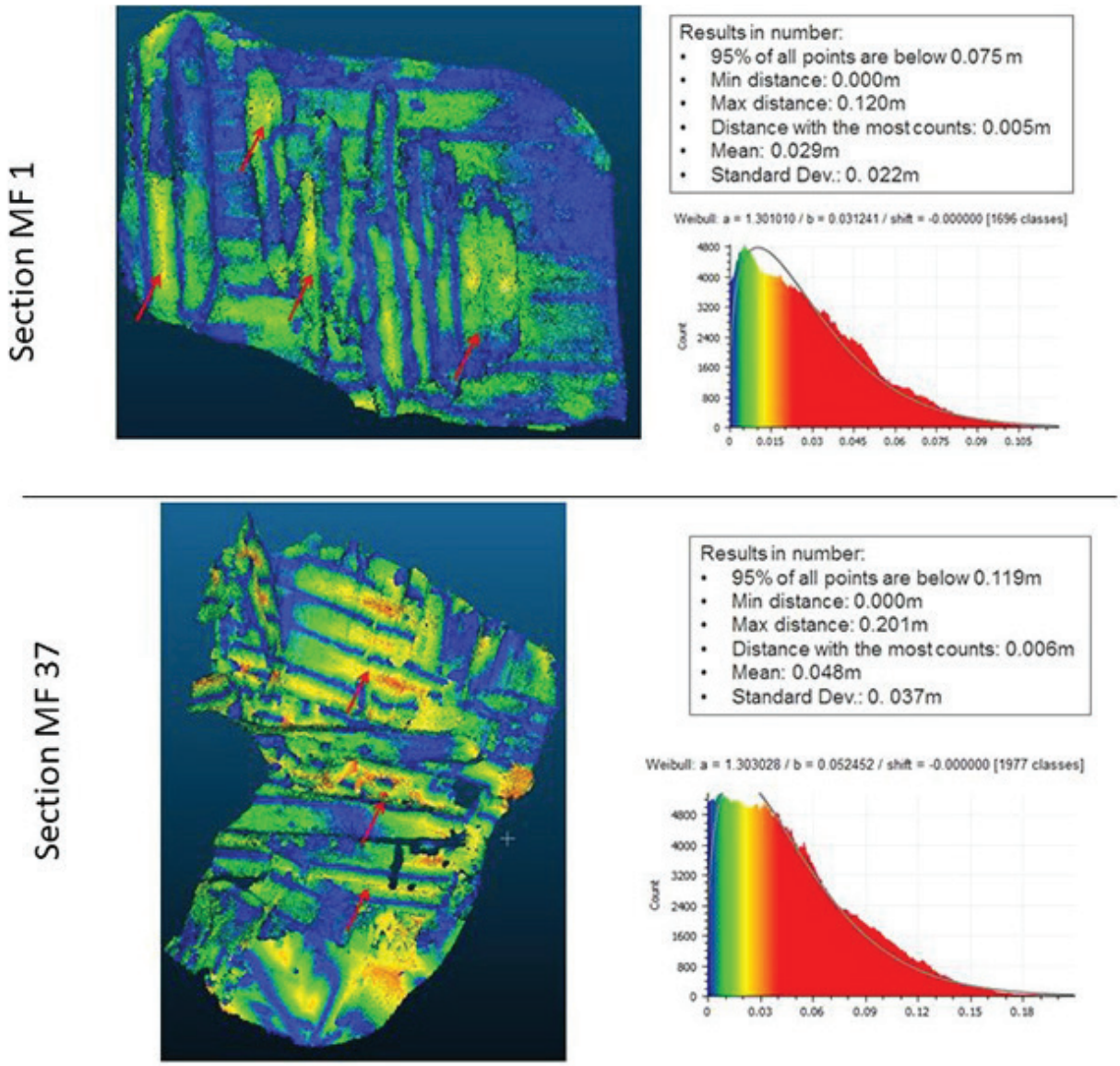

Figure 6. Results of the point-to-model comparison of the point clouds of site MF1 (top) and MF37 (bottom). On the left: heat map, top right: statistical information, lower right: histogram with Weibull parameters. The red arrows indicate the areas with the largest differences between the point clouds.

The method used in comparing point clouds is important and depends on three main factors: the density, accuracy and the precision of the point clouds compared. While the influence of the comparison method can be different for different complex scenes, a comparison using the P2M method over the P2P method is preferable. It can deal with differences in densities of point clouds, as well as lower accuracy and precision.

During this work, it could be demonstrated that the choice of the comparison method has an impact on the comparison results. This means that even though the same point clouds for the comparison were used, the results are different. The choice of the comparison method has a lower impact when dealing with less complex scenes and higher sampling densities, i.e. the results are 'more similar'. This conclusion is made based on the similar statistical values (mean, standard deviation, max and most common difference) of the P2P and P2M comparison of site ME 37 in Table 1 (column 2 and 3). However, if the scene is more complex and the two point clouds compared have higher differences this is not true anymore. In Table 1 (column 4 and 5) a clear 'improvement' of the values are visible when using the P2M over the P2P comparison method.

Paying special attention to the shift parameter (D) of the histogram a bias away from '0' can be observed in the P2P comparison, but not in the P2M comparison. In general, while it can be observed that the shape parameter (a) is larger for the P2P comparison and smaller for the P2M. The shape parameter (b) is similar

Table 1: Comparison of the statistical values and the histogram-derived values for ME37 and ME1 using the point-to-point (P2P) and the point-to-model (P2M) method.

	ME 37 – P2P	**ME 37 – P2M**	**ME 1 – P2P**	**ME 1 – P2M**
95% of the values are within [m]	0.071	0.075	0.130	0.119
Max [m]	0.140	0.140	0.211	0.201
Distance with the most counts [m]	0.007	0.005	0.027	0.006
Mean [m]	0.031	0.029	0.058	0.048
Standard Dev. [m]	0.021	0.022	0.038	0.037
Weibull shape a	1.505897	1.301010	1.558815	1.303028
Weibull scale b	0.033814	0.031241	0.064994	0.052452
Weibull shift	0.000116	0	0.000165	0

for ME37, so it does not matter which comparison method was used, however this parameter is shown to be significantly higher for ME1. This means that the ME1 histograms are more stretched (due to more missing timber elements). Within the ME1 site the value for *b* is slightly lower when using the P2M method compared to the P2P method.

Nevertheless, even though there are differences in the comparison results using the different methods, the comparison of the derived point clouds is preferable over comparison of just a few discrete measurements, as it shows missing elements and systematic trends more significantly.

Acknowledgements

Jeremy Green, Department of Maritime Archaeology, Western Australian Museum, Fremantle, for arranging the opportunity to laser scan the physical reconstruction of the *Batavia* hull in the WA Shipwrecks Museum.

References

Annesley, D. 2014. Comparison the accuracy of different 3D construction software on the Batavia Wreck. *iVEC Internship report*, 2014.

Balletti, C., C. Beltrame, E. Costa, F. Guerra and P. Vernier 2015. Underwater Photogrammetry and 3D reconstruction of marble cargos shipwrecks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 7-13.

Bandiera, A., C. Alfonso and R. Auriemma 2015. Active and passive 3D imaging technologies applied to waterlogged wooden artefacts from shipwrecks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 15-23.

Besl P.J. and N.D. McKay 1992. A Method for Registration of 3-D Shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. Los Alamitos, CA, USA: IEEE Computer Society. 1992; 14 Suppl 2: 239–256.

Bräuer-Burchardt, C., M. Heinze, I. Schmidt, P. Kühmstedt and G. Notni 2015. Compact handheld fringe projection based underwater 3D scanner. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XL-5/W5: 33-39.

Bruno, F., A. Lagudi, A. Gallo, M. Muzzupappa, B. Davidde Petriaggi and S. Passaro 2015. 3D documentations of archaeological remains in the underwater park of Baiae. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 41-46.

Bucci, G. 2015. Padus, Sandalus, Gens Fadiena – Underwater surveys in Palaeo-Watercourses (Ferrara District - Italy). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 55-60.

CloudCompare. CloudCompare v2.6.1 - User manual. http://www.cloudcompare.org. Accessed March 2017.

D'Amelio S., V. Maggio and B. Villa 2015. 3D Modelling for underwater archeological documentation: metric verifications. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 73-77.

Davidde Petriaggia, B. and G. Gomez de Ayalab, 2015. Laser scanner reliefs of selected archeological structures in the submerged Baiae (Naples). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XL-5/W5: 79-83.

Diamanti, E. and F. Vlachaki 2015. 3D Recording of underwater antiquities in the South Euboean Gulf. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 93-98.

Doneus, M., I. Miholjek, G. Mandlburger, N. Doneus, G. Verhoeven, C. Briese and M. Pregesbauer 2015. Airborne Laser Bathymetry for documentation of submerged archaeological sites in shallow water. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 99-107.

Drap, P., D. Merad, B. Hijazi, L. Gaoua, M.M. Nawaf, M. Saccone, B. Chemisky, J. Seinturier, J.-C. Sourisseau, T.Gambin and F.Castro, 2015. Underwater Photogrammetry and object modelling: a case study of Xlendi wreck in Malta. *Sensors* 15: 30551-30384.

Drap, P. 2012. Underwater Photogrammetry for Archaeology, Special Applications of Photogrammetry, in D.C. Da Silva (ed.) ISBN: 978-953-51-0548-0, *InTech*, Available from: http://www.intechopen.com/books/special-applications-of-photogrammetry/underwater-photogrammetry-forarchaeology.

Ekkel, T., J. Schmik, T. Luhmann and H. Hastedt 2015. Precise laser-based optical 3D measurement of welding seams under water. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 117-122.

Green, J.N. 1989. The loss of the *Verenigde Oostindische Compagnie retourschip BATAVIA*, Western Australia 1629 (British Archaeological Reports International series 489). Oxford: Archaeopress.

Green, J., S. Matthews and T. Turanli 2002. Underwater archaeological surveying using PhotoModeler, VirtualMapper: different applications for different problems. *The International Journal of Nautical Archaeology*. 31.2: 283-292.

Hollick, J., S. Moncrieff, D. Belton, A.J. Woods, A. Hutchison and P. Helmholz 2013. Creation of 3D models from large unstructured image and video datasets. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-1/W1: 133-137.

Leica Geosystems, Leica ScanStation C10 Datasheet. http://www.hds.leica-geosystems.com/en/Leica-scanStation-c10_79411.html Accessed November, 2016.

Martorelli, M., C. Pensa and D. Speranza 2014. Digital Photogrammetry for Documentation of Maritime Heritage. *Journal of Maritime Archaeology* 9: 81-93.

McCarthy, J. and J.Benjamin 2104. Mulit-image photogrammetry for underwater archaeological site recording: an accessible driver-based approach. *Journal of Maritime Archaeology*9: 95-114.

Menna, F., E. Nocerinoa, S. Troisi and F. Remondino 2015. Joint alignment of underwater and above-the-water photogrammetric 3D models by independent models adjustment. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 143-151.

Metres, J., T. Thomsen and J. Gully, 2014. Evaluation of structure from motion software to create 3D models of late nineteenth century great lakes shipwrecks using archived diver-acquired video surveys. *Journal of Maritime Archaeology* 9: 173-189.

Repola, L., R. Memmolo and D. Signoretti 2015. Instruments and methodologies for the underwater tridimensional digitization and data musealization. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 183-190.

Troisi, S, S. Del Pizzo, S. Gaglione, A. Miccio and R.L. Testa 2015. 3D Models comparison of complex shell in underwater and dry environments. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 215-222.

Van Damme, T. 2015. Computer Vision Photogrammetry for underwater archaeological site recording in a low-visibility environment. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-5/W5: 231-238.

Van Duivenvoorde, W. *2015. Dutch East India Company Shipbuilding: The Archaeological Study of Batavia and Other Seventeenth-Century VOC Ships*. College Station, United States of America: Texas A&M University Press.

Woods, A., N. Oliver, J. Hollick, J. Green, P. and Baker 2016. 3D Reconstruction of the Batavia (1629) wreck-site from historical (1970s) photography. *IKUWA6, 2016*. Publication pending.

Wu, C., 2013, June. Towards linear-time incremental structure from motion. In 3DTV-Conference, 2013 International Conference:. 27-134.