Article

Data Gap Classification for Terrestrial Laser Scanning-Derived Digital Elevation Models

Matthew S. O'Banion [1],* , **Michael J. Olsen [2]** , **Jeff P. Hollenbeck [3]** and **William C. Wright [1]**

[1] Department of Geography and Environmental Engineering, United States Military Academy,
    West Point, NY 10996, USA; william.wright@westpoint.edu
[2] Department of Civil and Construction Engineering, Oregon State University, Corvallis, OR 97331, USA;
    michael.olsen@oregonstate.edu
[3] Forest and Rangeland Ecosystem Science Center, U.S. Geological Survey, Corvallis, OR 97330, USA;
    jhollenbeck@usgs.gov
* Correspondence: matthew.obanion@westpoint.edu

check for updates

Abstract: Extensive gaps in terrestrial laser scanning (TLS) point cloud data can primarily be classified into two categories: occlusions and dropouts. These gaps adversely affect derived products such as 3D surface models and digital elevation models (DEMs), requiring interpolation to produce a spatially continuous surface for many types of analyses. Ultimately, the relative proportion of occlusions in a TLS survey is an indicator of the survey quality. Recognizing that regions of a scanned scene occluded from one scan position are likely visible from another point of view, a prevalence of occlusions can indicate an insufficient number of scans and/or poor scanner placement. Conversely, a prevalence of dropouts is ordinarily not indicative of survey quality, as a scanner operator cannot usually control the presence of specular reflective or absorbent surfaces in a scanned scene. To this end, this manuscript presents a novel methodology to determine data completeness by properly classifying and quantifying the proportion of the site that consists of point returns and the two types of data gaps. Knowledge of the data gap origin can not only facilitate the judgement of TLS survey quality, but it can also identify pooled water when water reflections are the main source of dropouts in a scene, which is important for ecological research, such as habitat modeling. The proposed data gap classification methodology was successfully applied to DEMs for two study sites: (1) A controlled test site established by the authors for the proof of concept of classification of occlusions and dropouts and (2) a rocky intertidal environment (Rabbit Rock) presenting immense challenges to develop a topographic model due to significant tidal fluctuations, pooled water bodies, and rugged terrain generating many occlusions.

Keywords: data gaps; occlusions; dropouts; terrestrial laser scanning; TLS; lidar; digital elevation models

1. Introduction

Data gaps/voids (i.e., the absence of data) are a common occurrence that plague remote sensing data including terrestrial laser scanning (TLS) 3D point cloud data. TLS point cloud data gaps can have an adverse effect on subsequent point cloud-derived products, including digital surface models (DSMs), bare-earth digital elevation models (DEMs), triangulated surface meshes, and 3D solid models, among others. A point cloud data gap of significant size and extent is unable to provide geometric or radiometric information to the chosen spatially continuous product; therefore, assumptions must be made to span the data gap, which inherently adds uncertainty to the derived product.

TLS data gaps stem from two primary sources (Figure 1): A line-of-sight obstacle resulting in an occlusion, and a dropout [1] resulting from a specular reflective or absorbent surface preventing the

energy from a given laser pulse from returning to the TLS instrument. The extent of 3D point cloud data gathered by a TLS instrument is limited to what is directly visible by the scanner; line-of-sight obstacles (e.g., topographic high points, trees) result in occlusions on the side of the obstacle opposite the scanner. Hence, a comprehensive survey of a complex site requires multiple scan positions from varied points of view to mitigate occlusions [2]. TLS data gaps can also stem from bodies of water [3] and other specular reflective or highly absorbent surfaces. If the surface of an object is such that a laser pulse emitted by the scanner is received and then reflected away never to return to the scanner, no point will be recorded at the current location in the scene and a data gap will result. Although dropouts can be difficult to distinguish from occlusions, the distinction between the two is important.
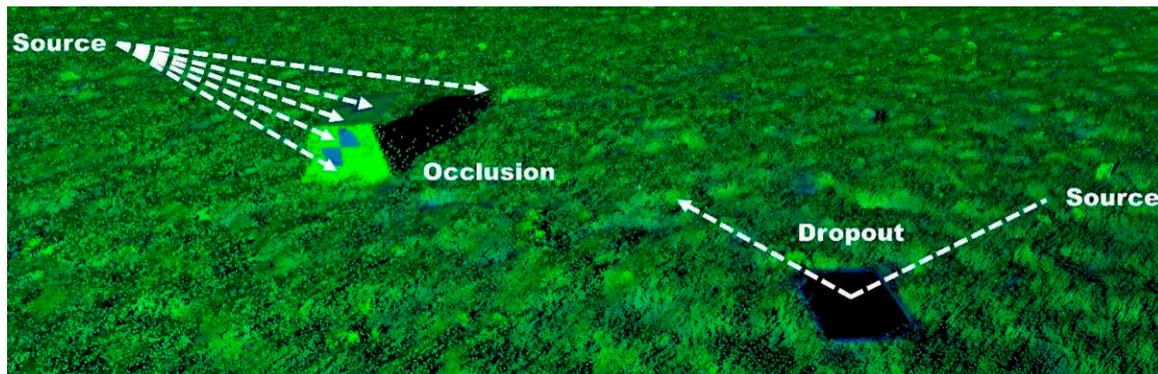

Figure 1. Examples of occlusion and dropout data gaps in terrestrial laser scanning (TLS) data.

The relative proportion of occlusions in TLS data is often indicative of the survey quality. Recognizing that regions of the scanned scene occluded from one scan position are commonly visible from another location, a prevalence of occlusions in point cloud data can indicate an insufficient number of scans and/or poor scanner placement. Conversely, a prevalence of dropouts is not tied as closely to survey quality. Generally, the scanner operator cannot control the presence of specular reflective or absorbent surfaces, be they water puddles, ponds, large bodies of water, or glass, at certain incidence angles. Consequently, data gaps resulting from dropouts typically cannot be avoided by even the most careful and comprehensive of TLS surveys. In some cases, however, the timing of the survey, particularly in locations with tidal or seasonal fluctuations, can have a substantial influence on the presence of pooled water and can be considered.

To date, there is a lack of literature concerning both the identification and classification of data gaps in TLS data as well as TLS-derived products. Existing work has explored data gap filling methods (e.g., [4]) and mitigation of occlusions [2]; however, no prior work has been identified that differentiates between occlusion and dropout data gaps. Classification of data gaps can enable optimization of DEM data gap filling. For instance, proper classification of data gaps in a DEM can enable one to select the appropriate method to interpolate such as using a thin plate spline method [4] to fill occlusions and a hydro-flattening type [5] technique to fill water-derived dropouts.

With respect to the classification of water, the literature seems to solely focus on applications relevant to airborne laser scanning (ALS) [5–8]. Unfortunately, methods for identifying and/or classifying bodies of water in ALS data are not relevant to TLS point cloud data due to differences in airborne and terrestrial points of view relative to horizontally oriented bodies of water. For instance, it is common in ALS data to have some points representative of the water surface with both low and very high laser pulse energy levels (intensity), whereas, because of the commonly oblique incidence angle of TLS observations to the ground surface, it is likely no water surface points will be captured.

An important quality metric for TLS point cloud data and derived DEMs is completeness. Given that the presence of any data gaps in a point cloud can bring the survey quality into question and lead to increased levels of DEM uncertainty from over-interpolation, a need exists for a methodology to

properly classify these data gaps, and quantify how much of the scanned area consists of point returns as well as the two types of data gaps. Knowledge of data gap origin can facilitate the judgement of TLS survey quality and the identification of pooled water in a scanned scene. Having the ability to quantify the presence of occlusions in a DEM provides the opportunity to evaluate the influence of TLS data acquisition and DEM creation parameters on the overall completeness of a given DEM [9]. Examples of these TLS parameters include angular resolution of TLS data, quantity of scans per unit area, DEM resolution, and minimum required points per DEM pixel. The proposed methodology can also communicate important information to those using TLS-derived products for scientific applications. For example, identifying pooled water has implications for habitat modeling and mapping in ecological research focused on species that respond substantially to variation in the submergent–emergent boundaries found in a rocky intertidal ecosystem [10–13].

As a result, we developed a novel data gap classification methodology that included two major steps. The first of these steps flagged the boundaries of dropout-based gaps in a projected 2D representation of the point cloud data (2D TLS Image), while the second step used the flags to classify the individual data gaps present in a TLS-derived DEM. We then applied this methodology to a field site located in the rocky intertidal ecosystem to assess a real-world application.

2. Methods

2.1. Data Gap Classification

The proposed methodology consisted of two steps: Identification of dropouts, followed by classification of individual data gaps as either occlusions or dropouts. Currently, all code for the data gap classification methodology was written in C/C++ for efficient implementation. A flow chart representing the proposed data gap classification methodology is presented in Figure 2.
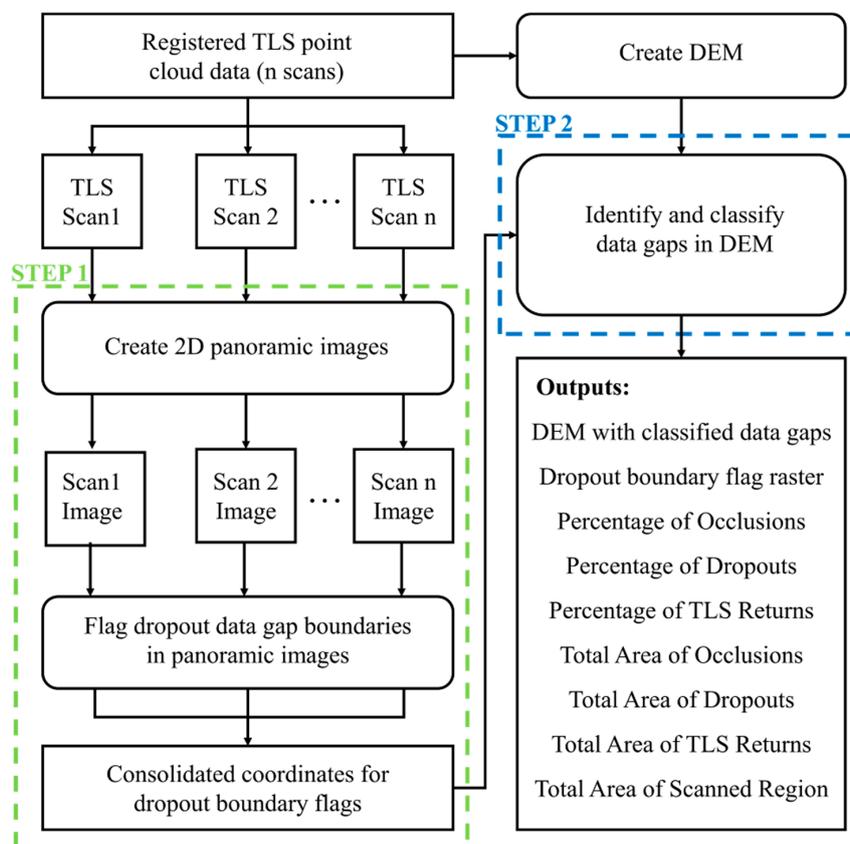

Figure 2. General flow chart for the proposed data gap classification methodology.

This data gap classification methodology required the scan data to be organized based on the data acquisition pattern, meaning the gridded structure/order in which the point returns were collected by the scanner had to be preserved in the individual registered scans. For this study, the text-based PTX format [14] was used, which supported the grid structure. Alternatively, the ASTM E57 [15] format can also preserve this information. The vertical scan lines could be reconstructed as a 2D panoramic image (Figure 3) that represents the point of view (POV) from the scanner origin where each pixel in the image represents a point return. The VZ-400 (RIEGL Laser Measurement Systems GmbH, Horn, Austria) TLS instrument used for this study collected points in a spherical coordinate system. With each pixel representing a point return, pixels in the TLS 2D image also had associated XYZ coordinates in the chosen reference frame. Pixels of the image in Figure 4 colored in shades of gray represent point cloud data colored by intensity and the black pixels represent data gaps. With an individual scan represented as a panoramic image, 2D image processing could be utilized to identify features within the scan data [16–18].
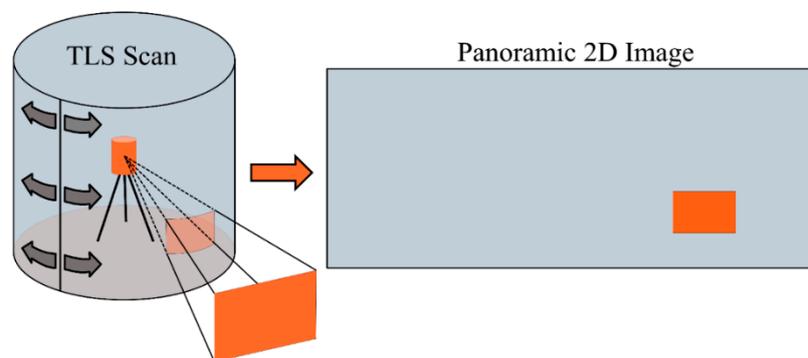

Figure 3. Conceptual depiction of projecting TLS data from native cylindrical coordinate system to panoramic 2D image.

Figure 4. Example of a TLS 2D panoramic image, colored by intensity (grayscale). Black pixels indicate no laser pulse returns and are present in locations with standing water (specular reflection), regions with low reflectivity, and locations beyond the maximum measurement range of the scanner (e.g., the sky and distant horizon).

2.1.1. Identify Dropout Boundaries—Step 1

Step 1 utilized the TLS 2D image based on the acquisition pattern. Because the TLS image generated for each scan represented the scanner's POV, data gaps caused by occlusions were not visible in the image. However, dropout data gaps were visible, and therefore the non-occluded portions of the dropout boundaries could be identified.

Prior to flagging dropout boundaries in the individual TLS images, vertical passes through the imagery were performed to minimize the flagging of data gaps as dropouts along the near and far boundaries of the scanned scene, which corresponded to the bottom and top of the TLS image,

respectively. The vertical passes were performed by iterating through the top and bottom rows of the image and for each column, moving both in a top-down and bottom-up fashion, tagging each data gap pixel until a valid return pixel was reached. The result of the vertical passes is indicated by the green regions in Figure 5.
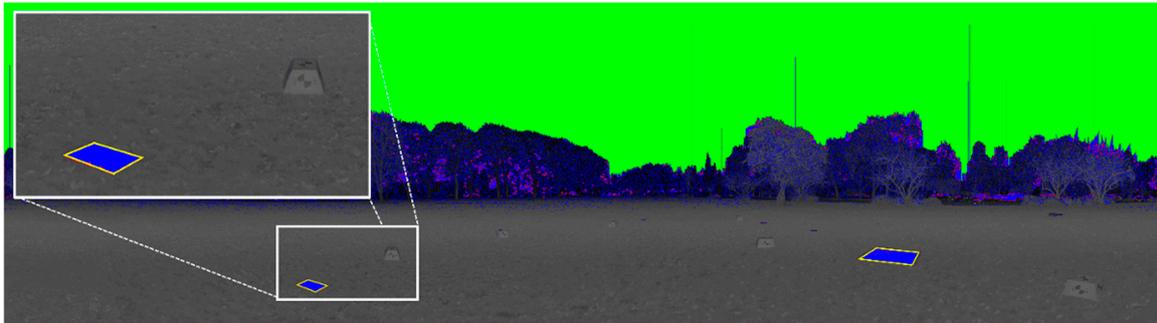

Figure 5. Example of a TLS 2D image with data gaps (blue), flagged dropout boundaries (yellow), and top and bottom vertical scan (green).

These green pixels were ignored during subsequent dropout boundary flagging, which reduced the amount of total flagged dropout pixels by minimizing the quantity of flags generated from long range-derived dropouts and those caused by the laser grazing distant topography at large incidence angles. As for the bottom-up passes, ignoring these pixels ensured that the scanner-based occlusion found beneath a given scan position was not surrounded by flags and subsequently misclassified as a dropout. If a dropout-generating surface existed outside the range limit of a given scan position, an additional scan position that was closer to said surface would be required to properly identify the resulting data gap as a dropout. The assumption was that the scanner would normally be setup over dry land. In a situation where the scanner was setup near standing water, intervention in the data gap classification process may be needed to account for the transition from field of view (FOV) occlusion to dropout. This is further discussed for the Rabbit Rock Study site.

Pixels that lay on the boundary of dropouts (Figure 5) were identified using a 3 × 3 pixel roving window (modified for image boundary pixels). For a given pixel in the TLS image, all eight neighbors were checked to see if any were no-data pixels (data gap). In an effort to omit small dropouts and focus on large extensive data gaps, the following neighboring pixel threshold was implemented: If greater than half (4) of the neighboring pixels had no-data, the current pixel was flagged as a dropout boundary point. Following the identification of the dropout pixels, the XYZ coordinates for each flagged pixel were exported to a text file.

2.1.2. Classification of Data Gaps—Step 2

The dropout boundary flag coordinates generated for each scan in Step 1 were compared to the DEM raster generated using all available TLS data. To ensure the dropout boundary flags properly aligned with the DEM, the DEM had to be in the same coordinate reference frame as the registered and georeferenced TLS source data. Figure 6 provides a visual progression of the data gap classification methodology, which is further discussed below.

To begin, a unique identification number (ID) had to be assigned to each data gap in the DEM. Insignificant data gaps (e.g., single pixel gaps) were omitted from the ID assignment process via a neighbor check. If a given no-data cell did not have at least four neighbors that were also no-data cells, it was not further analyzed. This second wave of omitting small data gaps from the classification processes reduced the occurrence of classifying small data gaps as dropouts when they were in close proximity to dropout boundary flags. DEM pixels that were found to be part of an extensive data gap were grouped together and assigned an ID using a two-pass connected components algorithm [19,20]. The first pass iterated through the grid from left to right and top to bottom, examining the four

neighboring DEM pixels found to the left and above the current pixel. Based on the following conditions, a label ID for each pixel was generated: If all four neighboring pixels had the initial value of zero, a new label was assigned to the current pixel; if only one neighbor had a non-zero label, its label was assigned to the current pixel; lastly, if neighboring pixels had different labels, one of them was assigned to the current pixel and the equivalency with the other labels was recorded. Following the assignment of labels, the occurrence of any adjacent cells with differing labels (equivalencies) was recorded in a lookup table for later use (second pass). The first pass results in all pixels receiving a label; however, contiguous regions may have still contained multiple label IDs. Following a convention that favored the minimum label for a region, a second pass was conducted in which the lookup table was used to consolidate label IDs. The resulting connected components raster represented each contiguous cluster of cells with a unique ID that could be queried for later processing and analysis.
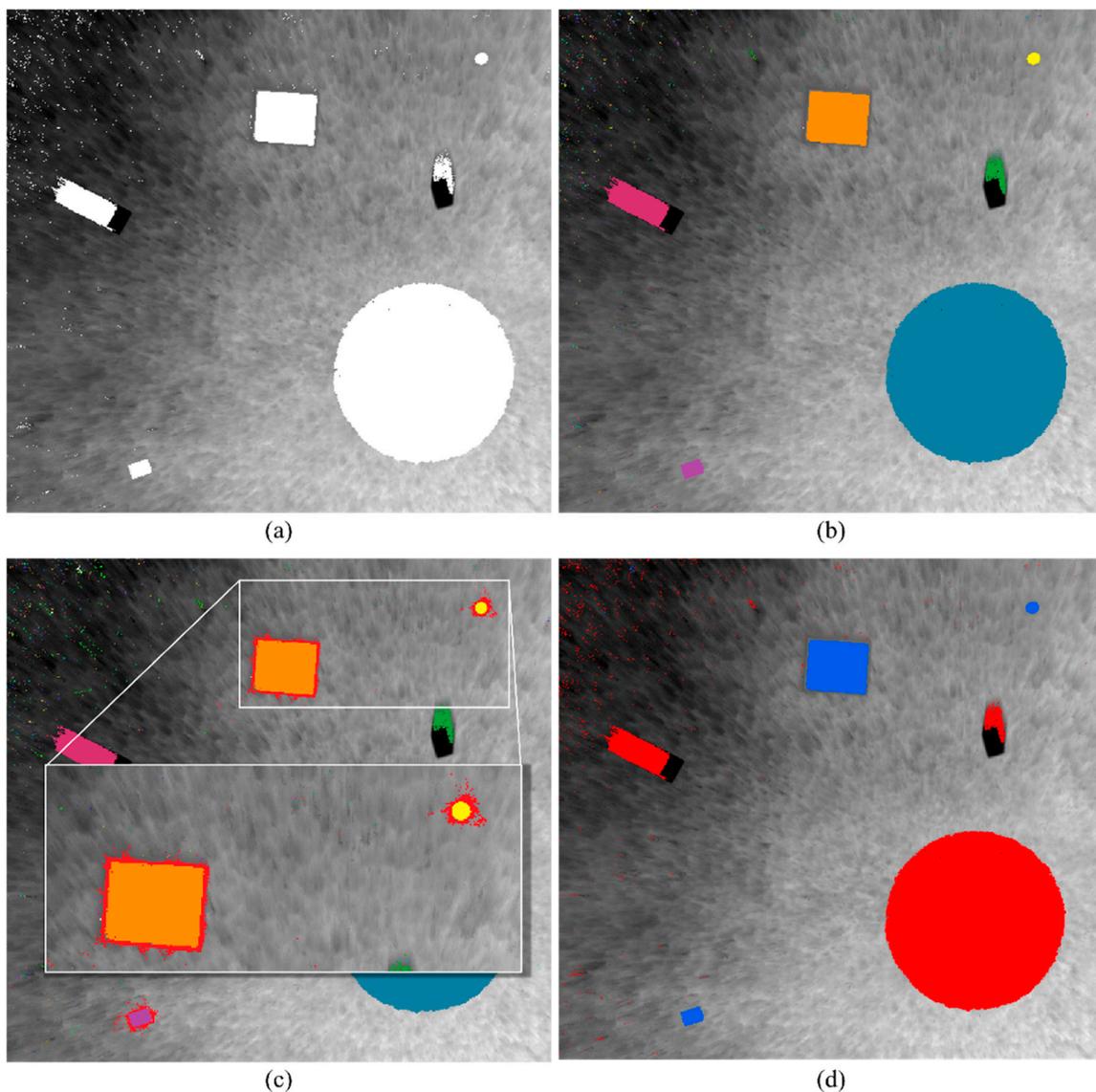

Figure 6. Demonstration of the TLS digital elevation model (DEM) data gap classification methodology. The DEM is colored with a grayscale color ramp (black is higher elevation). (**a**) Unclassified data gaps in DEM identified in white. (**b**) Individual data gaps are assigned unique identification number (ID). Each unclassified data gap is assigned a unique random color for visualization purposes. (**c**) The dropout boundary flag raster (red pixels surrounding some of the data gaps) is introduced. (**d**) The dropout boundary raster is used to classify data gaps as either occlusions (red) or dropouts (blue).

Using the unique IDs, each data gap was spatially compared against a dropout boundary raster (Figure 6c). The dropout boundary raster is a binary raster where cells assigned a "1" coincide with a location where a dropout boundary flag exists. Creating a dropout boundary raster from the individual flag coordinates reduced the quantity of flag data since numerous flags were likely to exist within a cell based on the selected DEM resolution and the observation of the boundary from multiple scan positions. The algorithm then iterates through the dropout boundary raster and classifies data gaps that are surrounded by flags as dropouts. To minimize the occurrence of large occlusions being misclassified based on the presence of a small quantity of dropout boundary flags, at least 10 dropout flags had to be associated with a given data gap to result in a dropout classification. The dropout flag minimum quantity parameter could be adjusted by the user to account for differences in site conditions and TLS resolutions. Following classification of dropout gaps, all remaining data gaps were classified as occlusions.

The last step of the classification methodology addressed a situation where a dropout connected with the occlusion formed beneath a TLS scanner position. For example, this occurs if a TLS scan position was placed adjacent to a puddle of water. In the case where a scanner-based occlusion was not filled in from another POV, the occlusion and the connected/intersecting dropout would be fused together by the connected components process and the data gap would be classified as a dropout. Under the assumption that a TLS instrument would likely not be setup in the center of a pool of water, we included functionality that reclassified any dropouts beneath a given scan position as an occlusion. The area beneath a scan position analyzed for reclassification was determined based on the field of view (FOV) of the specific TLS instrument used and an assumed scanner height (in this case, ~1.8 m). Areas outside of the scanner-based occlusion maintained their original classification.

2.2. Validation

A controlled test site was used for validation of the data gap classification methodology. The controlled test site was a rectangular area established in a flat, grassy field, measuring approximately 15×20 m (Figure 7). Six cardboard boxes and six shallow receptacles filled with water were placed in the scanned area to generate occlusion and dropout data gaps, respectively. For the purposes of this study, TLS data were acquired from five scan positions (Figure 8) at an angular resolution of 0.02°. The TLS data were used to create two separate DEMs of the test site using Bin 'N' Grid software [21]. Bin 'N' Grid is a point cloud rasterization tool that allows for the selection of a spatial resolution and the method by which the points within a given cell are used to determine an elevation. The mean elevation of points within a given cell was used to calculate the final DEM pixel elevation. The first DEM only used data from the one scan position located in the center of the site, and the second DEM used data from four scan positions located at each corner of the site. A preliminary registration (initial orientation) of the four scan positions was completed using the black and white paper targets attached to the faces of the cardboard boxes. A final registration was performed using a cloud-to-cloud registration technique implemented in the PointReg v.3 software [22]. Creating a DEM using only the center scan position ensured the DEM would contain extensive occlusion and dropout data gaps that the classification methodology would have to differentiate. Being much more complete, the second DEM did not contain large occlusion gaps, but it offered a second opportunity to validate if the classification methodology was properly identifying the dropout gaps in the scene. TLS data collected at the test site served to validate results of the data gap classification methodology because the location of occlusion and dropout data gaps were known. In addition, the surface area of pooled water in the scanned scene was calculated from inner dimension measurements of the water receptacles, which had well-defined shapes.

Visual inspection of the data gap classification results for the test site DEMs (Figure 9) indicate all extensive occlusion and dropout data gaps were properly classified. Details of the data gap classification results for the test site are included in Table 1. Areas in Figure 9 colored in red represent occlusions and areas in blue represent dropouts, which in this case also represent pooled water.
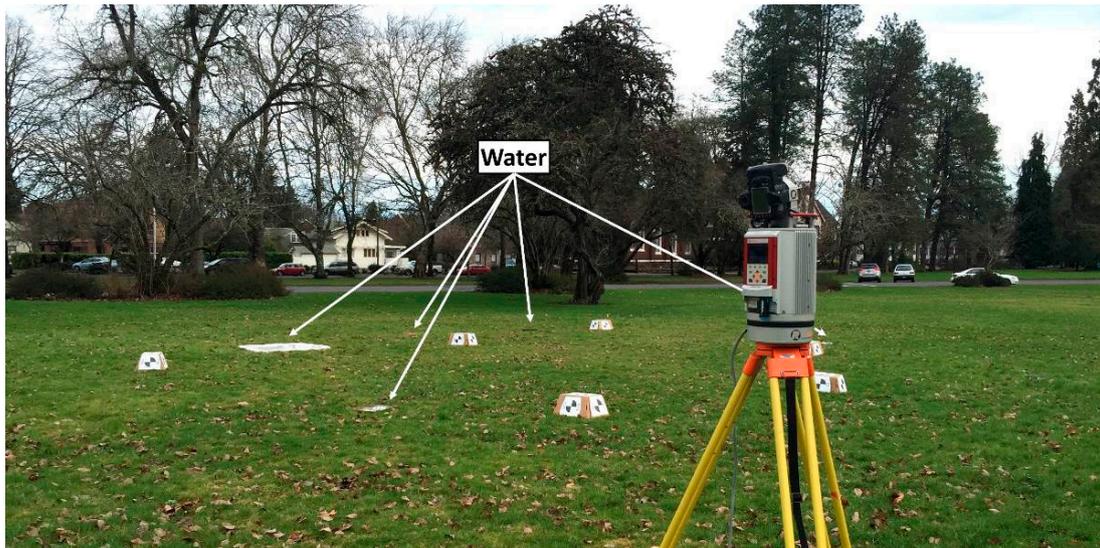
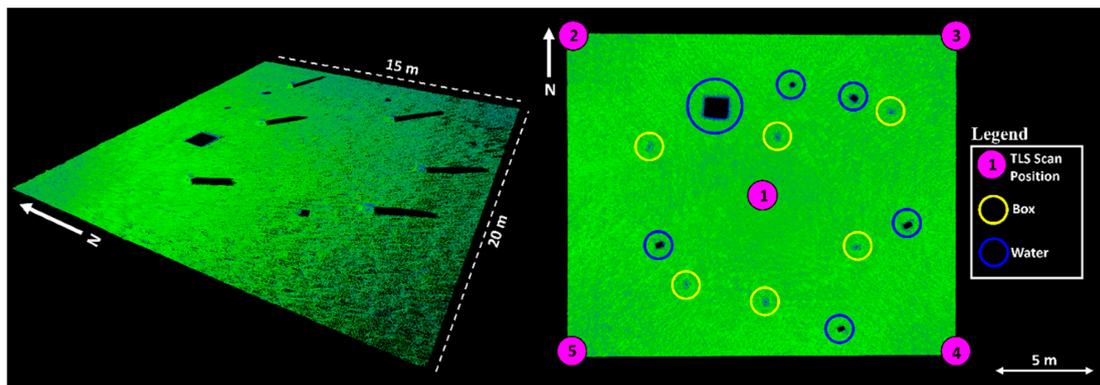**Figure 7.** Photograph of the test site with the Riegl VZ-400 TLS, cardboard boxes, and water receptacles.

Figure 8. Overview of the test site (**left**) and layout map indicating scan locations and data gap sources (**right**).
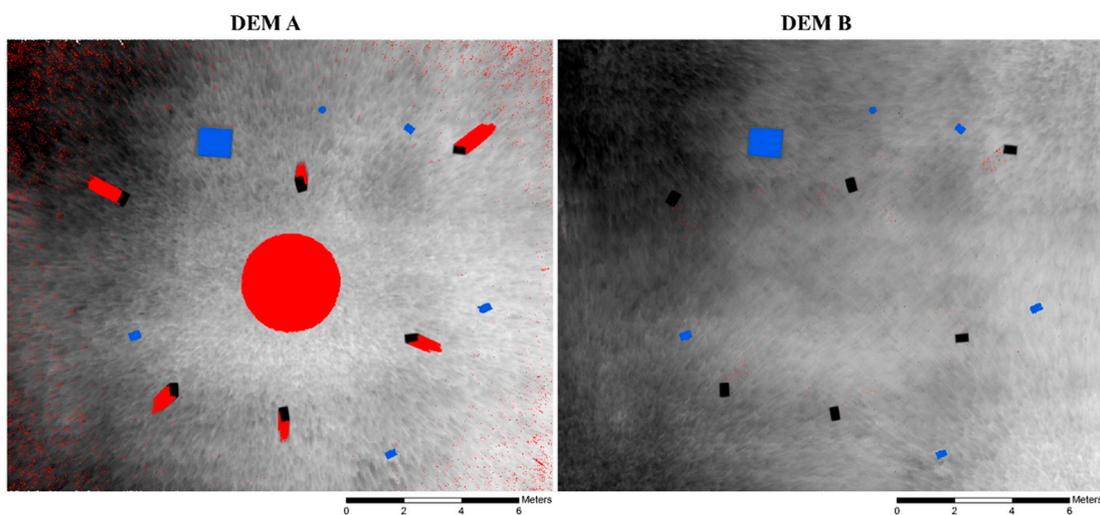

Figure 9. Data gap classification results for test site DEM A and DEM B. The DEMs are colored using a grayscale color ramp (black is higher elevation). Areas in red represent occlusions and dropouts are colored blue. Locations of the cardboard boxes are colored black.

Table 1. Results of data gap classification for test site DEMs A and B.

DEM	Occlusions (%)	Dropouts (%)	Returns (%)	Occlusion Area (m^2)	Dropout Area (m^2)	Return Area (m^2)	Total Area (m^2)
DEM A	5.04	0.54	94.42	15.1	1.6	283.8	300.6
DEM B	0.07	0.51	99.42	0.2	1.5	298.8	300.6

For both the one-scan (DEM A) and four-scan (DEM B) position scenarios, the location of the larger data gaps was as expected. For DEM A, relatively large occlusions occurred behind each of the cardboard boxes where a laser pulse shadow was cast, and beneath the scan position, which was outside the FOV of the TLS instrument. Additional small occlusions were present throughout DEM A that were attributed to laser pulse shadows cast by blades of grass and other small ground cover vegetation. No extensive occlusions were observed in DEM B, but small occlusions were observed throughout the DEM similar to those found in the one-scan position model. The lack of major occlusions in DEM B was attributed to the acquisition of TLS data from four scan positions with different points of view; what was not seen from one scan position was filled in by another, including the large occlusion in the center of DEM A. A decrease in the percentage of occlusions of ~5% was observed when using four scan positions instead of a single scan to capture the test site (Table 1). This was not a dramatic difference; however, the obstacles (cardboard boxes) creating the occlusions in the test site were relatively small. Common obstacles such as trees, cars, and buildings can cause very extensive occlusions, which can occupy the majority of a given scanned scene if additional scans from different positions are not performed. In addition, a second scan from a new position would ordinarily be done to fill in the occlusions below the first scan position.

The total area of dropouts identified for DEM A was 1.62 m^2, which agreed well with the precisely measured area of pooled water at the test site of 1.66 m^2. For test site DEM A, the percent difference was ~3%; however, for DEM B it increased to ~9%. Further investigation of this discrepancy included a difference comparison of the classification results for DEMs A and B. Individual pixels that changed from being a dropout in DEM A to a return in DEM B were identified and found to be predominantly located along the boundary of all pooled water features. The observed change in classification along the boundary of dropouts was attributed to minor misalignment (registration error) amongst the four scans used to generate DEM B. Subtle shifts in the TLS data stemming from registration error would always result in the boundary of a dropout to creep into the data gap, thereby causing a decrease in dropout area. The TLS data used to generate DEM A did not require registration because only one scan position was used. The total area of dropout boundary pixels that changed from a dropout to a return in DEM B was 0.11 m^2. When added to the original total dropout area for DEM B of 1.52 m^2, a dropout area of 1.63 m^2 could be estimated for a DEM with minimal to no registration error. All of the artificial pooled water sources were correctly classified as dropouts in both of the test site DEMs. In addition, small dropouts were identified throughout the DEMs that were likely valid and attributed to varied reflective/absorption conditions in the grass-covered ground surface.

3. Rabbit Rock Study Site

Following validation, the proposed data gap classification methodology was performed on TLS-derived DEMs generated for a natural environment known as Rabbit Rock, a rocky intertidal site located on the Oregon Coast. The Rabbit Rock site served as an example of applying the data gap classification to a real-world site in support of ecological research where the presence of pooled water influences habitats. The Rabbit Rock site is a complex, rocky intertidal environment (Figure 10) located along the central Oregon Coast, approximately 3.5 km north of Depoe Bay, OR along Hwy. 101. TLS data were collected at this location on two separate occasions during very low (minus) tides to model and identify the foraging habitat for the black oystercatcher (*Haematopus bachmani*), a rocky-intertidal obligate shorebird [23]. TLS scan positions 1–14 were acquired on 18 May 2011 and scans 15–21 were acquired on 1 June 2011 (Figure 11). The second set of scans were acquired to fill in

areas of the site inaccessible during the May survey because of higher tidal conditions. All TLS scans were acquired at angular resolutions of 0.03 or 0.05 degrees. Registration and geo-referencing of the point cloud data was performed with a constrained cloud-to-cloud registration technique implemented in PointReg v3 [22] based on GNSS coordinates, sensor inclination, and an estimated yaw angle for the TLS instrument at each scan position. Post-processed GNSS coordinates for the individual scan positions were generated using the rapid–static processing available through the National Geodetic Survey's Online Positioning User Service (OPUS-RS). Two 10-cm TLS-derived DEMs were generated for the Rabbit Rock site, one using only TLS scans 1–14 (DEM RR1) and the second using data from all 21 scan positions (DEM RR2) (Figure 12).

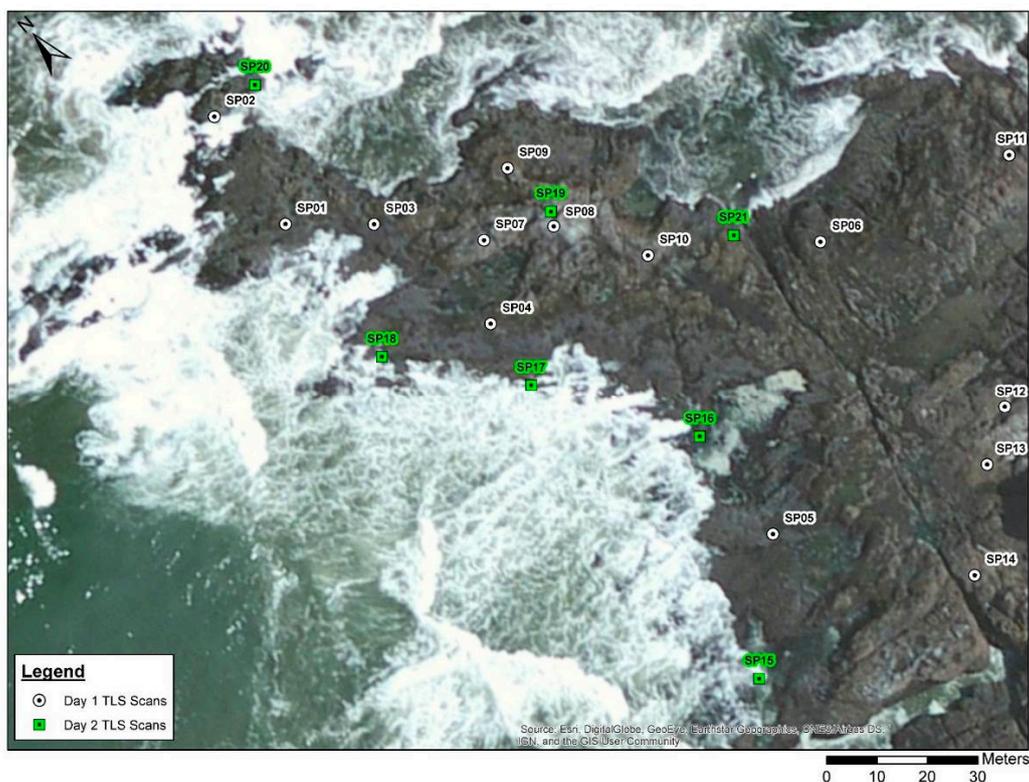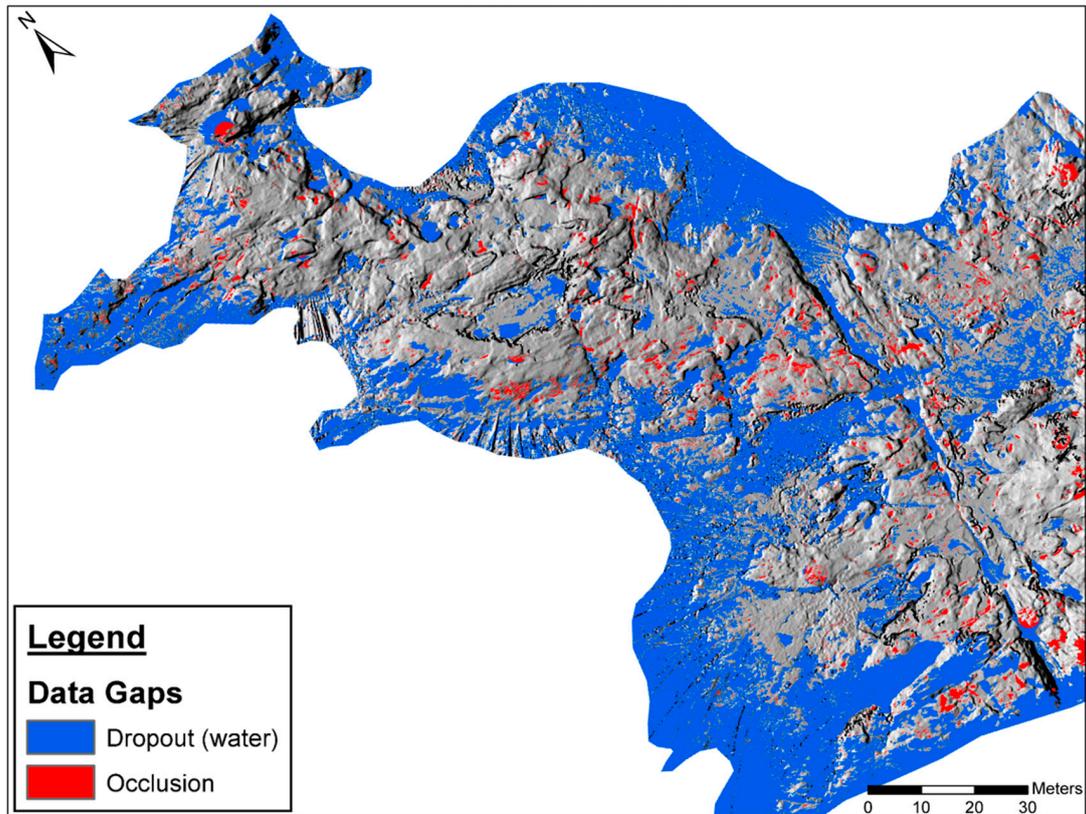**Figure 10.** An image at the Rabbit Rock site showing the undulating rock and pooled water conditions.

Figure 11. Overview map of Rabbit Rock site with locations of TLS scan positions.
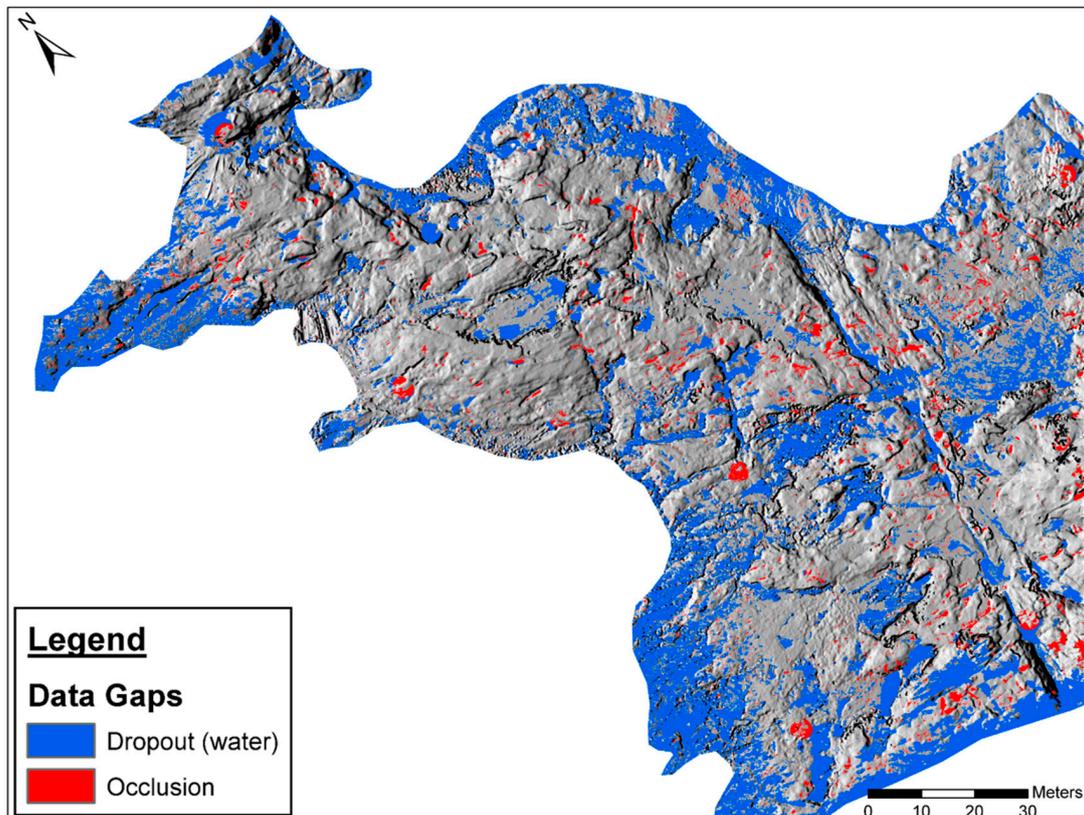
DEM RR1

DEM RR2

Figure 12. Clipped data classification results for DEMs RR1 (top) and RR2 (bottom).

Given the presence of undulating rock and numerous pools of water at the Rabbit Rock site (Figure 10), there were many opportunities for both occlusions and dropouts to exist in the scanned scene. When examining the Rabbit Rock DEMs with unclassified data gaps, two questions arose: How well was the site captured (TLS survey quality) and what regions of the DEM were occupied by pooled water? The presence of pooled water within the Rabbit Rock site was important for identifying and modeling the shorebird foraging habitat. In the unclassified DEM, occlusions and dropouts caused by pooled water were indistinguishable. To assess the survey quality and identify regions of the Rabbit Rock site occupied by pooled water, the proposed data gap classification methodology was utilized. To avoid classifying erratic dropouts associated with the dynamic ocean water surrounding the site, a site boundary was used to focus on the region targeted in the TLS topographic survey.

Results and Discussion

Results of the data gap classification (Figure 12 and Table 2) indicated that for DEM RR1, ~61% of the site was occupied by elevation data (TLS returns), ~36% was occupied by dropouts, but only ~2.6% was attributed to occlusions. For DEM RR2, ~72% of the site was occupied by TLS returns, ~25% was occupied by dropouts, and we saw a similar but slightly lower relative percentage of occlusions at around ~2.4%.

Table 2. Results of data gap classification for the cropped Rabbit Rock DEMs 1 and 2.

DEM	Occlusions (%)	Dropouts (%)	Returns (%)	Occlusion Area (m²)	Dropout Area (m²)	Return Area (m²)	Total Area (m²)
DEM RR1	2.56	36.19	61.25	436.3	6166.2	10,436.5	17,039.1
DEM RR2	2.39	25.48	72.13	406.4	4341.8	12,291.0	17,039.1

The classification result rasters for DEMs RR1 and RR2 were differenced in ArcGIS Pro v.2.6.1 [24] to identify any changes in pixel classification that occurred. Results of this comparison indicated ~88% of the pixels experienced no change in classification and ~10% of pixels classified as a data gap in DEM RR1 became a return in DEM RR2—~93% of which stemmed from the occurrence of a dropout becoming a return. Less than 1% of pixels underwent a change in classification from a return to a data gap or data gap switching (e.g., a dropout becoming an occlusion and vice versa). The reclassification of an RR1 return as a data gap in RR2 should not occur; however, only 0.05% of pixels experienced this change. Further examination of the comparison results indicated approximately 1727 m² of dropout pixels changed to a return pixel in DEM RR2, which represented ~93% of the total dropout area reduction presented in Table 2. Based on the observed decrease in the identified dropouts and the minimal change in percent occlusions, from DEM RR1 to RR2, the majority of the ~11% increase in returns was attributed to filling in areas that were obscured by high water conditions in the previous TLS survey. The similar percent occlusions observed for DEMs RR1 and RR2 may have been attributed to the combination of consistent TLS surveying techniques/scanner placements and the repeated undulating nature of the Rabbit Rock terrain.

There were a few locations where dropouts were classified near a TLS scan location. In these regions, the photographs taken from the scan position could be used to perform a qualitative validation of the results. The TLS-based imagery for scan positions SP15 and SP16 are presented in Figure 13. The pools of water visible in the imagery corroborated the classification results of extensive dropouts surrounding the scan positions. Circular occlusions were observed beneath SP15 and SP16 in Figure 13 due to other scan positions' inability to fill in these areas. It is important to note that these scanner-based occlusions were adjacent to pooled water, which made it difficult to determine where exactly the pooled water stopped and the occlusion began. In this case, we decided to be conservative with respect to judging survey quality; therefore, we assumed that all data gaps within a certain radius of a given scan position should be classified as occlusions. If we wanted to ensure that we were capturing all the

potential water pools, we could have changed the algorithm to classify the entire, merged data gap as a dropout.
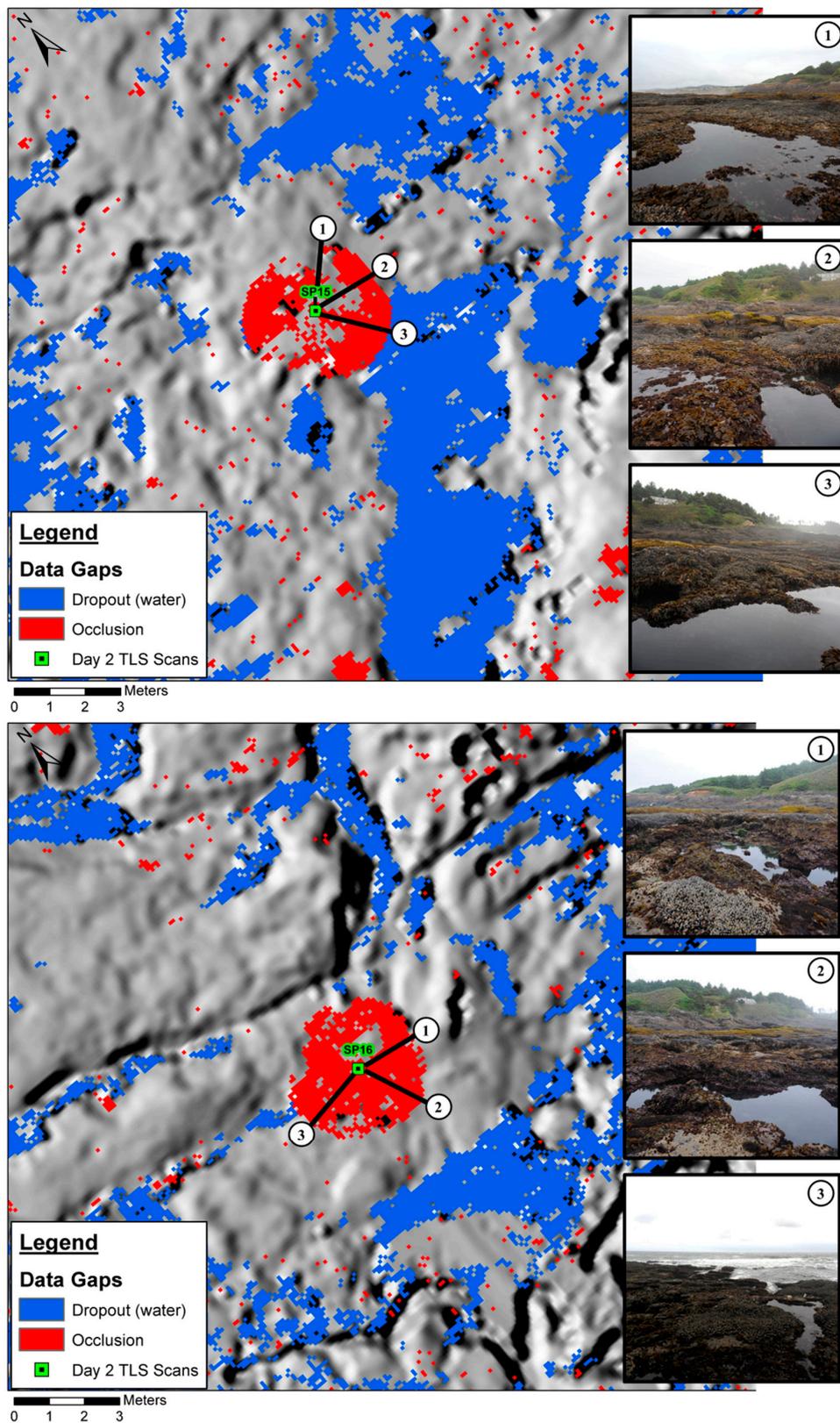

Figure 13. Comparison of results with co-acquired TLS scanner-based imagery for scan positions SP15 (top) and SP16 (bottom).

A quantitative validation of the Rabbit Rock classification results was performed using the near-infrared (NIR) channel from National Agriculture Imagery Program (NAIP) 1-m spatial resolution aerial imagery collected in 2011, 2014, and 2016. An unsupervised classification of each image was performed with twenty classes using the Iso Cluster Unsupervised Classification tool in ArcGIS Pro v.2.6.1 [24]. Each image was individually analyzed to identify which of the resulting classes were associated with extensive pools of water. Validated water classes were then used to generate a binary raster identifying "water" and "non-water" pixels for each NAIP image. Next, the three binary rasters were added together using the raster calculator functionality in ArcGIS Pro. The resulting raster had pixel integer values ranging from 0–3, where the value of a given pixel represented in how many NAIP images that location was classified as water. For example, a pixel value of "2" corresponded to a location that was classified as water in two out of the three NAIP images. To target persistent water pools that existed at the Rabbit Rock site, the validation analysis focused on regions that were classified as water in all three NAIP images (i.e., pixel value of 3). Groups of pixels with a value of "3" were converted to polygon features and then used to generate a set of 100 random points within these locations. The 100 points were then cross referenced with the DEM RR2 data gap classification raster to identify what percentage of the locations were correctly classified. For the 100 random points sampled form persistent water pools identified using the three NAIP images, ~89% were correctly classified as water in DEM RR2, ~3% were classified as occlusions, and ~8% were identified as TLS return pixels with elevation data. While we could not be absolutely certain that the water pools identified in the NAIP imagery represented the site conditions during the TLS surveys, the consistency observed in this validation indicated that the proposed data gap classification methodology was capable of identifying extensive water pools that were present in aerial imagery collected over a five-year period.

4. Conclusions

The proposed data gap classification methodology differentiated between occlusion and dropouts in a TLS-derived DEM using structured TLS point cloud data (PTX), and the associated DEM. The test site results showed a high degree of correct classification of occlusions and dropout-based data gaps and identification of a similar surface area of pooled water present in the scanned scene. The results for the Rabbit Rock site analysis indicated the identified dropouts correlated well with the presence of water, and the quality of the Rabbit Rock TLS survey was high given the low percent of occlusions.

For the test and Rabbit Rock sites, the extensive dropouts could be attributed to pooled water present in the scanned scene. If this classification methodology were applied to a dataset that included other highly reflective objects (e.g., glass windowpanes), dropouts could not be solely attributed to the presence of water; however, they would still be separated from the occlusions, which is an important distinction to make. For assessment of TLS survey quality, data gaps due to dropouts had to be identified and removed before the relative percentage of data gaps due to occlusions was determined. The proposed data gap classification methodology enabled us to make this required distinction.

In a complex environment such as Rabbit Rock, we could assume the primary source of dropouts was attributed to water. Thus, TLS offers tremendous potential for ecological studies in the rocky intertidal ecosystem. Due to the nature of this highly limited (spatially) ecosystem, TLS-derived DEMs may provide the foundation for scale-appropriate habitat models and simulations. Previous work using TLS-derived DEMs for modeling the shorebird foraging habitat demonstrated substantial capability [23]. However, a missing key attribute that influenced the development of habitat models was the accurate identification of submergent areas during low tides (i.e., tidepools). Because the rocky intertidal ecosystem was very spatially limited, submergent areas (ostensibly dropouts) may have comprised a considerable proportion of the total area of interest (see Table 2) and may become important for subsequent habitat assessments and modeling. Also of importance was the identification of water area (tidepools) boundaries. This interface was a region of many important interactions between the terrestrial and aquatic components of the intertidal ecosystem. For example, the black oystercatcher foraging habitat model developed by Hollenbeck et al. (2014) lacked the ability to identify local

regions of key prey items (limpets) that congregate at tidepool boundaries. Consequently, the ability to differentiate submergent areas from emergent areas that are exposed to the terrestrial component of the intertidal ecosystem is paramount for scale-appropriate habitat analyses and TLS-derived DEMs, processed to differentiate data dropouts and occlusions, and may hold significant promise for intertidal research. Previously available methods for delineating these locales have often involved very intensive field-survey methods [13] or digitizing the DEM or point cloud, which requires substantial effort and is often not feasible in many ecological studies.

Author Contributions: Conceptualization, Matthew S. O'Banion, Michael J. Olsen, and Jeff P. Hollenbeck; methodology, Matthew S. O'Banion, Michael J. Olsen, and Jeff P. Hollenbeck; validation, Matthew S. O'Banion, Michael J. Olsen, and William C. Wright; investigation, Matthew S. O'Banion and Michael J. Olsen; data curation, Matthew S. O'Banion and Michael J. Olsen; writing—original draft preparation, Matthew S. O'Banion; writing—review and editing, Michael J. Olsen, Jeff P. Hollenbeck, and William C. Wright; visualization, Matthew S. O'Banion and William C. Wright; supervision, Michael J. Olsen; project administration, Matthew S. O'Banion and Michael J. Olsen; funding acquisition, Michael J. Olsen and Jeff P. Hollenbeck. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. LiDAR Glossary | AGRG. Available online: http://agrg.cogs.nscc.ca/resources/lidar-glossary/ (accessed on 11 June 2017).
2. Telling, J.; Lyda, A.; Hartzell, P.; Glennie, C. Review of Earth science research using terrestrial laser scanning. *Earth-Sci. Rev.* **2017**, *169*, 35–68. [CrossRef]
3. Höfle, B.; Griesbaum, L.; Forbriger, M. GIS-based detection of gullies in terrestrial LiDAR data of the Cerro Llamoca Peatland (Peru). *Remote Sens.* **2013**, *5*, 5851–5870. [CrossRef]
4. Olsen, M.J.; Wartman, J.; McAlister, M.; Mahmoudabadi, H.; O'Banion, M.S.; Dunham, L.; Cunningham, K. To fill or not to fill: Sensitivity analysis of the influence of resolution and hole filling on point cloud surface modeling and individual rockfall event detection. *Remote Sens.* **2015**, *7*, 12103–12134. [CrossRef]
5. Deshpande, S.S.; Yilmaz, A. A semi-automated method to create a lidar-based hydro-flattened DEM. *Int. J. Remote Sens.* **2017**, *38*, 1365–1387. [CrossRef]
6. Höfle, B.; Vetter, M.; Pfeifer, N.; Mandlburger, G.; Stötter, J. Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surf. Process. Landf.* **2009**, *34*, 1635–1649. [CrossRef]
7. Smeeckaert, J.; Mallet, C.; David, N.; Chehata, N.; Ferraz, A. Large-scale classification of water areas using airborne topographic lidar data. *Remote Sens. Environ.* **2013**, *138*, 134–148. [CrossRef]
8. Wei, L.-X.; Yang, B.; Jiang, J.-P.; Cao, G.-Z. Adaptive algorithm for classifying LiDAR data into water and land points by multifeature statistics. *J. Appl. Remote Sens.* **2016**, *10*, 045020. [CrossRef]
9. O'Banion, M.S.; Olsen, M.J. Efficient planning and acquisition of terrestrial laser scanning–derived digital elevation models: Proof of concept study. *J. Survey. Eng.* **2018**, *145*, 06018003. [CrossRef]
10. Underwood, A.J. Experimental ecology of rocky intertidal habitats: What are we learning? *J. Exp. Mar. Biol. Ecol.* **2000**, *250*, 51–76. [CrossRef]
11. Beck, M.W. Comparison of the measurement and effects of habitat structure on gastropods in rocky intertidal and mangrove habitats. *Mar. Ecol. Prog. Ser.* **1998**, *169*, 165–178. [CrossRef]
12. Meager, J.J.; Schlacher, T.A.; Green, M. Topographic complexity and landscape temperature patterns create a dynamic habitat structure on a rocky intertidal shore. *Mar. Ecol. Prog. Ser.* **2011**, *428*, 1–12. [CrossRef]
13. Schoch, G.C.; Dethier, M.N. Scaling up: The statistical linkage between organismal abundance and geomorphology on rocky intertidal shorelines. *J. Exp. Mar. Biol. Ecol.* **1996**, *201*, 37–72. [CrossRef]

14. Leica Geosystems. *Leica Cyclone 8.0 Help*; Leica Geosystems: St. Gallen, Switzerland, 2014.

15. Huber, D. The ASTM E57 file format for 3D imaging DATA exchange. In Proceedings of the SPIE Electronics Imaging Science and Technology Conference (IS&T), 3D Imaging Metrology, Pittsburgh, PA, USA, 22–27 January 2011; Volume 7864A.

16. Barnea, S.; Filin, S. Registration of terrestrial laser scans via image based features. In Proceedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 Part 3, Espoo, Finland, 12–14 September 2007; pp. 32–37.

17. Olsen, M.J.; Ponto, K.; Kimball, J.; Seracini, M.; Kuester, F. 2D open-source editing techniques for 3D laser scans. In Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, CAA, Granada, Spain, 6–9 April 2010; pp. 47–50.

18. Mahmoudabadi, H.; Olsen, M.J.; Todorovic, S. Efficient terrestrial laser scan segmentation exploiting data structure. *ISPRS J. Photogram. Remote Sens.* **2016**, *119*, 135–150. [CrossRef]

19. Dillencourt, M.B.; Samet, H.; Tamminen, M. A general approach to connected-component labeling for arbitrary image representations. *J. ACM* **1992**, *39*, 253–280. [CrossRef]

20. Fisher, R.; Perkins, S.; Walker, A.; Wolfart, E. Connected Components Labeling. Available online: http://homepages.inf.ed.ac.uk/rbf/HIPR2/label.htm (accessed on 23 October 2020).

21. Olsen, M. Bin 'N' Grid: A Simple Program for Statistical Filtering of Point Cloud Data. *LIDAR Magazine*. 2011. Available online: https://lidarmag.com/2011/05/29/bin-n-grid-a-simple-program-for-statistical-filtering-of-point-cloud-data/ (accessed on 23 October 2020).

22. Olsen, M.; Johnstone, E.; Kuester, F.; Driscoll, N.; Ashford, S.A. New automated point-cloud alignment for ground-based light detection and ranging data of long coastal sections. *J. Survey. Eng.* **2011**, *137*, 14–25. [CrossRef]

23. Hollenbeck, J.P.; Olsen, M.J.; Haig, S.M. Using terrestrial laser scanning to support ecological research in the rocky intertidal zone. *J. Coast Conserv.* **2014**, *18*, 701–714. [CrossRef]

24. ESRI. *ArcGIS Pro*; v.2.6.1; ESRI: Redlands, CA, USA, 2020.