MACHINE LEARNING APPROACH FOR TENAGA NASIONAL BERHAD (TNB) OVERHEAD POWERLINE AND ELECTRICITY POLE INVENTORY USING MOBILE LASER SCANNING DATA

Mohd Saiful Anuar Mohd Rapheal¹, Anis Farhana², Mohd Radhie Mohd Salleh³, Muhammad Zulkarnain Abd Rahman², Zulkifli Majid³, Ivan Amri Musliman², Ahmad Fikri Abdullah³, Zulkiflee Abd Lati³

¹ Project GIS, Distribution Network, Tenaga Nasional Berhad, Universiti Tenaga Nasional, Kajang, Selangor, Malaysia
² Department of Geoinformation, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor, Malaysia
³ Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400, Serdang, Malaysia

anisfarhana880@gmail.com, mohdradhie.gis@gmail.com, mdzulkarnain@utm.my, zulkiflimajid@utm.my, ivanamri@utm.my, anisfarhana880@gmail.com

© Author(s) 2022. CC BY 4.0 License.

KEY WORDS: Mobile Laser Scanner (MLS), Electricity Assets, Machine Learning, Overhead Powerline, Electricity Poles.

ABSTRACT

Electricity assets recognition and inventory is a fundamental task in the geospatial-based electrical power distribution management. In Malaysia, Tenaga Nasional Berhad (TNB) aims to complete their assets inventory throughout the country by 2022. Previous research has shown that a method for assets detection especially for TNB is still at an early stage, which mainly relied on manual extraction of the assets from different data sources including mobile laser scanner (MLS). This research aims at evaluating a geospatial method based on machine learning to classify the TNB assets using high density MLS data. The MLS data was collected using Riegl VMQ-1 HA scanner and supported by the base station and control points for point cloud registration purpose. In the first stage the point clouds were classified into ground and non-ground objects. The non-ground points were further classified into different landcover types i.e. vegetation, building, and other classes. The points classified as other classes were used for overhead powerline and electricity poles classification using random forest-based Machine Learning (ML) approach in LiDAR 360 software. Based on the classified point clouds, detailed characteristics of electricity poles (i.e. number of poles, height, diameter and inclination from ground) and overhead powerlines (number of cable segments) were estimated. This information was validated using field collected reference data. The results show that the detection accuracy for electricity poles and overhead power line are 65% and 63% respectively. The estimation of length, diameter and height of the spun pole from point clouds has produced Root Mean Square Error (RMSE) value of 0.081cm, 0.263 cm and 0.372 cm respectively. Meanwhile for the concrete pole, the length, diameter and height has been successfully estimated with the value of RMSE of 0.034 cm, 0.029 cm and 0.331 cm respectively. The length of overhead powerline was estimated with 59.02 cm RMSE. In conclusion, the MLS data had show promising results for a semi-automatic detection and characterization of TNB overhead powerlines and poles in the sub-urban area. Such outcome can be used to support the inventory and maintenance process of the TNB assets.

1. INTRODUCTION

LiDAR technology has significantly changed the landscape of land surveying and mapping since the introduction of high processing power of computer and Global Navigation Satellite System (GNSS). The technology capable of delivering detailed and accurate three-dimensional point clouds data to support various terrestrial applications. Nowadays, the newly emerges LiDAR technology has shifted to multi-platforms approach namely UAV-based and mobile laser scanner (MLS). Commonly the MLS system is mounted to a vehicle and backpack, which provides detailed geographic information of terrestrial objects in the form of high-density point clouds. The MLS system works with the support from GNSS, inertial navigation system (INS) and distance measurement indicator (DMI) which georeferenced the point clouds. The MLS system commonly integrates camera sensor to colorize the point clouds.

MLS technology that provides high density point clouds may accurately characterized a complete pole object. In Malaysia, the poles and overhead cables are owned either by local authorities, Telekom Malaysia or TNB. These poles and cables are different in size, materials and functionality. Most of the previous studies have been focusing on extracting light poles in urban environment (Ma et al., 2018). The existing light poles detection on point clouds can be grouped into data driven and knowledge driven approaches. The knowledge-driven method operates based on the facts that poles are commonly located near the road curbs and poles stand vertically on the ground within a certain height (Ma et al., 2018). The knowledge-driven method is further classified into matching-based and rule-based extractions. The matching-based extraction operates by matching the selected portions of point clouds with known object models. Yu et al. (2015) compared the pre-defined 3D model with point clouds based on a pair-wise comparison. Based on the comparison, the local and global dissimilarity is calculated. Cabo et al. (2014) partitioned the point clouds using voxel model, which help in reducing the initial data. Small two-dimensional segments were selected that are compatible with a target pole. The 3D voxel of the detected pole was determined to classify individual poles in the dataset.

On the other hand, the rule-based detection method used geometric and spatial features of poles as rules in the process of
selecting the points. Teo & Chiu (2015) used rule-based approach with geometric features i.e. area of cross-section, position, size, orientation, and shape to remove other objects. Meanwhile, pole parameters (e.g., location, radius, and height) were computed to extract feature and the hierarchical filtering method was employed to remove the unwanted objects. Rodríguez-Cuenca et al. (2015) applied vertical conditions of pole as parts of the rule-based pole extraction process. The rule-based methods are efficient, simple and relatively fast to apply on the point clouds. However, there is not standard on the quantity of rules (Ma et al., 2018). Furthermore, the quality of rules should be carefully determined either from the data or the knowledge of experts. Commonly the rule-based method tends to over segmented the point clouds with numerous rules or under-segmented with small number of rules. The data-driven method requires set of features to be designed and applied to train learning algorithms in detecting the poles. The training sets should be collected for different type of poles and the training process would involve massive labelled point clouds (Ma et al., 2018). Haiyan Guan, Yongtao Yu, Jonathan Li, & Pengfei Liu (2016) developed a contextual visual library of light poles from 50 training sets along a specific stretch of roads. The library was used to detect light poles from the ground filtered point clouds data. Wu et al. (2017) integrated two dimensional (2D) images with point clouds for poles detection. The pole features (i.e. the height, the average height, the standard deviation of height, the estimated volume, the number of pole points, pixel intensity and the number of the super-voxels whose area of the convex hull) were computed and used in the support vector machine (SVM) and random forest classifiers.

Wang et al. (2018) classified MLS point clouds into powerline class using data driven approach over urban and rural environment. They highlighted that the major issue was the close and high vegetation canopy that blocked GNSS signal during data MLS observation process. The GNSS outage has caused 10cm positioning errors for data observed in rural area. Another problem raised by the inhomogeneous point cloud distribution over the powerline for point clouds obtained from the backpack MLS system. This has significantly affected the powerline detection over rural area. This paper aims at evaluating the performance of high-density point clouds produced by MLS system over sub-urban area. This paper aims at evaluating the performance of high-density point clouds produced by MLS system over sub-urban area. The study area is characterized by residential zones which the point clouds include various objects i.e. cars, trees, heavy traffic. This might affect the homogeneity and density of point clouds over overhead powerline and poles. In addition, this may cause problems in the classification and extraction process of these electricity assets.

2. MATERIALS AND METHOD

2.1 Overall Methodology

The methodology can be divided into 4 phases namely, 1) data acquisition, 2) data pre-processing, 3) classification of point clouds, 4) detection of overhead powerlines and electricity poles and 5) validation of results.

Figure 1. Workflow of the research methodology

2.2 Description of Study Area

The main study area for this research is in Taman Sri Pulai, Johor (1°33´N 103°37´E) (peninsular Malaysia) as shown in a map (Figure 2). This study area was focused on the suburban area located in Taman Sri Pulai. Therefore, the data collection was taken from this covered area for the TNB asset which is utility pole and overhead cable. This area was selected as one of the study areas because as it can see the mixed-use or residential area is representing the land cover, hence there is a lot of TNB assets, especially electrical poles and overhead cable as one of the contributions of sources to distribute electric power to users. Moreover, these areas are easy to access because they are located in a road environment which is an important roadway asset. The data acquisition was collected by team field data capture from GPS Lands Sdn Bhd and Topcon Malaysia.

Figure 2. Location of Study Area at Taman Sri Pulai, Johor
2.3 Data Acquisition

The MLS data acquisition is made over sub-urban area of Taman Sri Pulai, Johor. The area consists of residential which is characterized by terrace houses, moderately cover by vegetation. The data collection was done during the off-peak time to reduce traffic obstacle during MLS data collection. The MLS system employed Riegl VMQ-1 HA scanner and supported by the base station and control points for point cloud registration purpose.

2.4 Data Pre-Processing

Data Pre-Processing is one of the most crucial parts in the methodology of flowchart. In this phase, performing 1) Georeferenced data, 2) Land use land cover classification and 3) Feature extraction is using LiDAR360 software. The method used for this research study is Machine Learning (ML) method using random forest algorithm because it aimed at classifying a model that the user declared. The way ML works is based on the parameters setting of height and size of the training model.

2.5 Classification of Point Clouds

Raw LiDAR point clouds consist of individual points that represent the 3D spatial positions of the pulse-reflecting targets of the laser scan present at the time of data capture in the survey area. Details related to the echo return number (in multiple return laser scanners), intensity metrics, RGB image values, and classification codes may be attributed to these coordinate value sets.

2.5.1 Georeferenced Data

Three-dimensional point cloud data files from Mobile LiDAR Scanning data systems include not only X, Y, and Z point coordinates, but additional information such as GPS time, scan angle, or reflectance intensity information for the millions of points found in the point cloud. The geo-referenced or recorded point clouds are formed in a desired coordination system once the trajectory has been produced based on observation of DMI and IMU. In this research paper, the coordinate system that will be used is WGS 84 / UTM zone 48N, EPSG:32648 as it located in Johor, Malaysia.

2.5.2 Land Use Land Cover for Point Cloud Data

Land use land cover for point cloud data is second phase of the methodology where to get the results for extracted TNB assets, land use land cover and ground points filtering should be done first before the feature extraction for poles and power lines were conduct. One of the important factors which affects the accuracy of mobile LiDAR data is a data filtering technique, which is the process of isolating LiDAR point clouds so that only ground point will be use. In other words, ground points and non-ground points must be separated. Besides, the land use land cover was also conducted to minimize the others features of point cloud such as ground, building, and trees were detecting as pole and powerline. Hence, this phase will perform the classification for ground, building and trees, then follow with the pole and powerline. Therefore, the chances to other features to detect pole and powerline is low and it will be eased to clean the data for the next process.

2.6 Detection of Overhead Powerlines and Electricity Poles Using Machine Learning

ML approach we use to classify LiDAR point clouds. LiDAR360 provides a ML approach to point cloud classification. The ML classification tool makes use of a random forest method for determining individual point classifications based on a statistical model of user-defined feature types. To apply the ML approach, the training model must be developed first by defining the pole and powerline. The way ML works are based on the building parameters of maximum building size: 60m and minimum building height: 3m. After the training data is ready, using classify by trained ML model, it can be applied to classify all the point clouds features to detect pole and powerline.

2.6.1 Filtering Data

In this part of the methodology, the software that is used to filter the data is ArcGIS and MATLAB. Both software use the same method of Minimum Bounding Geometry using circle fitting for Spun Pole Low Voltage and rectangle fitting for Concrete Pole. The Minimum Bounding Geometry creates a feature class containing polygons which represent a specified minimum bounding geometry enclosing each input feature or each group of input features. The output polygon features, and their attributes will vary depending on the specified geometry type and grouping choices. For the powerline, the height was filtered based on the aggregation method using ArcGIS software to get the exact height of the pole. Aggregation method is one of the tool resamples that works to multiply the cell resolution of the input raster by the factor specified by the cell factor parameter.

2.7 Validation of Results

From the mobile LiDAR data, accuracy assessment can be made according to the data feature extraction results using ML classification tools. By performing accuracy examination, it can figure out how accurate the mobile LiDAR data is to detect pole or power line extraction. It also can detect the location of the TNB assets, the height of poles, and diameter of poles. All of these must be validated by using a few approaches based on the completeness of data reduction. In this paper, the assessment and comparison of the classification and feature extraction results will be carried out using complete confusion matrix to prove the overall accuracy of the method used. Table 1 shows the example of confusion matrices that will be used for this research paper. The purpose of using this confusion matrices is to evaluate metrics for ML approach or to evaluate the model’s performance. This is considered as the predicted point clouds detected using mobile LiDAR data vs actual point clouds.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Pole</td>
<td>True Negative</td>
</tr>
<tr>
<td>Pole</td>
<td>False Negative</td>
</tr>
</tbody>
</table>

Table 1. Example confusion matrices to evaluate data.

The assessment method to calculate the accuracy is based on the precision and recall. The following formula were used to compute precision, recall and accuracy (Braun, 2020):

\[
\text{Precision} = \frac{TP}{TP + FP} \\
\text{Recall} = \frac{TP}{TP + FN} \\
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]
To perform the accuracy assessment to identify the accuracy of the diameter, circumference, length, height of pole and the number segment of powerline detected in mobile LiDAR data is by computing the Root Mean Square Error (RMSE) (Braun, 2020).

\[
RMSE = \sqrt{\frac{\sum (\text{Predicted} - \text{Actual})^2}{N}}
\]

To know whether the validation of extraction pole and powerline achieve the target of TNB’s requirements, and the accuracy needed by TNB are based on Table 2 that shows TNB’s measurement of pole size according to the requirements of getting the accuracy. The highlighted part is the one that focuses on this study area at Taman Sri Pulai, Johor because it only consists of concrete pole and spun pole low voltage. All the assessments are made based on the measurement that has been fixed.

### Table 2. TNB’s measurement of pole size

<table>
<thead>
<tr>
<th>TNB Diameter Pole</th>
<th>Standard Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>Width</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------</td>
</tr>
<tr>
<td>TNB Concrete Pole</td>
<td>150mm</td>
</tr>
<tr>
<td>Up</td>
<td>230mm</td>
</tr>
<tr>
<td>TNB Spun Pole Low Voltage</td>
<td>140mm</td>
</tr>
<tr>
<td>Medium Voltage</td>
<td>100mm</td>
</tr>
<tr>
<td>190mm</td>
<td>323mm</td>
</tr>
</tbody>
</table>

### 3. RESULTS AND DISCUSSION

This section will show the result in detail of data acquisition for mobile LiDAR Data. Two (2) main types of data have been used in this study. First, Mobile LiDAR Data that has been acquired in LAS format that provide GPS time, point source ID, scan direction flag, number of returns, scan angle rank, edge of flight line, minimum and maximum x, y, z. The second data used in this study is ground truth data that came from field survey manually using digital inclinometer which is Nikon Forestry Pro II Laser Rangefinder to measure the height of pole and measuring tape to measure the circumference of pole.

The characteristics and description of this Mobile LiDAR Data are shown in the Table 3 below while the result of Mobile LiDAR Data Acquisition is shown in Figure 3.

### Table 3. Description of Mobile LiDAR Data

<table>
<thead>
<tr>
<th>LiDAR System</th>
<th>RIEGL VMX-1HA (vehicle)</th>
<th>Geoslam ZEB Discovery (backpack)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cover Area (per m²)</td>
<td>77,259</td>
<td></td>
</tr>
<tr>
<td>Average Point Density (per m²)</td>
<td>203,285</td>
<td></td>
</tr>
<tr>
<td>File Format</td>
<td>LAS Format</td>
<td></td>
</tr>
</tbody>
</table>

### 3.1 High Density MLS point clouds

This section will explain in detail about the significant result based on the experimental result that used random forest (RF) method to detects and validate how good Mobile LiDAR data based on the accuracy of point cloud of pole and point cloud of powerline. The result of classification for Mobile LiDAR Data is shown in Figure 4. Figure 5 shows the the pole and powerline detected using the ML approach, Figure 6 shows the extracted result for pole and powerline after cleaning data and Figure 7 shows the types of TNB assets after features extraction.
Based on Figure 4, it classifies the pole and powerline according to the classification of pole and powerline in training data. Hence, it only detects poles and powerlines that are covered by the study area. Based on Table 4, the overall result for this Machine Learning approach using random forest method for pole and powerline are 0.65 and 0.63 respectively for the accuracy of the assessment method.

### 3.2 Detection of Overhead Powerlines and Electricity Poles

The overall accuracy assessment for automatic extraction using random forest method for pole and powerline that consists of the precision, recall and accuracy for pole and powerline after performing the machine learning approach. The range of values for the correlation coefficient is 0.0 to 1.0. The closer the value to 1.0, the higher the accuracy of mobile LiDAR detected the point clouds of pole and powerline.

<table>
<thead>
<tr>
<th>Assessment Method</th>
<th>Pole</th>
<th>Powerline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>0.830</td>
<td>0.801</td>
</tr>
<tr>
<td>Recall</td>
<td>0.751</td>
<td>0.746</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.650</td>
<td>0.630</td>
</tr>
</tbody>
</table>

**Table 4.** The Precision, Recall and Accuracy of Pole and Powerline

Figure 6 shows the result of circle fit by Taubin where the point clouds of the pole detected using Mobile LiDAR data have been used to analyze the radius of pole in the field. There were 13 poles observed as spun pole low voltage and it gives the radius and coordinate of the pole. The radius value is based on the point clouds detected.

This result is obtained from the previous result which is pole and powerline extraction using random forest method. From these results, there are two types of TNB poles in this study area which are Spun Pole (Low Voltage) in (a) and Concrete Pole (b) in Figure 7, that have both different shapes and sizes. Therefore, from these findings, the assessment method is performed by calculating the root mean square error (RMSE) for the diameter, circumference, and height of pole for spun pole low voltage and the length, width, and height for concrete poles.

Figure 7. The types of TNB assets

<table>
<thead>
<tr>
<th>(a) Spun Pole (Low Voltage)</th>
<th>(b) Concrete Pole</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c) Conductor Segment</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 7.** The types of TNB assets

3.2 Detection of Overhead Powerlines and Electricity Poles

3.3 Detection and Inventory Accuracies

Based on Figure 9 below, there are 50 poles of concrete pole in the field study and the shape of the concrete pole is rectangle, hence it uses a rectangle fit in ArcMap using minimum bounding geometry. This result was providing the length and the area of the pole based on the point clouds detected. The same assessment as spun pole low voltage was used by applying to form the attribute table hence, the RMSE can be computed and validate the result that shows in Table 5 for spun pole low voltage, Table 6 for concrete pole and Table 7 for powerline segment.

**Table 6.** The Precision, Recall and Accuracy of Pole and Powerline
From the result obtained, the overall accuracy of point cloud classification using RF method were 0.650 and 0.630 for pole and powerline, respectively. Most of the misclassified point clouds occurred in poles and powerlines adjacent to trees. It detects point clouds that do not represent poles and powerlines. There are a few factors that affect this result, which is first, there are a few errors in the training data that may wrong classify the point cloud, for example, classify poles that are close to the tree that causes the tree to be detected as a pole. In addition, the second factor that affects the result is the classification of ground and building is not 100 percent accurate because there is a point cloud of the building that is not recognized as a building.

Furthermore, the RMSE values of height for both types of poles is relatively high which caused the error of DTM estimation that is not highly accurate and also influenced the results. At the end of this research study, all the information obtained can be performed on a 3D map in desktop web with the information that fits with TNB’s requirements.

The line graph shows in Figure 10 till Figure 15 is based on the attribute table acquired after cleaning data is performed to see the pattern of the graph for each number of poles detected as observed with mobile LiDAR compared to the actual number of poles detected in the field.

There are 6 figures of graph where the first 3 graphs is for spun pole low voltage and the other 3 graphs is based on concrete pole.

### 3.3.1 Result for Spun Low Voltage Pole

![Figure 10. The Diameter of Spun Low Voltage Pole](image)

Figure 10 shows the diameter of spun pole low voltage. Based on the graph, the value of diameter that observed in pole number 4 and pole number 9 are higher than the actual value which is 0.42m for both pole number 4 and pole number 9. The actual value for diameter is 0.24m.

![Figure 11. The Circumference of Spun Low Voltage Pole](image)

Figure 11 shows the circumference of spun pole low voltage. Based on the graph above, the same cases as the diameter where pole number 4 and pole number 9 are also has the higher value of circumference compared to the actual value which is 1.3188m. The actual value for circumference is 0.73m.

![Figure 12. The Height of Spun Pole Low Voltage](image)

From the result obtained, the overall accuracy of point cloud classification using RF method were 0.650 and 0.630 for pole and powerline, respectively. Most of the misclassified point clouds occurred in poles and powerlines adjacent to trees. It detects point clouds that do not represent poles and powerlines. There are a few factors that affect this result, which is first, there are a few errors in the training data that may wrong classify the point cloud, for example, classify poles that are close to the tree that causes the tree to be detected as a pole. In addition, the second factor that affects the result is the classification of ground and building is not 100 percent accurate because there is a point cloud of the building that is not recognized as a building.

Furthermore, the RMSE values of height for both types of poles is relatively high which caused the error of DTM estimation that is not highly accurate and also influenced the results. At the end of this research study, all the information obtained can be performed on a 3D map in desktop web with the information that fits with TNB’s requirements.
Figure 12 shows the length of concrete pole. Based on the graph above, the actual length of spun concrete pole is 0.23m. There are five poles that has the higher value and overall, the observed value is nearly to the actual value.

3.3.2 Result for Spun Concrete Pole

![Figure 13. The Length of Concrete Pole](image)

Figure 13 shows the length of concrete pole. Based on the graph above, the actual length of spun concrete pole is 0.23m. There are five poles that has the higher value and overall, the observed value is nearly to the actual value.

![Figure 14. The Width of Concrete Pole](image)

Figure 14 shows the width of spun concrete pole. Based on the graph above, there are one pole that has higher observed value which is pole 4, 0.323106m compared to the actual value which is 0.2m. Overall, most of the observed pole value is nearly to the actual value.

![Figure 15. The Height of Concrete Pole](image)

Figure 15 shows the height of concrete pole. Based on the graph above, the patterns of the graph showing there is the higher and the lower value of observed compared to the actual value.

There are several assessment methods that has been made in this study, which is compute the precision, recall getting the accuracy of the mobile LiDAR point clouds that detect pole and powerline. Moreover, from the extracted features itself were computed the RMSE value for pole and powerline to identify how good the mobile LiDAR data classify and detect the TNB asset.

4. CONCLUSION

The accuracy results for this Machine Learning approach using random forest method for pole and powerline are 65% and 63%, respectively. The validation result obtained does not achieve the target at least in 80% of the accuracy of Mobile LiDAR detects the point clouds. However, the result is still considered in good range because it considers a few factors which is most of the misclassified point clouds occurring in poles and overhead cables adjacent to trees.

In conclusion, the aim and objectives of this study have been achieved by employing geospatial approach for TNB asset mapping using mobile LiDAR data in Taman Sri Pulai, Johor. The study has shown that point clouds obtained from high density point clouds can be successfully used to support the semi-automatic approach of TNB asset extraction and inventory. In addition, the outcomes can be used to support the TNB asset maintenance routine.

ACKNOWLEDGEMENTS (OPTIONAL)

Special thanks to team field data capture from GPS Lands Sdn Bhd and Topcon Malaysia for collecting and providing the Mobile LiDAR Data at Taman Sri Pulai, Johor. A big appreciation to Mohd Radhie Mohd Salleh, as being a part of my thesis journey that was supervised and guided on how to do data pre-processing.

REFERENCES


Ibrahim, E.-H. & Lichti, D. 2013. Detecting road poles from


