# EXTRACTION OF POINT CLOUD-BASED INFORMATION FOR POWERLINE CORRIDORS

C. Askit[1], D. Ates[1], I. Bakir[1], S. Seyfeli[1, 2], A. O. Ok[1]

[1] Hacettepe University, Department of Geomatics Engineering, Ankara, Turkey
[2] Hacettepe University, Graduate School of Science and Engineering, Ankara, Turkey
(cerenaskit, dilanates, irembakir, semanur.seyfeli, ozgunok)@hacettepe.edu.tr

**Commission IV, WG IV/7**

**KEY WORDS:** LiDAR, Feature Extraction, Random Forest, Power Line Classification, Obstacle Detection.

**ABSTRACT:**

It has been a challenge for electric power management to automatically extract power lines from LiDAR point clouds. However, environmental and technical issues have made management more challenging in complicated areas where power lines are in close proximity to buildings and/or trees. In this study, the structure and types of the data captured by a LiDAR sensor in regions containing line corridors were analysed. The crucial stage is appropriately identifying from the data the essential parts of a power line corridor route. The point cloud dataset used in the study belongs to the Borssele region in Zeeland, the Netherlands. By manually labelling the dataset, three classes were identified: wire, pylon, and others. For the classification of point clouds, the Random Forest method was utilised. To assess the obstacles posed by the class wire, 5 m, 10 m, and 15 m 3D buffer zones are created. The visual presentation of obstacles within the buffer zone is achieved by assigning them a separate class code and indicating that they are inside and partially within. Based on the results, the correctness values of the classes of wire and others are considered to be satisfactory. However, the class pylon contains points with incorrect labels after the classification. As a result, the accuracy of the pylon class is much lower than the accuracy of the other two classes.

## 1. INTRODUCTION

The demand for electricity has now reached an all-time high level. This means that keeping the load powered up at all times is vital to keeping clients satisfied, which is particularly significant due to the COVID-19 pandemic. To ensure that users receive uninterrupted power, corporations have prioritised the ability to maintain and regulate power lines rapidly and affordably. Thus, the power lines must be thoroughly inspected to detect and eliminate potential threats. The inspection process mainly differs depending on the characteristics of the settlement, i.e., urban or rural areas. While it mostly affects the usual flow of life in urban areas, it also results in numerous concerns, such as health, that cannot be compensated for by any interruption. Forest fires can arise as a consequence of any line breakage in forest areas, which are the heart of the living ecosystem, and this scenario can put the living life on an irreversible path in a rural area.

There are several advantages to accurately mapping power lines: maximising power continuity by identifying infractions and creating safe line corridors; protecting the ecosystem from severe damage if a line breaks by limiting dangers that may affect living things in various locations; and automatically adjusting a helicopter's flight path in accordance with the mapping system's outputs; and so on. However, because electricity wires can reach great distances, traditional field-based inspection is difficult or impossible in many cases. As a result of its technical advances, Light Detection and Ranging (LiDAR) technology has gained the lead role in such studies. The LiDAR, which can directly record high-precision 3D point cloud data of the power line corridor, is preferred to significantly shorten the time needed for the field survey. The point cloud dataset handled using LiDAR provides accurate data for the development of powerful and flexible methods for collecting information about power lines, regardless of the observed region's type and structure. The LiDAR, which creates a three-dimensional point model of a region, enables the detection of obstacles on power lines and the analysis of the region in which the dataset is being used. This makes it easier to identify and mitigate potential threats along the power line corridor.

The primary objective of this work is to accurately map power lines using supervised point cloud data classification. This study utilised LiDAR data acquired by the ESRI Netherlands Map Society for the Borsselle region of Zeeland, Netherlands to locate obstacles in the power line corridor. Point-based supervised classification and obstacle detection are the two primary components of the method presented in this study. Local characteristics were retrieved by k-nearest neighbour (kNN) and entered as an input to the Random Forest (RF) classifier using the neighbourhood relations of the locations. After classification, a buffer region is formed for electrical wires to detect obstacles, with the obstacles shown in a new class code and a table containing their state as inside or partially within. The methodological overview is presented in Figure 1. In this work, CloudCompare (Vehiclemetrics Inc., 2022), MATLAB (The MathWorks Inc., 2021), and ArcGIS Pro (Esri Inc., 2021) were utilised for dataset segmentation, classification, and other related processes, respectively.

The remainder of this work is organised in the following fashion. In the second section, some of the earlier work that has been done in this context is presented. In Section 3, we provide the information that has been gathered regarding the dataset that was used as well as the methodology. The outcomes of the experiments are discussed in Section 4. Section 5 summarises the findings and recommendations.

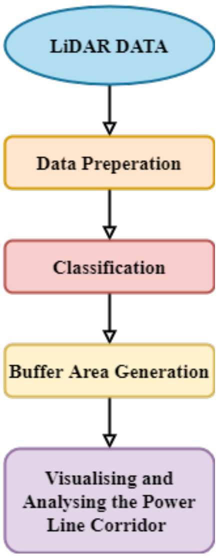



**Figure 1.** Workflow of the methodology.

## 2. RELATED WORK

Previous research analysed the difficulties encountered when managing a power line corridor; their origins; how they can be identified and mitigated; and how the applied work can improve the flow of life. The work conducted by Kim and Sohn (2013) covering the classification of power line objects using airborne laser scanning (ALS) data, considered Random Forests, a point-based supervised classification method that enables the identification of corridor objects using LiDAR data. In a different study, an improved machine learning approach was proposed by Guo et al. (2016) to extract power lines from ALS data and reconstruct the gaps of power lines. A similarity detection method was used to identify the distribution properties of power-line sets. After that process, although most parts of the power lines were found to be clear, several openings were detected. For that reason, the RANSAC rule was utilised. Wang et al. (2017a) presented a power line extraction strategy whose four main components were candidate filtering, multi-scale neighbourhood selection, extraction of spatial topological features, and support vector machine (SVM) classification. In an experiment for an urban region, they discovered that feature extraction based on the multiple-scale spherical neighbourhood yielded superior classification results compared to a single-scale neighbourhood. Wang et al. (2017b) presented a semi-automatic design for the extraction of power line corridors. The direction of the power line corridor was introduced for candidate point filtering, and a multi-scale inclined cylindrical neighbourhood was presented for the extraction of structural features.

In the novel work of Yang and Kang (2018), a voxel-based method was developed to automatically extract transmission lines from airborne LiDAR point cloud data. They also preferred Markov Random Field (MRF) model-based extraction for generating optimal results both locally and globally. Concerning an autonomous vision-based investigation, Nguyen et al. (2018) analysed power line mapping with deep learning algorithms, observing plants that threaten power lines, detecting icing, and monitoring disaster-induced breaches. Therefore, viable beginning points for constructing a system with deep learning methods using an unmanned aerial vehicle (UAV) were provided. Awrangjeb (2019) presented a sequential method for the recovery of power line corridors, pylons, and wires. First, corridors for power lines were recovered from the input point cloud data. Each power line corridor was described as a collection of rectangular regions, and only locations inside each rectangular region were evaluated for the location and extraction of pylons. The non-ground points between two consecutive pylons in the same power line corridor were then used to isolate individual wires. Chen et al. (2020) proposed a filtering process as a partition-based elevation histogram method based on spatial feature analysis of LiDAR data. The triangle modelling method was used in the land modelling process. The kd-tree data structure was employed to build the sphere search model, and safe distance detection of the tree barrier's dangerous points was achieved for power line points and vegetation points recovered by classification.

## 3. MATERIAL AND METHODOLOGY

### 3.1 Dataset

The point cloud data used in this study includes the power lines in close proximity to a nuclear power plant located in the Borssele (Zeeland, Netherlands) region (see Figure 2). The Netherlands elevation data of AHN3 (Current Elevation File Netherlands) were generated from that point cloud data in which both ground-level and off-ground objects (trees, buildings, bridges, and other objects) were resampled from the point cloud to a 0.5-meter grid. The dataset has EPSG:28992 horizontal and EPSG:5709 vertical reference systems and contains approximately 40 million points. In order to make the dataset more manageable, the data has been broken up into multiple small parts. Overall, 10% and 2% of the data belong to the classes of *wire* and *pylon*, respectively, and the rest (88%) belongs to the class *others*. The average point density is ≈ 12 points/m$^2$.

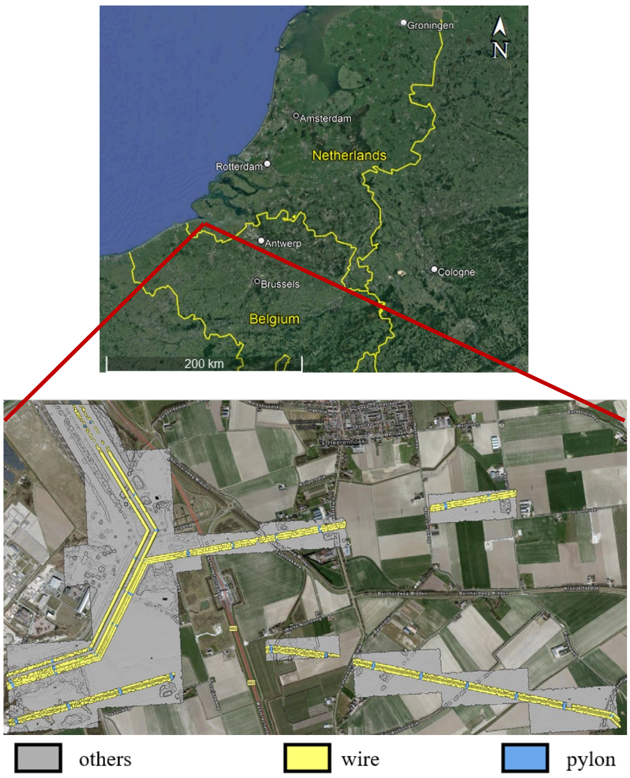



**Figure 2.** Overview of the point cloud data of Borssele, Netherlands with three manually labelled classes.

### 3.2 Point Cloud Classification

At this stage, we perform a point based supervised classification to detect three classes: *wire* (power line), *pylon* (conductor), and *others*. Extracting different features from the LiDAR point cloud as an input and the selection of a classifier are crucial steps. The features are usually computed from points and neighbours that represent the local object structures. For determining the spatial structure for each point under consideration, various neighbourhood types, such as spherical (Brodu and Lague, 2012), cylindrical (Filin and Pfeifer, 2005), and k-nearest neighbourhoods (kNN) (Weinmann et al., 2015; Seyfeli and Ok, 2022), are favoured. In this context, establishing a neighbourhood relationship with kNN is one of the main issues. The 3D kNN relation is in fact a flexible radius spherical neighbourhood. In this approach, a fixed k value is chosen, and the kd-tree algorithm (Bentley, 1975) is performed to find the k points that are closest to each point in the point cloud based on the 3D Euclidean distance. Twelve features, grouped as geometric-based and eigen-based, also known as shape properties, are recovered for each of the points in the neighbourhood to enrich the point attributes.

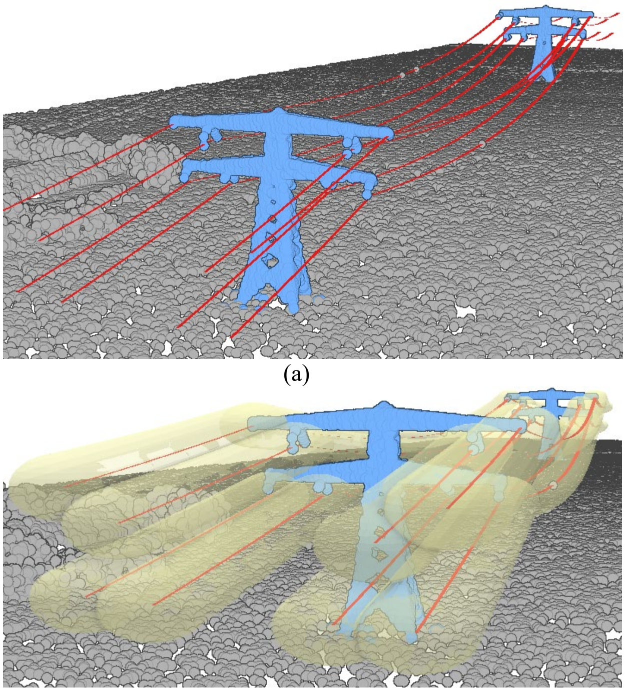
The majority of publicly available 3D point cloud databases only include geometric data expressed as spatial 3D coordinates. The height of the point (Z) is one of the basic geometric features that are taken from the point cloud. The height difference between the lowest and highest points ($\Delta H = h_{max} - h_{min}$) and the standard deviation of the height ($\sigma_H$) of all points in the local relations are additional features obtained from height. Verticality (V) is the final geometric property (Seyfeli and Ok, 2022). Moreover, the eigenvalues $\lambda_{1,2,3}$ have great potential to calculate local shape properties, including dimensionality (linearity, planarity, and sphericity) in the local neighbourhood (Demantké et al., 2012), and other measures such as omnivariance, anisotropy, eigenentropy, sum of eigenvalues, and change of curvature. The eigenvalues listed as $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ and the measures listed in Equation 1 are eigen-based properties.

$$
\begin{aligned}
&\mathit{Linearity} \;\rightarrow\; L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}\\
&\mathit{Planarity} \;\rightarrow\; P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1}\\
&\mathit{Sphericity} \;\rightarrow\; S_\lambda = \frac{\lambda_3}{\lambda_1}\\
&\mathit{Omnivariance} \;\rightarrow\; O_\lambda = \sqrt[3]{\lambda_1\lambda_2\lambda_3}\\
&\mathit{Anisotropy} \;\rightarrow\; A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1}\\
&\mathit{Eigenentropy} \;\rightarrow\; E_\lambda = -\sum_{i=1}^{3} \lambda_i \ln(\lambda_i)\\
&\mathit{Sum\ of\ eigenvalues} \;\rightarrow\; \Sigma\lambda = \lambda_1 + \lambda_2 + \lambda_3\\
&\mathit{Change\ of\ curvature} \;\rightarrow\; c_\lambda = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}
\end{aligned}
\tag{1}
$$

During the classification step, all features retrieved from the neighbourhood of points are fed into a classifier, which assigns them to one of the given (semantic) classes. For simplicity, efficiency, and application in terms of reproducibility, the focus is on individual point classification using the Random Forest (Breiman, 2001) approach, which is frequently recommended for rapid classification of dense LiDAR point clouds.

### 3.3 Obstacle Detection

All living things in the forest ecosystem, especially trees and predators that grow uncontrollably, are potential obstacles. The obstacle detection approach that is based on classified point clouds is divided into five steps. The power lines are converted from points to lines first, as illustrated in Figure 3a (by using the Extract Power Lines from Point Cloud tool). The Buffer 3D tool is then used to assign a buffer on each power line (see Figure 3.b), and violations in buffers are visualised (by using the Locate Las Points by Proximity tool). The status as entirely and partially violated points for each buffer is displayed in the fourth step using the Inside 3D tool. The last optional step is to save the table in Excel format.



(a)

(b)

**Figure 3.** (a) Extracted power lines from point cloud, and (b) the generation of buffer area.

### 3.4 Parameters

In this study, a part of the data was selected from the datasets that contained an obstacle close to the power line. The training set contains over 30 million points, while the test set contains over 10 million points. While the random forest model is estimated using training data, it is subsequently used for the test set to independently evaluate classification outcomes.

The necessary parameters to initiate the methodology are presented in Table 1. Accordingly, a fixed value (100 points) was chosen for the neighbourhood in the point-based classification stage of the study. In total, 12 features were extracted from the adjacent points. For the random forest classifier, the number of trees was taken as 25. As a comparison criterion for safe zone detection, buffer zones were set based on the distances of 5 m, 10 m, and 15 m.

| Section | Parameter Definition | Parameters |
|---|---|---|
| 3.2 | k value | 100 points |
| | # of features | 12 |
| | # of trees of RF | {5, 10, 15, 20, “25”, 30} |
| | # of variables (n) to select for each decision split, √n | 3 |
| | minimum # of observations per tree leaf | 1 (default) |
| 3.3 | Buffer radius (in meter) | 5, 10, 15 |

**Table 1.** Parameter settings.

## 4. RESULTS AND DISCUSSION

In Table 2, the columns represent predicted values for the classes of wire, pylon, and others based on the results of the confusion matrix. The training set's overall, user, and producer accuracy values are all fairly high (all above 99%). However, the accuracy values for the classes of wire and others are satisfactory in the test set, whereas the pylon class contains many mislabelled points (see Table 3). As a result, the user (27.9%) and producer accuracy (61.7%) values for pylons are significantly lower than those for the other two classes. Such poor findings may be the result of imbalanced number of points in the classes, with the pylon class containing the fewest number of points. Despite the fact that the numerical overall accuracy value indicates that the classification performance is quite accurate, a thorough examination of the classified data reveals numerous mislabelled points (see Figure 4).

| | | Predicted Values | | |
|---|---|---|---|---|
| | | others | wire | pylon |
| Actual Values | others | 9,502,555 | 1,943 | 27,343 |
| | wire | 1,368 | 67,696 | 575 |
| | pylon | 6,531 | 168 | 10,807 |

**Table 2.** Confusion matrix of the test set.

| | others | wire | pylon |
|---|---|---|---|
| User Accuracy | 99.9 | 97.0 | 27.9 |
| Producer Accuracy | 99.7 | 97.2 | 61.7 |
| Overall Accuracy | 99.6 | | |
| Kappa | 80.5 | | |

**Table 3.** User, producer, overall accuracy and kappa coefficient of the test data (in percentage - %).

The geometric similarity of wires, pylons, and other objects is a source of confusion in the classification process. If the wires and pylons have any geometrical similarities with other objects, the model's predictions may be incorrect. Railroad or train lines that have a similar shape to power lines may cause classification errors. Similarly, considering the height feature used, wind turbines share a high degree of geometric similarity with pylons, and as a result, the model may classify wind turbines as pylons. The same can be said for the network poles and crane towers.

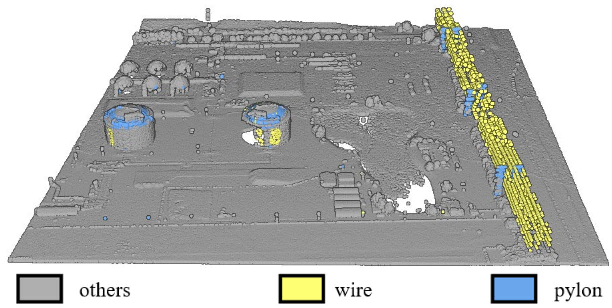
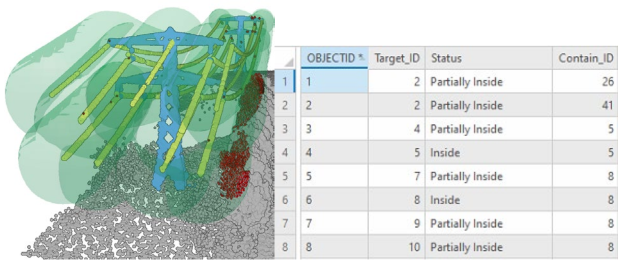



**Figure 4.** Classification result of a part of the test set.

The topography of the study area is a second point of discussion. Due to the difficulty of field-based inspection in regions such as forests and mountains, violation detection in rough terrain is essential for power line safety and the protection of animals. The topography of the data used for classification impacts the model's efficiency. Due to the comparatively smooth surface of the study area's topography, a rough surface may not produce the same results. In a flat environment, the absolute elevation of power lines does not significantly vary. However, this may not be the case in a rugged environment. As a result, the precision may possibly diminish over uneven terrain. To address this issue, the terrain can be represented using a digital terrain model with a resolution comparable to the input point cloud.

In Figure 5, violations within the 10-meter-radius buffer zone are represented as inside and partially within. The tables offer X, Y, and Z information for each obstacle. The 'Object ID' column in the table represents the buffer ID in the data, the 'Target ID' column represents the identity of the power lines, the 'Status' column represents the status of the obstacles entering the buffer zone, and the 'Contain ID' column represents the number of obstacles entering the buffer zone.



| | OBJECTID * | Target_ID | Status | Contain_ID |
|---|---|---|---|---|
| 1 | 1 | 2 | Partially Inside | 26 |
| 2 | 2 | 2 | Partially Inside | 41 |
| 3 | 3 | 4 | Partially Inside | 5 |
| 4 | 4 | 5 | Inside | 5 |
| 5 | 5 | 7 | Partially Inside | 8 |
| 6 | 6 | 8 | Inside | 8 |
| 7 | 7 | 9 | Partially Inside | 8 |
| 8 | 8 | 10 | Partially Inside | 8 |

**Figure 5.** Visualization of obstacles within a 10-meter radius with a status table.

As depicted in Figure 6, the buffer's radius and the number of obstacles it contains are directly proportional. Expanding the radius of the buffer area increases the number of obstacles within the buffer, making it simpler to discover violations. Calculating the radius of the buffer region is closely tied to the topographic structure of the data. As the radius expands, it becomes possible for unobstructed objects to cross the expanding buffer zone, resulting in misleading analysis results. In order to avoid erroneous results and construct a safe area that is convenient for the data, the topographic structure should be used to estimate the buffer radius. Besides, in this work, we preferred an individual buffer around each powerline extracted; however, one buffer around all extracted power lines, i.e. uniting the individual buffers, might be easier for dealing with the maintenance purposes.


https://doi.org/10.5194/isprs-archives-XLVIII-4-W6-2022-41-2023 | 

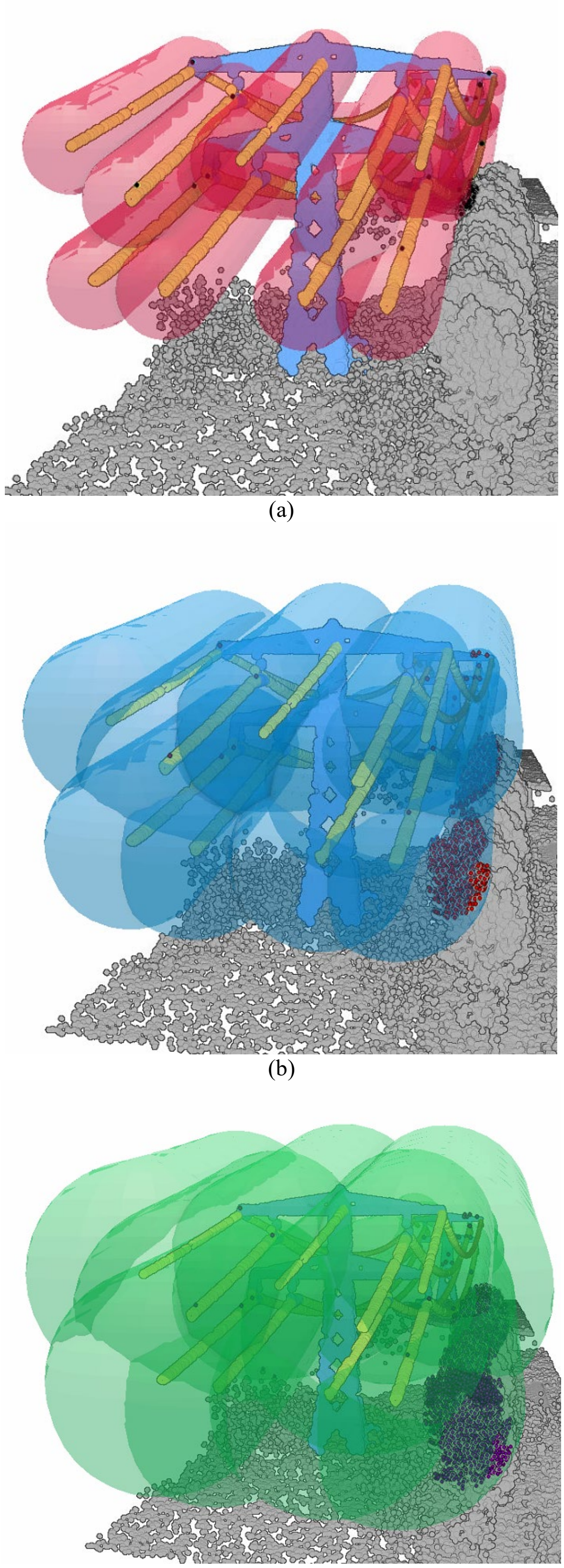



**Figure 6.** Visualization of obstacles within a 5m, 10m, and 15m radius.

## 5. CONCLUSION

The purpose of this study is to classify power lines and build a buffer zone based on their classification. Identifying existing obstacles near power lines using LiDAR data may reduce the amount of time and effort invested and significantly increase operating efficiency. Provided as input are point cloud data for power lines, power poles, and other objects in urban and forest environments. LiDAR data was classified using the Random Forest approach.

The supervised classification using the random forest approach achieved a producer accuracy of 97.2% for the class wire and 61.7% for the class pylon. A satisfactory level of performance has been obtained. The identification of obstructions that enter the buffer zone is accomplished effectively. The classification accuracy for electrical wires was determined to be relatively high. The wires' nearly straight and almost parallel structure assures a low error rate. Even if a moderate degree of accuracy is attained through the classification of pylons, misclassifications sometimes occur, especially for points with similar geometries. Cranes, antennas, trees, and other poles with comparable characteristics were classified as electrical poles in the point cloud data since only geometric properties of neighbourhood points were considered during the classification.

The classification performance may increase if the point cloud data can be supported with RGB features in addition to the geometric properties extracted from the point cloud. Parallel processing by dividing the study area into smaller grids can be utilised to improve the efficiency of the processing. In this study, three different classes were utilised, and the obstacles were defined in a single class (i.e., others). The number of classes can be increased to better identify obstacles. For example, it can be expressed that the majority of the detected obstacles are trees. For this reason, a tree class can be formed for power line corridors. In regions where the density of the point cloud data is different, it will not be appropriate to hold several parameters constant. In such a case, the parameters related to the neighbourhood should be updated according to the point density of the input dataset. In the context of supervised classification algorithms, this study could benefit from additional machine learning or deep learning strategies recently developed in the literature.

## REFERENCES

Awrangjeb, M., 2019. Extraction of Power Line Pylons and Wires Using Airborne LiDAR Data at Different Height Levels. *Remote Sensing* 11(15), 1798. https://doi.org/10.3390/rs11151798

Bentley, J.L., 1975. Multidimensional binary search trees used for associative searching. *Communications of the ACM* 18(9), 509–517. https://doi.org/10.1145/361002.361007

Breiman, L., 2001. Random Forests. *Machine Learning* 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brodu, N., Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing* 68, 121–134. https://doi.org/10.1016/j.isprsjprs.2012.01.006

Chen, X.H., Dai, J.Q., He, Y.R., Ma, W.W., 2020. Power line extraction and analysis based on LiDAR. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., XLII-3/W10, 91–96. https://doi.org/10.5194/isprs-archives-XLII-3-W10-91-2020

Demantké, J., Mallet, C., David, N., Vallet, B., 2012. Dimensionality based scale selection in 3D LiDAR point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XXXVIII-5/W12, 97–102. https://doi.org/10.5194/isprsarchives-XXXVIII-5-W12-97-2011

Esri Inc., 2021. A Professional Desktop GIS (ArcGIS Pro) Application, Version 2.8. https://pro.arcgis.com (13 May 2021)

Filin, S., Pfeifer, N., 2005. Neighbourhood Systems for Airborne Laser Data. *Photogrammetric Engineering & Remote Sensing* 71(6), 743–755. https://doi.org/10.14358/PERS.71.6.743

Guo, B., Li, Q., Huang, X., Wang, C., 2016. An Improved Method for Power-Line Reconstruction from Point Cloud Data. *Remote Sensing* 8(1), 36. https://doi.org/10.3390/rs8010036

Kim, H.B., Sohn, G., 2013. Point-based Classification of Power Line Corridor Scene Using Random Forests. *Photogrammetric Engineering & Remote Sensing* 79(9), 821–833. https://doi.org/10.14358/PERS.79.9.821

Nguyen, V.N., Jenssen, R., Roverso, D., 2018. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *International Journal of Electrical Power & Energy Systems* 99, 107–120. https://doi.org/10.1016/j.ijepes.2017.12.016

Seyfeli, S., Ok, A.O., 2022. Classification of Mobile Laser Scanning Point Cloud in an Urban Environment Using kNN and Random Forest, in: Ben Ahmed, M., Boudhir, A.A., Karaş, İ.R., Jain, V., Mellouli, S. (Eds.), Innovations in Smart Cities Applications Volume 5. Springer International Publishing, Cham, 963–973. https://doi.org/10.1007/978-3-030-94191-8_78

The MathWorks Inc., 2021. A Programming and Numeric Computing (MATLAB) Software, Version R2021a. https://mathworks.com/products/matlab (16 Mar 2021)

Vehiclemetrics Inc., 2022. 3D Point Cloud and Mesh Processing (CloudCompare) Software, Version 2.12.1 Kyiv. Open-Source Foundation. https://cloudcompare.org (7 May 2022)

Wang, Y., Chen, Q., Li, K., Zheng, D., Fang, J., 2017a. Airborne lidar power line classification based on spatial topological structure characteristics. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.,* IV-2/W4, 165–169. https://doi.org/10.5194/isprs-annals-IV-2-W4-165-2017

Wang, Y., Chen, Q., Liu, L., Zheng, D., Li, C., Li, K., 2017b. Supervised Classification of Power Lines from Airborne LiDAR Data in Urban Areas. *Remote Sensing* 9(8), 771. https://doi.org/10.3390/rs9080771

Weinmann, M., Jutzi, B., Hinz, S., Mallet, C., 2015. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing* 105, 286–304. https://doi.org/10.1016/j.isprsjprs.2015.01.016

Yang, J., Kang, Z., 2018. Voxel-Based Extraction of Transmission Lines From Airborne LiDAR Point Cloud Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11(10), 3892–3904. https://doi.org/10.1109/JSTARS.2018.2869542