World Academy of Science, Engineering and Technology
International Journal of Geological and Environmental Engineering
Vol:14, No:6, 2020

Extracting Terrain Points from Airborne Laser Scanning Data in Densely Forested Areas

Ziad Abdeldayem, Jakub Markiewicz, Kunal Kansara, Laura Edwards

***Abstract*—**Airborne Laser Scanning (ALS) is one of the main technologies for generating high-resolution digital terrain models (DTMs). DTMs are crucial to several applications, such as topographic mapping, flood zone delineation, geographic information systems (GIS), hydrological modelling, spatial analysis, etc. Laser scanning system generates irregularly spaced three-dimensional cloud of points. Raw ALS data are mainly ground points (that represent the bare earth) and non-ground points (that represent buildings, trees, cars, etc.). Removing all the non-ground points from the raw data is referred to as *filtering*. Filtering heavily forested areas is considered a difficult and challenging task as the canopy stops laser pulses from reaching the terrain surface. This research presents an approach for removing non-ground points from raw ALS data in densely forested areas. Smoothing splines are exploited to interpolate and fit the noisy ALS data. The presented filter utilizes a weight function to allocate weights for each point of the data. Furthermore, unlike most of the methods, the presented filtering algorithm is designed to be automatic. Three different forested areas in the United Kingdom are used to assess the performance of the algorithm. The results show that the generated DTMs from the filtered data are accurate (when compared against reference terrain data) and the performance of the method is stable for all the heavily forested data samples. The average root mean square error (RMSE) value is 0.35 m.

***Keywords*—**Airborne laser scanning, digital terrain models, filtering, forested areas.

I. Introduction

ALS is an active remote sensing technology that provides three dimensional coordinates of the scanned area. Raw ALS data represent all natural (bare earth, trees, lakes, rivers, etc.) and manmade (buildings, vehicles, bridges, roads, etc.) features on the surface of the ground [1], [2]. DTMs could be generated after extracting only the bare earth (ground) points from the raw ALS data [3]. Extracting the ground points in heavily forested areas is a challenging task as the dense trees prevent the laser pulses from reaching the ground surface. Therefore, there is a need to develop methods that focus on filtering raw ALS data in forested areas specifically, such as [4]. Across the past few years, many filtering methods were presented. These methods could be classified into three main categories: the segmentation based [5], [6], the morphological operations based [7], [8] and the interpolation based methods [9], [10]. However, there are some filtering approaches that do not belong to any of the above-mentioned categories, such as the statistical based methods and the TIN (triangulated irregular network) based methods [11]-[13].

In this paper, a method for filtering ALS data in forested areas is presented. The approach is a surface interpolation based and mainly uses the smoothing splines criterion to fit the noisy forests point clouds and a weight function for assigning a weight value for each point of the raw ALS data for improving the classification process into ground or non-ground point.

Ziad Abdeldayem* and Kunal Kansara are with the Department of Civil Engineering, Faculty of Engineering and Technology, Liverpool John Moores University, Liverpool, UK (*e-mail: z.abdeldayem@ljmu.ac.uk).
Jakub Markiewicz is with the Department of Photogrammetry, Faculty of Geodesy and Cartography, Warsaw University of Technology, Warsaw, Poland.
Laura Edwards is with the School of Biological and Environmental Sciences, Faculty of Science, Liverpool John Moores University, Liverpool, UK.

II. Smoothing Splines Filtering Approach

A. Cubic Smoothing Splines Criterion

A cubic splines function, uses the cubic polynomial to interpolate the data on each interval.

$$f_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \quad (1)$$

where $a_i, b_i, c_i, \& \, d_i$ are the coefficients that can be determined using one of the available mathematical methods, such as [14]-[16]. Fig. 1 illustrates the noisy raw ALS data and how the cubic splines interpolation function fits these data measurements passing through each point.

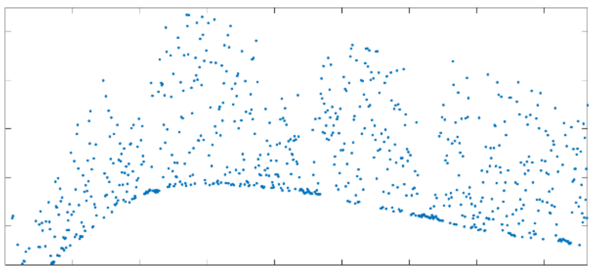

Fig. 1 (a) Raw ALS data (blue dots)

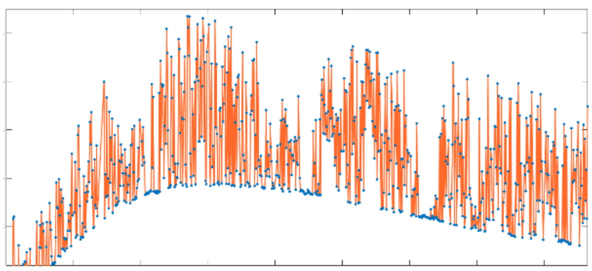

Fig. 1 (b) A cubic splines function fitting the raw data by passing through each point (orange solid line)

For filtering raw ALS data, we need kind of a function that passes close to the data points but not exactly through each one. Cubic smoothing splines provides a good balance of the fitted curve as it passes between the ground and non-ground ALS data as shown in Fig. 2. The criterion for smoothing splines,

$$\alpha \sum_{i=1}^{n} w_i(z_i - f(x_i))^2 + (1-\alpha)\int_{x_1}^{x_n}\{f''(x)\}^2 dx \qquad (2)$$

consists of two parts, the first one, $\alpha \sum_{i=1}^{n} w_i(z_i - f(x_i))^2$, measures how closely the function fits the data measurements, and the second one, $(1-\alpha)\int_{x_1}^{x_n}\{f''(x)\}^2 dx$, measures the smoothness of the fitted curve.

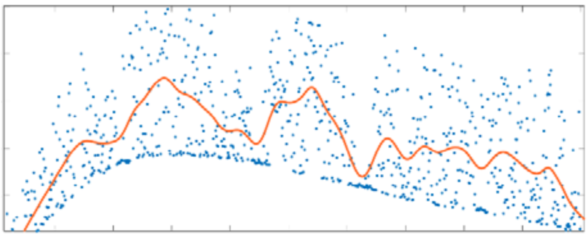

Fig. 2 The fitted curve using cubic smoothing splines criterion

B. ASL Data Filtering

The suggested filtering approach stages are presented in Fig. 3. MATLAB software is used to implement all the processing stages by designing a robust code specifically for this aim. A brief description of the processing operations will be provided in the next section.

Stage 1: Creating a Uniform Grid for the Raw ALS Data

Raw ALS data measurements are irregularly distributed points. Therefore, it is important to rearrange these random points into a regularly spaced grid [17]. The number of rows (N_r) and columns (N_c) of the grid can be calculated as:

$$N_r = (Y_{max} - Y_{min}/Cell\ size) + 1$$
$$N_c = (X_{max} - X_{min}/Cell\ size) + 1 \qquad (3)$$

The selection of the grid cell size depends on the raw data density. It represents the spatial resolution. As a first step towards removing the non-ground (trees) points, the algorithm will remove the higher point (larger elevation value) if more than one have fallen in one cell. This will produce a uniform 2D grid; each cell contains the coordinates of one point of the ALS data.

Stage 2: ALS Data Fitting and Computing the Residual Values

The cubic smoothing splines interpolation function - Criterion (2) - is applied to each row of the abovementioned grid. After normalizing the data, the function fits them using a value of 0.9999 for the smoothing parameter α. This value is fixed for all the processing stages. It is determined after conducting many tests using different values of the smoothing parameter. The generated curve follows the general trend of the ALS data without passing through each point as shown in Fig. 2. The difference in elevation values between the ALS point $z_{i\,(ALS)}$ and the fitted curve $z_{i\,(curve)}$ is called the residual value res_i,

$$res_i = z_{i(ALS)} - z_{i(curve)} \qquad (4)$$

Stage 3: Removing the Non-Ground Points and Computing the Weight Value for Each Point

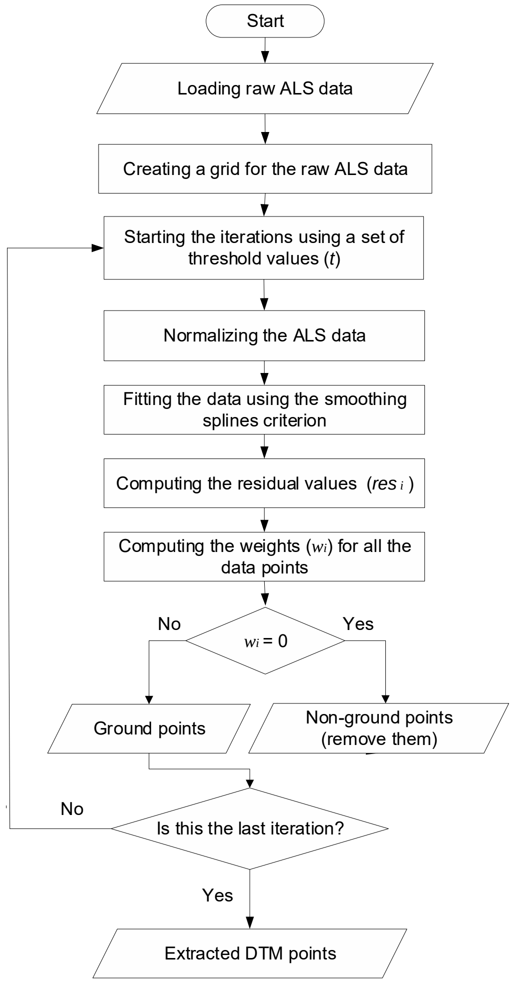

Fig. 3 Flowchart illustrating all the ALS data filtering stages, starting from loading the raw data till extracting the filtered DTM points

All the ALS points located above the fitted curve by a predefined threshold value t will be classified as trees (non-ground points) and removed. The threshold value represents a buffer to help the algorithm to avoid removing the ground points [18]. In the first iteration, all the points would be given an equal weight values. After that, (5) will be applied to assign a weight value for each point.

$$w_i = \begin{cases} 1, & res_i < \sigma \\ 1 - 2\left(\frac{res_i - \sigma}{t-\sigma}\right)^2, & \sigma \le res_i \le (\sigma + t)/2 \\ 2\left(\frac{t - res_i}{t-\sigma}\right), & (\sigma + t)/2 \le res_i \le t \\ 0, & res_i \ge t \end{cases} \qquad (5)$$

The initial threshold value could be a small number, $t=0.25$, so that most of the trees points could be detected and removed in the first iteration. Then, to avoid removing the ground points, the second iteration will start with a big value and then it would be decreased gradually from 7 to 1.

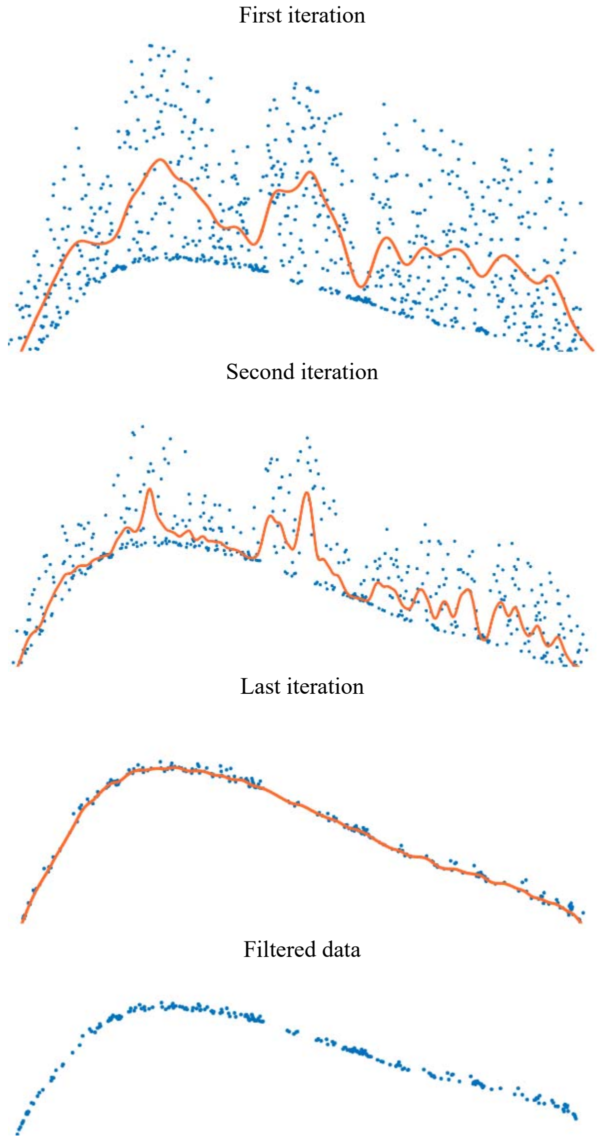

Fig. 4 Filtering raw ALS data (blue points) by fitting the data using the cubic smoothing splines approach. The data above the fitted curve (the solid line) will be classified as trees and removed

Fig. 4 illustrates how the algorithm removes the trees points (above the fitted curve) and how the ground points attract the fitted curve after each iteration.

III. Experiments and Results

The performance of the presented filtering approach is tested using three samples of forested areas in the United Kingdom. These samples represent different tree types and data densities, and range from moderate sloped areas to steep mountains. The area of each sample is about 1 km², and the data density for samples 1 (Dartmoor forest, England) and 2 (Drumtochy forest, Scotland) is 1 point/m², whereas sample 3 (Caerwent forest, Wales) represents very dense data, 4 points/m². The cubic smoothing splines algorithm is applied on these three forests' samples.

DTMs are generated from the extracted ground points for each sample and then they are compared cell-by-cell against very accurate reference data (DTMs) provided by the UK Environment Agency with a vertical accuracy of ± 0.05 m [19]. The RMSE value is calculated for each sample using,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(E_{(i)})^2}{n-1}} \tag{6}$$

where $E_{(i)}$ is the error for each point,

$$E_{(i)} = Z_{(i)Reference\ DTM} - Z_{(i)Filtered\ DTM} \tag{7}$$

In addition, the three samples are filtered using a professional algorithm (ENVI LiDAR) designed by the world-leading company, Harris Geospatial Solutions. Then, the RMSEs are calculated and compared against the RMSEs using the presented raw ALS data filtering method. Table I shows the results of filtering the three samples of the forested areas using both the proposed smoothing splines method (SSM) and ENVI software.

TABLE I
ACCURACY ASSESSMENT OF THE FILTERED DATA

Forest Sample	RMSE (m)		Mean error (m)		Minimum error (m)		Maximum error (m)	
	SSM	ENVI	SSM	ENVI	SSM	ENVI	SSM	ENVI
Sample 1	0.19	0.41	-0.01	0.08	-2.56	-14.90	1.13	6.00
Sample 2	0.48	0.88	-0.46	0.70	-3.90	-6.13	5.70	10.95
Sample 3	0.39	0.76	-0.79	1.07	-2.86	-3.36	2.49	8.62
Average	0.35	0.68	-0.42	0.62	-3.11	-8.13	3.11	8.52

The average RMSE values using the presented filter and ENVI filter are 0.35 m and 0.68 m, respectively. These results clearly show the excellent performance of the suggested filtering algorithm. Fig. 5 illustrates the raw ALS data (before filtering) and the generated DTMs (after filtering). Sample 2 represents dense forest on steep mountains, which could be considered one of the most challenging areas for any filtering method, as it can mislead the filter and affect the algorithm's performance easily. Despite both filters have reached the largest RMSE value with this sample, the presented approach has succeeded in achieving higher accuracy for the resulted filtered data, as the RMSE is 0.48 m, in comparison with ENVI's RMSE, which is 0.88 m. In general, the statistical results in Table I show that the suggested algorithm has accurately filtered the raw ALS data. It performed very well with all samples, even with the very dense data (Sample 3).

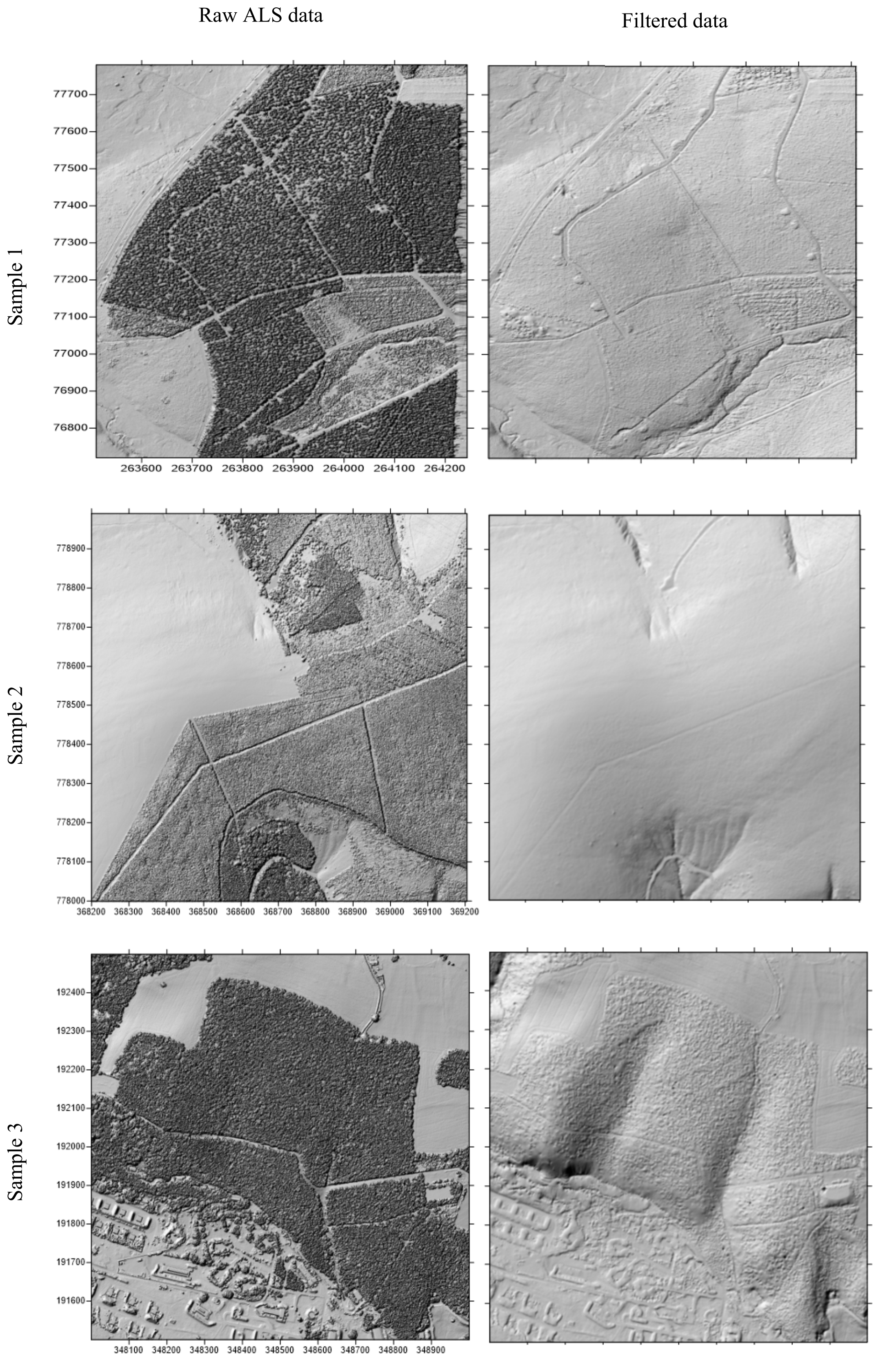

Fig. 5 Shaded relief images of the three forested areas' samples, before and after filtering, using the presented approach

IV. Conclusions

This paper presents a method for extracting the terrain points from raw ALS data in forested areas. The approach is based on using the smoothing splines function with a predefined smoothing parameter and set of threshold values. The algorithm removes the trees (non-ground points) iteratively; all the ALS points have the same weight in the first iteration. After that, a weight function that allocates a weight value for each ALS point is used to enhance the filtering process. For assessing the performance of the method, three raw ALS data samples represent different spatial resolution and terrain types are tested. The accuracy of the filtered data is assessed by calculating the RMSE values of the generated DTMs using the proposed method. In addition, the three samples are filtered using a professional software (ENVI LiDAR), and the RMSE results are calculated and compared against the proposed method's results. The average RMSE values are 0.35 m and 0.68 m using the presented method and ENVI software, respectively. All the results confirm the reliability of the proposed approach for filtering heavily forested areas.

Acknowledgment

The author would like to specially thank EDINA digimap LiDAR data service for providing the required data for the forested areas. They would also like thank V. Sridhar from Mathworks for his help in improving the algorithm performance.

References

[1] Y. Quan, J. Song, X. Guo, Q. Miao, and Y. Yang, "Filtering LiDAR data based on adjacent triangle of triangulated irregular network," Multimedia Tools and Applications, vol. 76, no. 8, pp. 11051–11063, 2017.

[2] Y. B. Yang, N. N. Zhang, and X. L. Li, "Adaptive slope filtering for airborne Light Detection and Ranging data in urban areas based on region growing rule," Survey Review, vol. 6265, no. June, pp. 1–8, 2016.

[3] C. A. Silva, C. Klauberg, Â. M. K. Hentz, A. P. D. Corte, U. Ribeiro, and V. Liesenberg, "Comparing the performance of ground filtering algorithms for terrain modeling in a forest environment using airborne LiDAR data," Floresta e Ambiente, vol. 25, no. 2, 2018.

[4] Z. Hui, Y. Hu, Y. Z. Yevenyo, and X. Yu, "An improved morphological algorithm for filtering airborne LiDAR point cloud based on multi-level kriging interpolation," Remote Sensing, vol. 8, no. 1, pp. 12–16, 2016.

[5] G. Sithole and G. Vosselman, "Filtering of airborne laser scanner data based on segmented point clouds," International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 36, no. 3/W3-4, pp. 66–71, 2005.

[6] A. Kobler, N. Pfeifer, P. Ogrinc, L. Todorovski, K. Oštir, and S. Džeroski, "Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain," Remote Sensing of Environment, vol. 108, no. 1, pp. 9–23, 2007.

[7] Z. Chen, B. Gao, and B. Devereux, "State-of-the-Art: DTM Generation Using Airborne LIDAR Data," Sensors, vol. 17, no. 1, p. 150, 2017.

[8] L. Liu and S. Lim, "A voxel-based multiscale morphological airborne lidar filtering algorithm for digital elevation models for forest regions," Measurement: Journal of the International Measurement Confederation, vol. 123, pp. 135–144, 2018.

[9] W. Schickler and A. Thorpe, "Surface estimation based on LIDAR," Proceedings of the ASPRS Annual Conference, no. April, pp. 23–27, 2001.

[10] H. Chen, M. Cheng, J. Li, and Y. Liu, "An Iterative Terrain Recovery Approach To Automated Dtm Generation From Airborne Lidar Point Clouds," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIX-B4, no. September, pp. 363–368, 2012.

[11] Q. Chen, H. Wang, H. Zhang, M. Sun, and X. Liu, "A point cloud filtering approach to generating DTMs for steep mountainous areas and adjacent residential areas," Remote Sensing, vol. 8, no. 71, 2016.

[12] Z. Hui, D. Li, S. Jin, Y. Y. Ziggah, L. Wang, and Y. Hu, "Automatic DTM extraction from airborne LiDAR based on expectation-maximization," Optics and Laser Technology, vol. 112, pp. 43–55, 2019.

[13] S. Cai, Z. Wuming, L. Xinlian,W. Peng, Q. Jianbo, Y. Sisi, Y. Guangjian, S. Jie, "Filtering Airborne LiDAR Data Through Complementary Cloth Simulation and Progressive TIN Densification Filters," Remote Sensing, vol. 11, no. 9, p. 1037, 2019.

[14] T. Ramsay, "Spline smoothing over difficult regions," Journal of the Royal Statistical Society. Series B: Statistical Methodology, vol. 64, no. 2, pp. 307–319, 2002.

[15] B. W. Silverman, "Some Aspects of the Spline Smoothing Approach to Non-Parametric Regression Curve Fitting," Journal of the Royal Statistical Society, vol. 74, no. 1, pp. 1–52, 1985.

[16] C. De Boor, "A Practical Guide to Splines," Springer, pp. 207-240, 2001.

[17] K. Zhang, S. C. Chen, D. Whitman, M. L. Shyu, J. Yan, and C. Zhang, "A progressive morphological filter for removing nonground measurements from airborne LIDAR data," IEEE Transactions on Geoscience and Remote Sensing (TGRS), vol. 41, no. 4 PART I, pp. 872–882, 2003.

[18] Z. Abdeldayem, "Automatic Weighted Splines Filter (AWSF): A New Algorithm for Extracting Terrain Measurements From Raw LiDAR Point Clouds," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 60-71, 2020.

[19] Environment Agency, "Environment Agency LIDAR data," Environment Agency LIDAR data, p. 16, 2016.