A Maximum Entropy Based Outlier Removal for Airborne LiDAR Point Clouds

Ge Jiang, *Student Member, IEEE*, Derek D. Lichti, Tiangang Yin, and Wai Yeung Yan, *Senior Member, IEEE*

***Abstract*—Airborne light detection and ranging (LiDAR) data often suffer from noisy returns hovering in empty space within the collected 3-D point clouds. This can be attributed to system-induced factors, such as timing jitter and range walk error, or instantaneous air conditions, such as smoke, rain, clouds, etc. These floating points are indeed outliers, which significantly affect the subsequent analytical processes. Though various point cloud denoising methods are proposed based on sparsity assumption and elevation, they are highly unlikely to remove both clustered and scattered noisy points, especially those located close to the point clouds. Meanwhile, the performance of existing methods does not perform well when noisy points appear close to the ground or on rugged terrain. Accordingly, we propose a maximum entropy based outlier removal (MEOR) method for airborne LiDAR point clouds. More specifically, the proposed method includes two stages, i.e., one global coarse outlier removal stage (MEOR-G) and the subsequent local refined outlier removal stage (MEOR-L). In each stage, the MEOR algorithm is exploited to 1) produce an elevation histogram for the point clouds, 2) search for the elevation threshold to distinguish noisy points and valid points, and 3) remove noisy points and preserve valid data points. We conduct several comprehensive experiments to compare the performance of our proposed MEOR against four other existing noisy point removal methods on four LiDAR datasets. The experimental results demonstrate that MEOR significantly outperforms four other denoising methods by simultaneously removing clustered and scattered noisy points and achieves an improvement by 0.126–99.815%, 0–100%, 0.001–8.454%, and 0.053–99.691% in terms of recall, precision, overall accuracy, and F1 score, respectively.**

***Index Terms*—Airborne LiDAR, clustered noisy points, maximum entropy, outlier removal, point cloud denoising, scattered noisy points.**

Received 10 May 2024; revised 20 August 2024; accepted 5 October 2024. Date of publication 10 October 2024; date of current version 28 October 2024. This work was supported in part by the FCE Start-up Fund (BE2U) of the Hong Kong Polytechnic University, in part by the General Research Fund Project under Grant 15221022 (Quantifying Airborne LiDAR Data Artifacts) and Grant 15215023 (Airborne LiDAR based Large-Scale Urban Object Classification via Semantic Scene Completion), and in part by the Research Grants Council of the Hong Kong Special Administrative Region. *(Corresponding author: Wai Yeung Yan.)*

Ge Jiang is with the Department of Land Surveying and Geo-informatics, Hong Kong Polytechnic University, Hong Kong, and also with the Department of Geomatics Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada (e-mail: 20034646r@connect.polyu.hk).

Derek D. Lichti is with the Department of Geomatics Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada (e-mail: ddlichti@ucalgary.ca).

Tiangang Yin is with the Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hong Kong (e-mail: tiangang.yin@polyu.edu.hk).

Wai Yeung Yan is with the Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hong Kong, and also with the Department of Civil Engineering, Toronto Metropolitan University, Toronto, ON M5B 0C3, Canada (e-mail: waiyeung.yan@polyu.edu.hk).

Digital Object Identifier 10.1109/JSTARS.2024.3478069

I. Introduction

LiDAR point cloud data analytics ride on the recent rapid development of artificial intelligence (AI) [1], [2], [3], [4], [5]. Despite that, as powerful as deep learning is nowadays, it can never be free of incorrect predictions or errors [6], [7], [8]. The concept of "Garbage in, Garbage out" still stands in AI, and thus data quality becomes increasingly important. Nevertheless, the airborne LiDAR community, at least at the moment, tends to overlook the importance of addressing the point cloud quality, which hinders the applications of airborne LiDAR point cloud, e.g., individual tree detection, 3-D modeling, and other fields, in contrast to the greater emphasis on geometric accuracy and consistency of the points [9], [10], [11], [12]. Indeed, typical airborne LiDAR data artifacts encompass various issues, not limited to striping, noise, data holes, and motion artifacts [13]. Very often, airborne LiDAR may contain points "hovering" in empty space due to rain, smoke, clouds, or unknown floating objects, resulting in instantaneous data points located above the topography [14], [15], [16], [17], [18]. Thus, outlier removal (OR) methods of point cloud become necessary before performing any analytical process.

Accordingly, many research efforts focus on removing noisy points of airborne LiDAR point clouds [19], [20], [21], [22], [23], [24], [25], [26], [27], [28]. Existing studies emphasized the importance of point cloud density and regard it as an index to differentiate noise from valid data points, i.e., ground and object point clouds. Among these, the spatial frequency OR (SFOR) is a commonly adopted filtering technique [19]. It aims to query neighbor points located within a specific radius of sphere in each of the data points. For SFOR, a data point with a number of retrieved neighbor points smaller than a predefined threshold is identified as a noisy return. On the other hand, the statistical OR (SOR) method proposed by [20] searches k nearest neighbor points for each individual data point and calculates the mean distance of k nearest neighbor points. The noisy points are assessed based on the sigma rule derived from the entire dataset, i.e., a data point with the average distance of k nearest neighbor points larger than N standard deviations from the entire mean distance is classified as a noisy point. The major drawback of SOR is its sensitivity to return density variations in certain areas, such as the decreasing density from the top to the bottom of tree canopies. To overcome this limitation, a novel method based on spatial hierarchical directional relationship (SHDR) and region growing specializing in detecting noisy points for airborne LiDAR data is proposed [21]. It first divides the original point cloud into 3-D grids and constructs 14 spherical neighborhoods along

the respective 14 directions for the initial seed point based on 3-D space topology. Then the farthest interior point along each direction in the spherical neighborhood is added to the seed point set. It iteratively executes all points by using seed points as new region growing seed points and distinguishes noisy points and valid points by examining whether they are inside the sphere. This method is effective in finding noisy points for airborne LiDAR data. However, the computational cost is fairly high and the selection of distance threshold, which is not a reliable and flexible behavior, can greatly affect the denoising results. From a different perspective, the morphological OR (MOR) method based on the thin plate splines to process the elevation information of point cloud is proposed [22]. MOR combines two fundamental morphological operations: 1) erosion that levels down the points with higher elevation to the lowest elevation along the column spline, and 2) dilation that levels up the points with lower elevation to the highest elevation along the column spline. Though MOR can remove outliers to a certain extent, it also introduces deformation to the valid point clouds with terrain relief. Another possible drawback is the requirement of transforming the point cloud to columns or voxels, which demand interpolation that may subsequently generate deviation to the raw data. Deep learning-based methods [17], [23], [24] have demonstrated a viable approach in LiDAR point cloud denoising lately. Nevertheless, high computational resources and sufficient well-labeled training data are required to yield the best denoising performance.

Although existing airborne LiDAR OR methods can identify either scattered or clustered noisy points, they fail to identify both types of noise simultaneously. Meanwhile, some of these methods may wrongly remove valid data points and destroy the integrity of the original point cloud in representing the topography. Besides, these methods rely on empirical threshold adjustment or parameter setting, which limits the generalization and efficiency of point cloud denoising. To this end, in this article, we propose a two-stage global-local model with the maximum entropy OR (MEOR) method to identify both clustered noisy points and isolated outliers based on the elevation of data points acquired by airborne LiDAR. A global, coarse denoising step, MEOR-G, first removes clustered and scattered noisy points that are far away from the valid data points. The subsequent local refined denoising step, MEOR-L, further identifies isolated outliers that are close to the valid data points. Various experimental results show that our proposed method effectively removes both clustered and scattered noisy points. Meanwhile, it preserves the details and integrity of valid point clouds based on elevation information that is consistent and intuitive compared to the four aforementioned classic denoising methods.

The rest of the article is organized as follows. We first present the proposed algorithm in Section II in detail. Section III describes the datasets and presents the experimental parameters settings. The produced experimental results and discussions are shown in Sections IV and V, respectively, including a comparison with four other existing point cloud denoising methods. The Section VI concludes our findings of the study.

II. Methodology

A. Overall Workflow

The proposed MEOR method consists of two stages, i.e., MEOR-G and MEOR-L, which are shown in Fig. 1. The objective of MEOR-G is to estimate a global threshold of elevation, represented as T' in Fig. 1(b), based on the maximum entropy [32]. The estimated T' thus differentiates noisy data points from valid data points representing the topography. This step removes clustered noisy points and isolated outliers with notable elevation differences compared to the valid point cloud. Fig. 1(c) shows a preliminary denoising result after implementing MEOR-G. Despite that, the point cloud may still suffer from scattered noisy points [as identified in white in Fig. 1(c)] close to the point cloud of the ground. The following step, i.e., MEOR-L, further identifies scattered outliers retained from the result of MEOR-G. In this step, the preliminary result of MEOR-G is divided into clusters, and the schematic representation of clusters is shown in Fig. 1(d). MEOR-L determines the elevation threshold T' for each cluster as shown in Fig. 1(f). It thus removes noisy points based on the maximum entropy principle, which is similar to that of MEOR-G. The coarse and refined results of denoising for each cluster are represented in Fig. 1(e) and (g), respectively. Ultimately, the results after running MEOR-L of all clusters are combined into a resulting point cloud. The final denoising result is shown in Fig. 1(h).

B. Maximum Entropy Principle

For a given system, Shannon [33] proposed to extend the concept of entropy to measure the statistical information uncertainty. The calculated entropy value represents the amount of information contained in a variable. If a system possesses n discrete states, the entropy of the whole system can be calculated based on the probability distribution of all states

$$S = -\sum_{i=1}^{n} p_i \log(p_i) \tag{1}$$

and

$$\sum_{i=1}^{n} p_i = 1,\ 0 \le p_i \le 1 \tag{2}$$

where p_i represents the probability of state i. The value of the entropy S varies according to the distribution of the discrete states. The greater uncertainty and amount of information existing in the system lead to a larger value of S of the variable.

As attributed to the merit of quantifying the amount of information contained in a given system, the concept of entropy can be further generalized to other applications [32]. The maximum entropy principle states that the uncertainty and accumulated entropy can reach the peak if the probabilities of all the states tend to be equal. In this case, the uniform distribution of states can maximize the accumulated entropy and realize the largest uncertainty and amount of information. Laplace's principle of insufficient reason further supports this argument [34], which

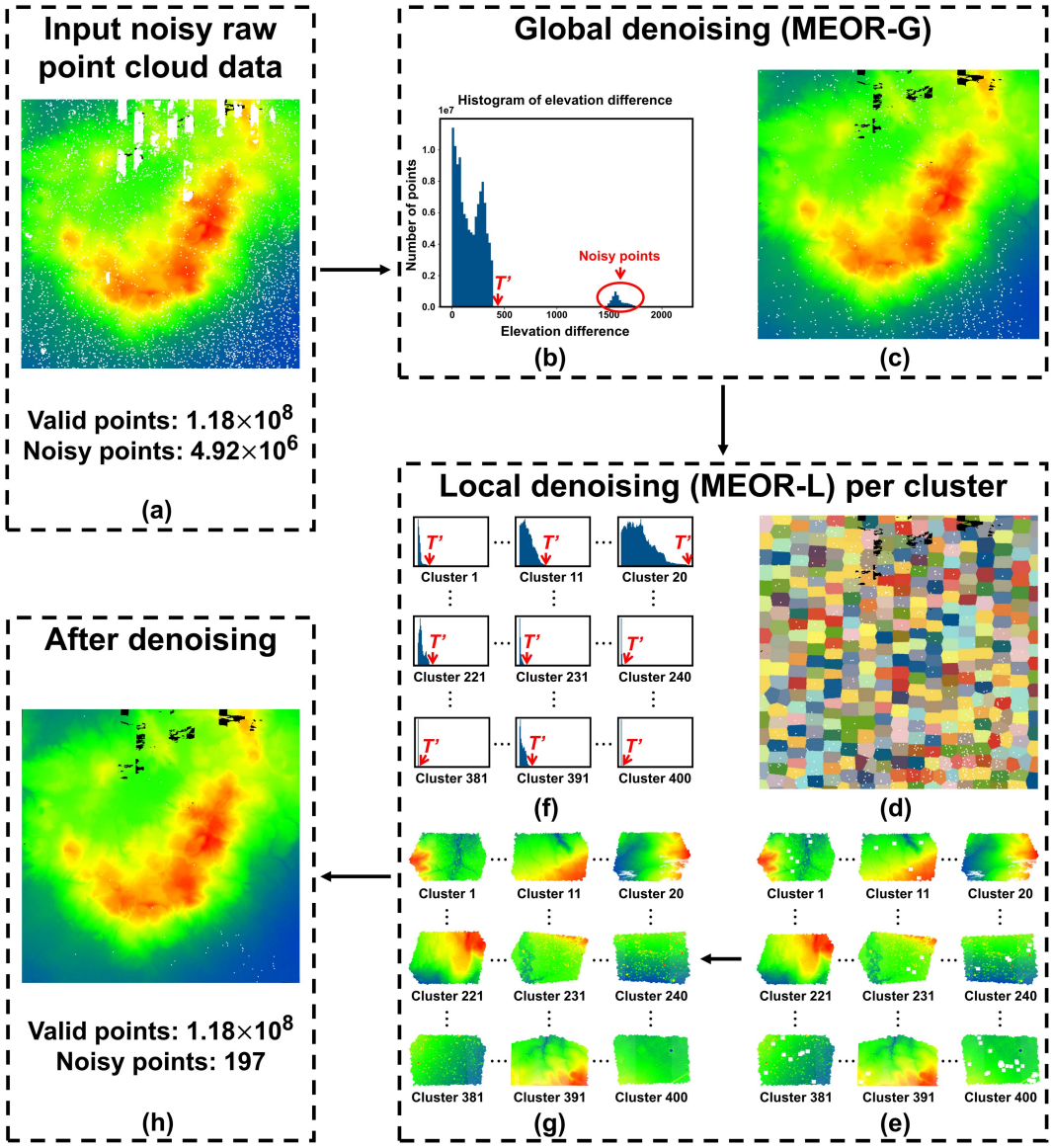


Fig. 1. Overall workflow of the proposed MEOR that includes (a) input point cloud with noisy points, (b) histogram of elevation difference for global point cloud, (c) coarse and global OR result, (d) schematic representation of clusters, (e) point cloud of each cluster, (f) histogram of elevation difference for point cloud within each cluster, (g) fine and local OR result for each cluster, and (h) final result.

states that the best strategy to discriminate between two or more events is to consider all equally distributed states in each event.

In MEOR, we apply the maximum entropy principle to discriminate noisy points and valid points with elevation values of the point cloud as information of the system. The two "events" of valid and noisy point clouds existing in this system usually possess the corresponding uniform density distribution within each "event." Therefore, an elevation threshold maximizing the accumulated entropy of all points' elevation can differentiate the valid and noisy point clouds.

C. Maximum Entropy-Based Outlier Removal (MEOR) Method

First, MEOR calculates the average elevation of the given dataset as the benchmark elevation $\bar{z}$. Then it generates the vector of absolute elevation difference $\mathbf{Z}_\delta = |\mathbf{Z}_n - \bar{z}|$ between elevation vector $\mathbf{Z}_n$ of the entire point cloud and the benchmark elevation $\bar{z}$. n is the total number of points in the dataset. Next, the vector of elevation absolute difference $\mathbf{Z}_\delta$ is divided into l levels, i.e.,

$$n = \sum_{i=1}^{l} n_i \tag{3}$$

where n_i is the number of points with elevation difference belonging to the ith level. Then the probability p_i of the i level is calculated according to the following, i.e.,

$$p_i = \frac{n_i}{n}, \quad 1 \le i \le l. \tag{4}$$

The strategy to determine the value of parameter l is related to the scale of elevation difference $\mathbf{Z}_\delta$. For instance, if $\mathbf{Z}_\delta$ ranges from

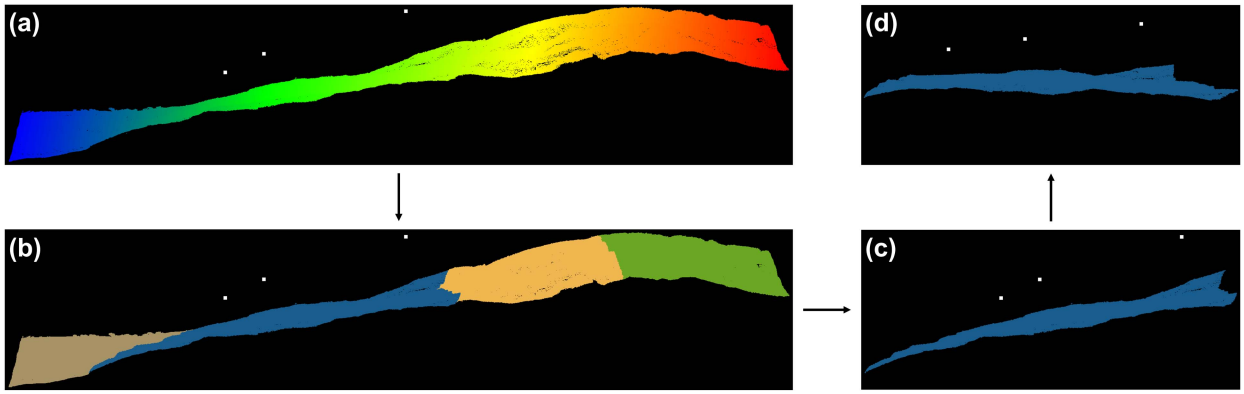


Fig. 2. Illustration of point cloud clustering for MEOR-L that includes (a) an example of point cloud with great terrain relief, (b) clustering result, (c) a cluster with noisy points, and (d) rotated cluster parallel with the ground surface.

0 to 100 m, we can divide the dataset into 20 levels with a 5 m range, i.e., $l = 20$. Before calculating the elevation threshold to distinguish noisy points and valid points, we determine whether noisy points exist in the point cloud. If the number of consecutive levels without points is larger than one given threshold, e.g., five, then we consider the point cloud to be noise-free. Otherwise, for the noisy point cloud, we assign the first level of l as an initial threshold (the threshold level $t = 1$). Then, the corresponding elevation threshold can be computed as $T = \bar{z} + \frac{\mathbf{Z}_{\delta\max}}{l} \times t;\ t \in [1, l]$, where $\mathbf{Z}_{\delta\max}$ is the maximum element of vector $\mathbf{Z}_\delta$, to divide $\mathbf{Z}_\delta$ into two parts ($\widetilde{\mathbf{Z}}^\Psi$ and $\widetilde{\mathbf{Z}}^\Upsilon$) according to the absolute elevation difference between $\bar{z}$ and the elevation of the point difference. The points with the difference smaller than T are categorized as initial valid point cloud $\widetilde{\mathbf{Z}}^\Psi$. Otherwise, those points can be categorized as initial noisy point cloud $\widetilde{\mathbf{Z}}^\Upsilon$.

Subsequently, Shannon entropy for the initial valid point cloud $\widetilde{S}_t^\Psi$ and initial noisy point cloud $\widetilde{S}_t^\Upsilon$ with specific t can be respectively calculated based on (1). The entropy's summation of valid points and noisy points can then be expressed as $S_t^{\Psi+\Upsilon} = \widetilde{S}_t^\Psi + \widetilde{S}_t^\Upsilon$. t is then raised by one and $S_t^{\Psi+\Upsilon}$ is subsequently calculated for each corresponding t. If t is no larger than l, then repeat until t is equal to l. According to the maximum entropy principle, the threshold level which can maximize the entropy's summation is regarded as the optimal threshold level $t' = \arg\max_{t\in l}(S_t^{\Psi+\Upsilon})$ dividing the original point cloud into valid point cloud $\mathbf{Z}^\Psi$ and noisy point cloud $\mathbf{Z}^\Upsilon$. The optimal elevation threshold can then be derived as $T' = \bar{z} + \frac{\mathbf{Z}_{\delta max}}{l} \times t'$. Finally, $\mathbf{Z}^\Psi$ and $\mathbf{Z}^\Upsilon$ are determined by comparing the absolute elevation difference between $\bar{z}$ and $\mathbf{Z}_n$ with the value of T' based on the following (5). If the absolute elevation differences are no larger than T', then the studied points can be regarded as $\mathbf{Z}^\Psi$, otherwise, the points are determined as $\mathbf{Z}^\Upsilon$

$$\mathbf{Z}_n = \begin{cases} \mathbf{Z}^\Psi, & \text{if } |\mathbf{Z}_n - \bar{z}| \le T' \\ \mathbf{Z}^\Upsilon, & \text{if } |\mathbf{Z}_n - \bar{z}| > T'. \end{cases} \tag{5}$$

D. Two-Stage Global-Local or for Point Cloud

The global stage MEOR-G removes the noisy points significantly distant from the valid points. Based on the predicted elevation threshold for a given dataset using the maximum entropy principle, MEOR-G can distinguish and remove the noisy point cloud from the valid point cloud, so the remaining point cloud is considered coarsely denoised. However, due to the terrain relief generated by land cover objects on the ground, e.g., mountains, buildings, trees, etc., some scattered noisy points close to the valid point cloud cannot be identified according to a single value of global elevation threshold. Therefore, the following MEOR-L stage generates a cleaner point cloud and preserves the details of the valid point cloud. In this stage, the initial coarse denoised point cloud is divided into clusters based on the region growing algorithm described in Section II-E. Subsequently, the whole MEOR procedure is repeated to calculate the local elevation threshold and identify noisy points for each individual cluster. Ultimately, we combine all clusters' refined denoising results and obtain a point cloud that is considered free of noise.

E. Point Cloud Clustering for MEOR-L

Fig. 2(a) shows an example that point clouds are acquired from a mountainous area with great topographic relief and noisy points are located close to the mountain. In this case, it is invalid to remove noisy points [the white scattered points in Fig. 2(a)] if we directly exploit the proposed MEOR method. The solution is to divide the point cloud into flat clusters, which are shown in different colors in Fig. 2(b) and rotate the point cloud cluster, e.g., the blue cluster in Fig. 2(c), to parallel with the ground surface shown in Fig. 2(d). At last, we can address the point cloud denoising by using MEOR on the rotated cluster in Fig. 2(d).

The basic idea of point cloud clustering is to merge the points that are close enough according to a smoothness constraint. Therefore, the result of this step is a set of clusters, where the points within each cluster are considered to belong to the identical approximated smooth surface. The clustering is based on region growing algorithm, the principle of which is the comparison of the normal angles between the neighbor points. The number of searched neighbor points η, normal angle difference threshold θ_{th}, curvature threshold c_{th}, and maximum number of cluster $c_{\max}$ are predefined thresholds. The details of the clustering method are shown in the following:

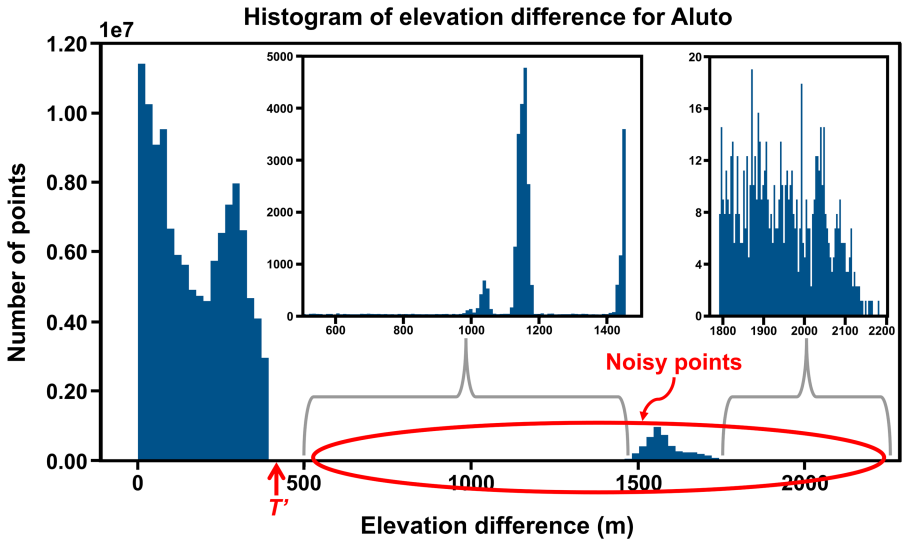


Fig. 3. Histogram of elevation difference for dataset Aluto.

Step 1: It begins with the point with minimum curvature value as the first seed point and adds it to the seed set. The region grows from the point with minimum curvature that is located in the flat area to reduce the total number of clusters.

Step 2: The algorithm searches η corresponding nearest neighbor points for individual seed points and measures the normal angle difference between the current seed point and its each neighbor point. Points with the difference larger than the given threshold θ_{th} is excluded from the current region. Otherwise, this neighbor point is popped to the current region from the available point set.

Step 3: A neighbor point already added to the current region is selected into the seed set if the associated curvature value is smaller than a threshold c_{th}.

Step 4: Steps 2 and 3 are repeated for all the seed points iteratively until the seed set is empty or the number of points within the current region reaches $c_{\max}$. For now, only one cluster should be formed.

Step 5: The algorithm returns to step 1 and executes steps 2 to 4 repeatedly and the algorithm stops when the available point set is empty. Ultimately, points are clustered into regions with approximately flat surfaces.

Step 6: Due to the terrain relief, most of the formed regions are not parallel to the ground surface. Since our proposed MEOR is implemented by the elevation information of the point cloud, the titled flat regions should be rotated until they become parallel with the ground surface. The angle difference θ between z-axis and the normal vector of the fitting plane for the points within each region is computed, so that the fitting plane can be adjusted to become parallel to the xy-plane.

III. Experiments

A. LiDAR Datasets

The proposed MEOR is tested using four airborne LiDAR datasets suffering from different scenarios of noisy points. The first dataset was originally collected by an airborne survey for the Aluto volcanic complex in Ethiopia to study the fluid pathways [35]. The coverage of the point cloud for this dataset is shown inside the red box at the top of Fig. 4(a), where the dataset suffers from patches of clouds and smoke as shown at the bottom of Fig. 4(a).

The second dataset was collected in 2012 on the Kameni islands, Santorini, Greece, for mapping the subaerial and submarine extensions of lava flows [36]. The corresponding area is shown at the top of Fig. 6(a), where a number of notable cumulus clouds are shown at the bottom of Fig. 6(a).

The third dataset covers a study region located in the British Columbia Institute of Technology, Vancouver, Canada, as shown at the top of Fig. 8(a). This dataset is contaminated by noise likely caused by ambiguous range measurements in multiple-pulses-in-air technique [37], leading to a notable string of noisy points hovering over the point cloud.

Finally, the fourth dataset covers a shore region near Scarborough, ON, Canada, including various land covers, e.g., tree canopies, houses, grassland, fences, and water region as shown in the red box at the top of Fig. 10(a). This dataset contains scattered noisy points located near the ground data points, as shown in white, at the bottom of Fig. 10(a).

In these four datasets, all the noisy points and valid points are manually labeled to serve the subsequent evaluation. The detailed information of each dataset, including the total number of points, the number and the corresponding proportion of noisy points, the number of valid points, and more information about the LiDAR sensor, e.g., flying height, pulse repetition frequency, and scan frequency can be found in Table I.

B. Evaluation Metrics

Since the four abovementioned datasets have been manually labeled with noisy and valid data points, we consider four

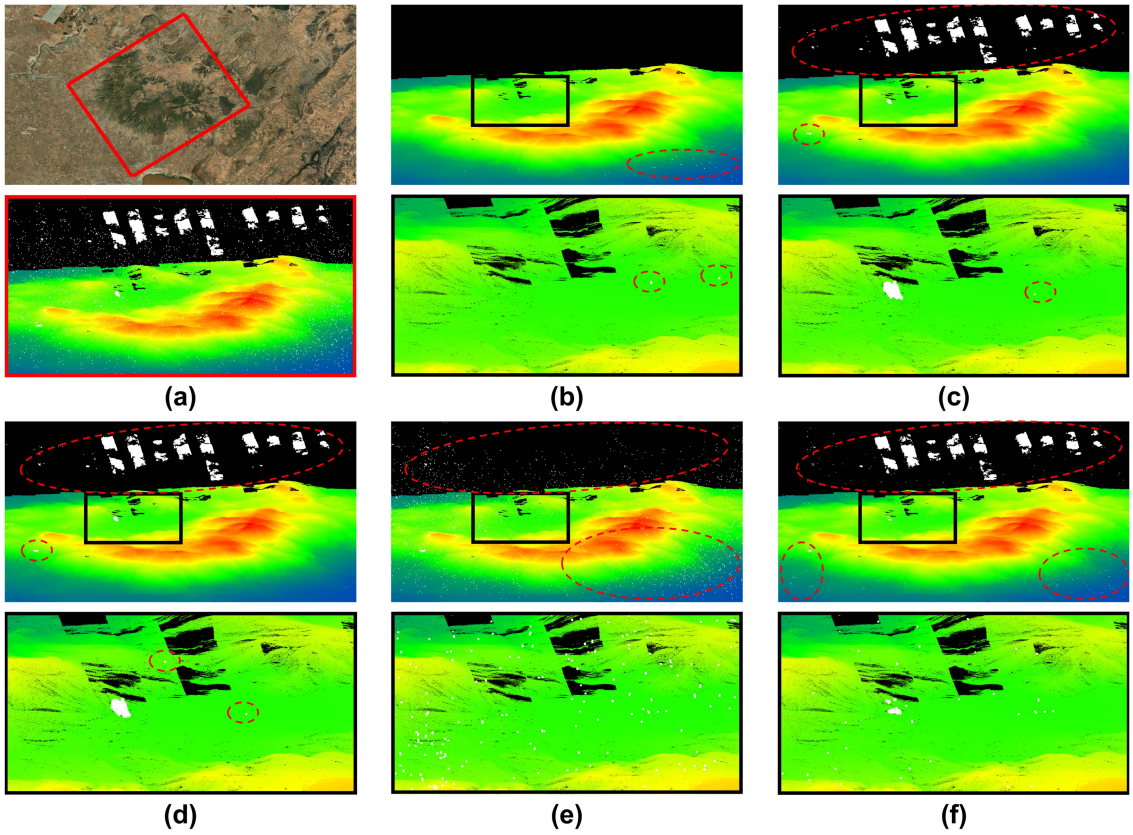


Fig. 4. (a) Raw point cloud, results of different OR methods, including (b) MEOR, (c) SFOR, (d) SOR, (e) SHDR, and (f) MOR in dataset Aluto.

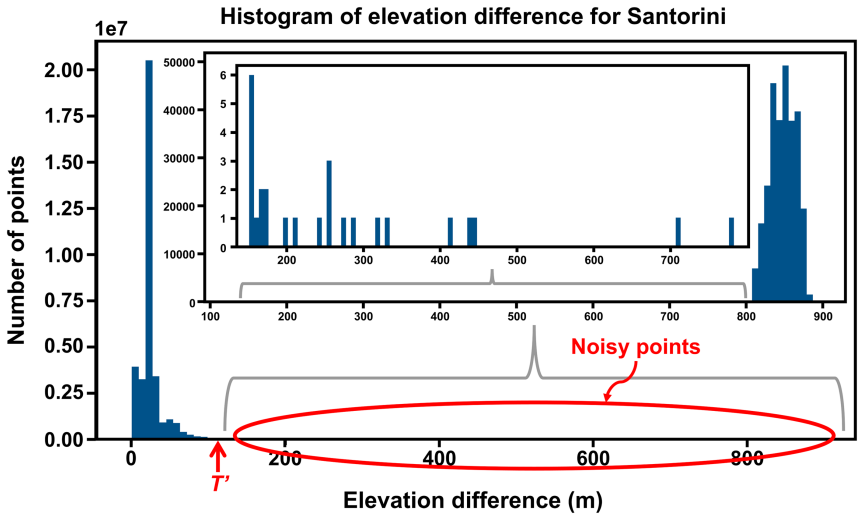


Fig. 5. Histogram of elevation difference for dataset Santorini.

commonly used metrics, i.e., recall (R), precision (P), overall accuracy (OA), and F1 score (F1) to assess the point cloud denoising results. They are calculated according to the corresponding True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN)

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{6}$$

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{7}$$

$$\mathrm{OA} = \frac{\mathrm{TP+TN}}{\mathrm{TP++TN+FP+FN}}. \tag{8}$$

Then, we can calculate the F1 score as

$$\mathrm{F1} = 2 \times \frac{\mathrm{P} \times \mathrm{R}}{\mathrm{P} + \mathrm{R}} \tag{9}$$

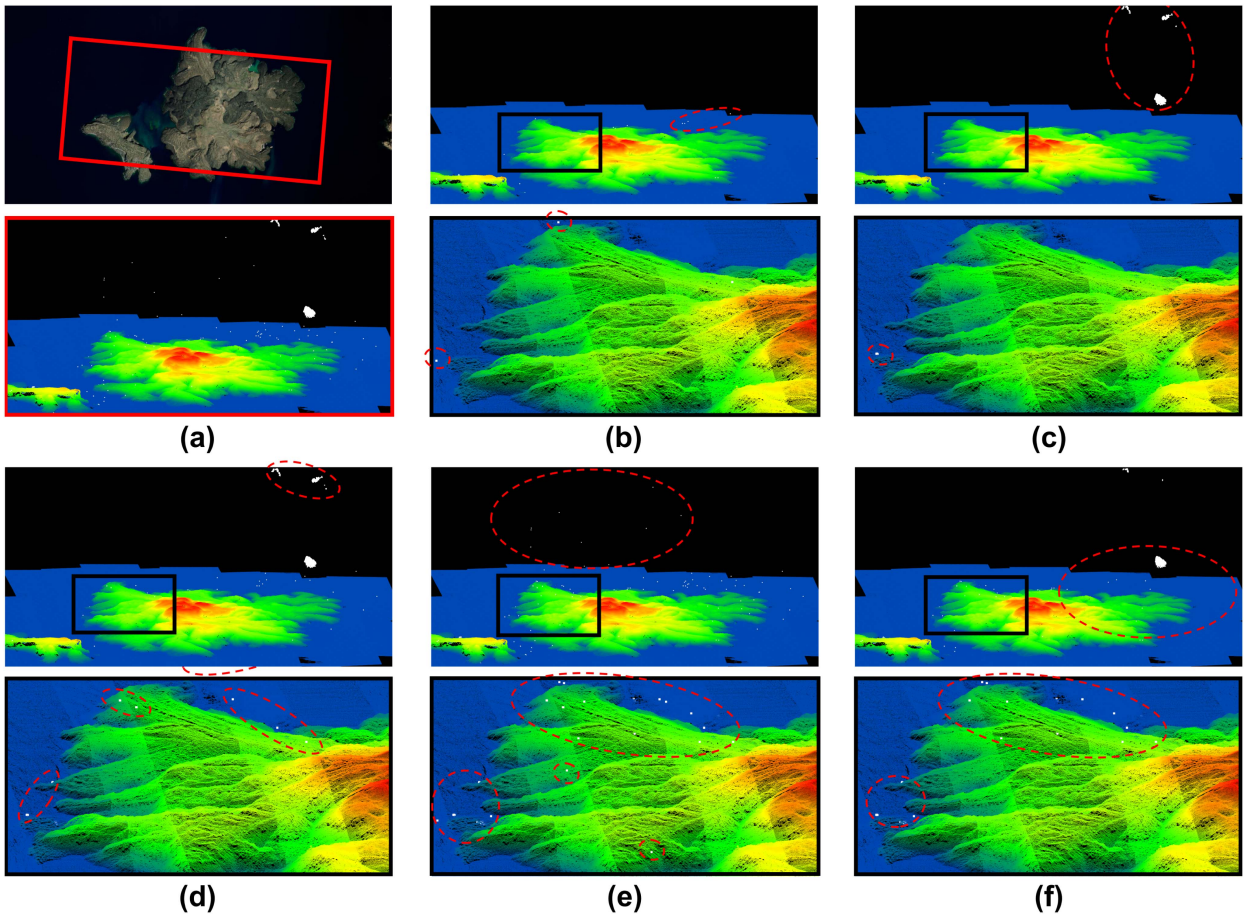


Fig. 6. (a) Raw point cloud, results of different OR methods, including (b) MEOR, (c) SFOR, (d) SOR, (e) SHDR, and (f) MOR in dataset Santorini.

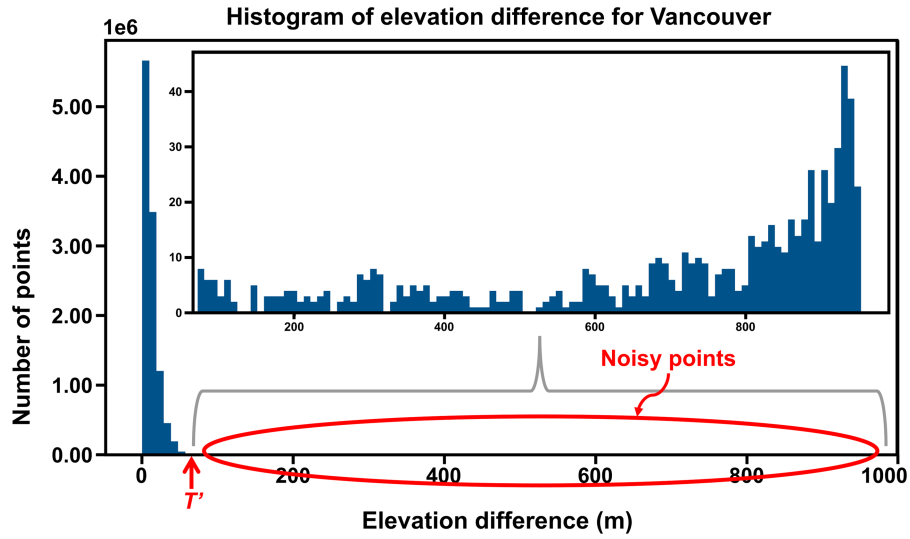


Fig. 7. Histogram of elevation difference for dataset Vancouver.

where TP represents the total number of noisy points that are correctly predicted as noisy points, FN refers to the number of noisy points misidentified as valid points, FP means the number of valid points wrongly identified as noisy points, and TN stands for the number of valid points correctly identified as valid points.

C. Experimental Settings

To justify the efficiency of the proposed MEOR, four existing point cloud denoising methods mentioned in Section I, SFOR [19], SOR [20], SHDR [21], and MOR [22], are included

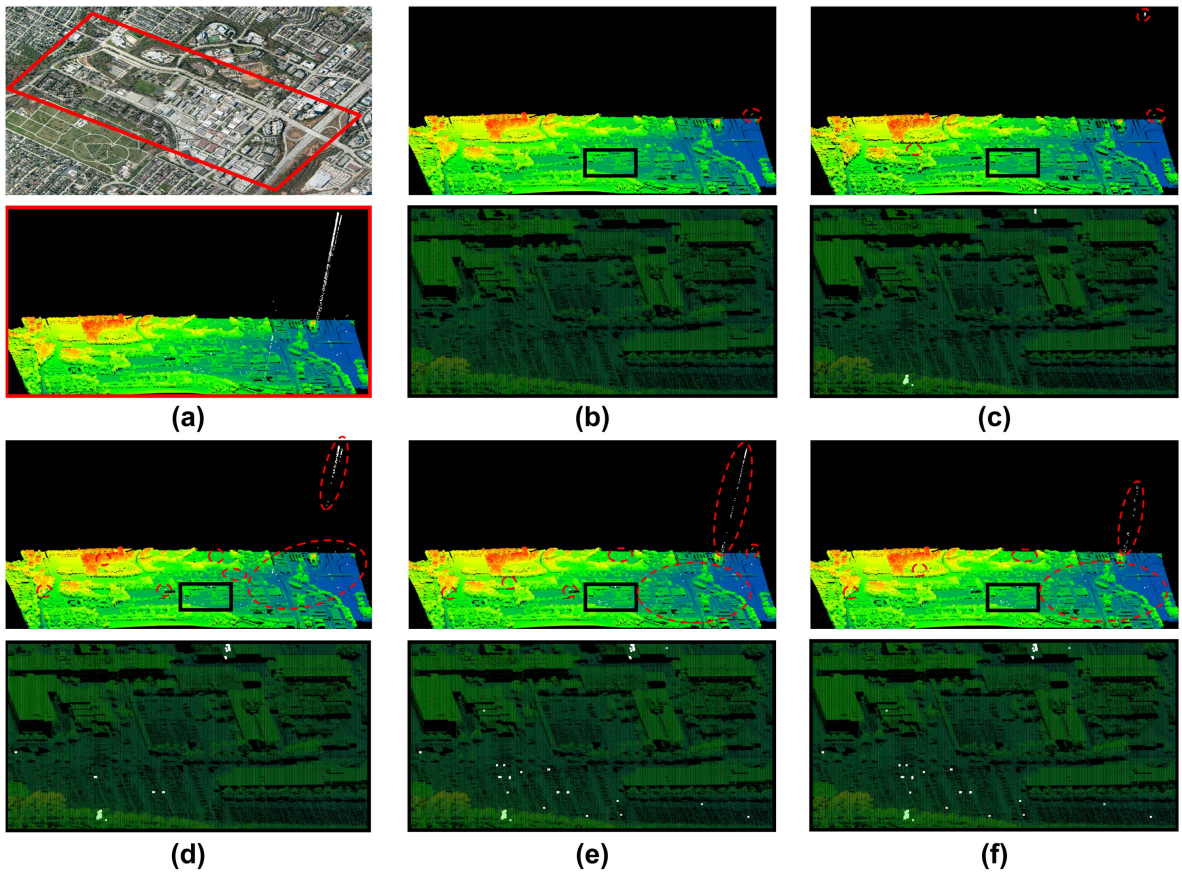


Fig. 8. (a) Raw point cloud, results of different OR methods, including (b) MEOR, (c)SFOR, (d) SOR, (e) SHDR, and (f) MOR in dataset Vancouver.

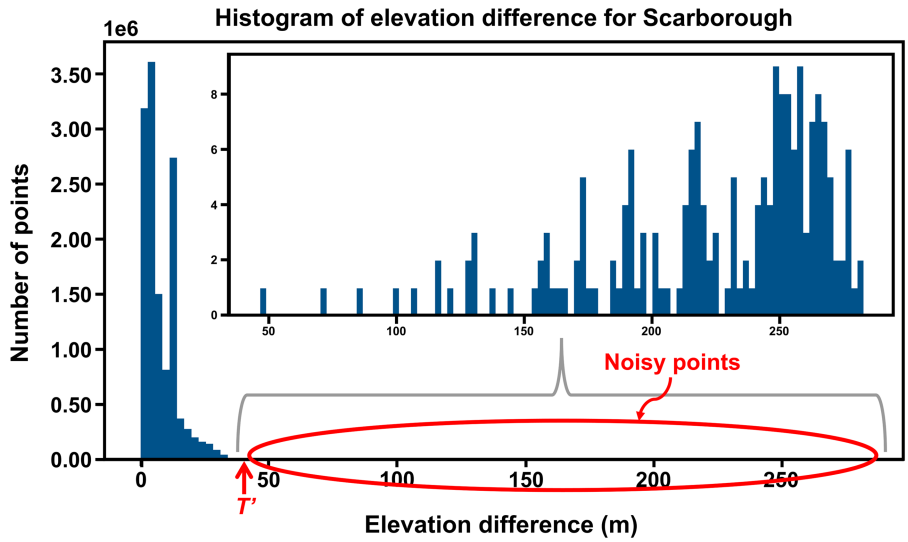


Fig. 9. Histogram of elevation difference for dataset Scarborough.

for comparison. As our proposed MEOR and these four methods are based on different principles and possess corresponding specific parameters, here we predefine and set the values of corresponding parameters for each of the individual methods.

Maximum Entropy based Outlier Removal (MEOR): The predefined experimental parameters of our proposed MEOR include the number of clusters that are segmented from the global coarse denoised point cloud and parameter l that divides the absolute elevation difference $\mathbf{Z}_\delta$ into l levels for maximum entropy principle mentioned in (3). We set the values of four parameters in the clustering step of MEOR-L according to $\eta = 5$, $\theta_{\text{th}} = 2.0$, $c_{\text{th}} = 0.00001$, and $c_{\max} = 100$, respectively,

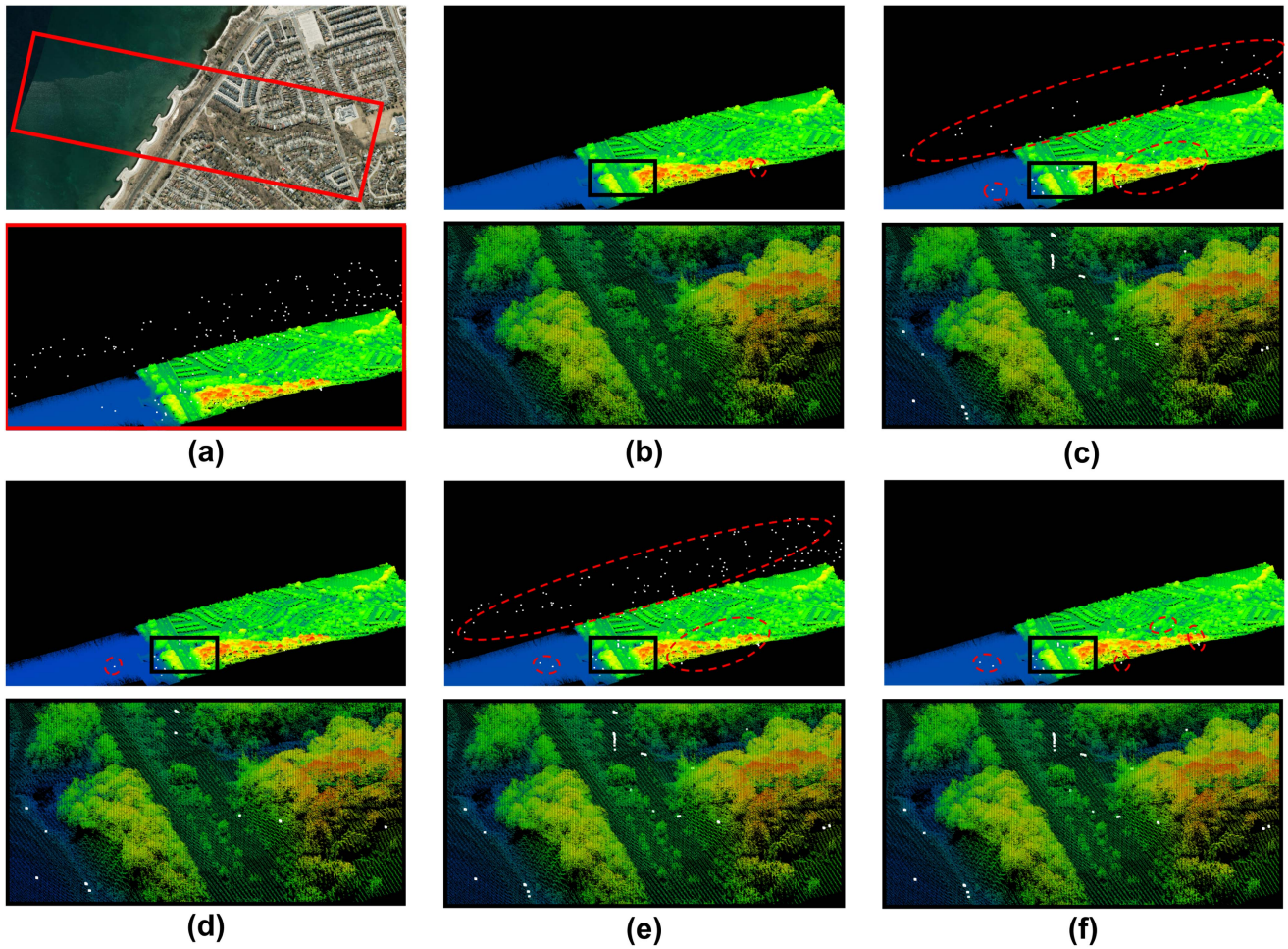


Fig. 10. (a) Raw point cloud, results of different OR methods, including (b) MEOR, (c) SFOR, (d) SOR, (e) SHDR, and (f) MOR in dataset Scarborough.

TABLE I
CHARACTERISTICS AND INFORMATION FOR FOUR AIRBORNE LIDAR DATASETS

Datasets	Type of noisy points	Type of cover area
Aluto	Cluster & Scatter	Dormant stratovolcano complex
Santorini	Cluster & Scatter	Island formed by lava flows
Vancouver	Cluster & Scatter	Campus of urban area
Scarborough	Scatter	Shore region of urban area

Datasets	OP	NP	VP	Sensor	Flying height	Repetition frequency	Scan frequency
Aluto	123,770,644	4,921,256 (3.976%)	118,849,388	Leica ALS50	-	46.2 kHz	29.9 Hz
Santorini	35,491,876	282,080 (0.795%)	35,209,796	Leica ALS50	2,000 m	94.7-119 kHz	55.7-58.2 Hz
Vancouver	11,090,291	995 (0.009%)	11,089,296	Leica ALS50	600 m	83 kHz	-
Scarborough	13,255,997	527 (0.004%)	13,255,470	Optech Titan	430 m	200 kHz,	40 Hz

*OP = Original number of data points, NP = Number of noisy points, and VP = Number of valid points.

which are determined by ablation study in Section V-D. The number of clusters varies among four datasets, i.e., 400, 250, 200, and 200 for Aluto, Santorini, Vancouver, and Scarborough, respectively, according to the corresponding total number of points. In terms of the value of parameter l, the experimental results show that when we set it to 90, it provides the best performance discriminating noisy points and valid points and consumes relatively lower computational time to accomplish the task.

Spatial Frequency Outlier Removal (SFOR): As described in this method, neighbor points within a specific radius of the sphere are constructed for each point in the LiDAR dataset. If the number of neighbor points within such a sphere is inferior to a certain threshold, the current point is considered noisy. We set

the radius of the sphere as 2 m since it is significantly larger than the mean point spacing of all four datasets [19]. Accordingly, the threshold radius is set as 2 m for all datasets.

Statistical Outlier Removal (SOR): For each point in the dataset, k nearest neighbor points are queried. The mean distance of each point to its k neighbor points is then calculated. The noisy points are identified using the sigma rule based on the entire dataset, that is, for one point, if the mean distance of its k neighbor points is not within N standard deviations from the mean distance, then this point is treated as a noisy point. We set the parameter k as six and parameter N as one in the experiments for all datasets [20].

Spatial Hierarchical Directional Relationship: The setting of the distance threshold determines whether the noisy points have a spatial connectivity relationship to the valid point set. If the distance threshold is set to an unsuitable value, then the noisy points are unlikely detected. The paper [21] shows that the optimal distance threshold is set to 3 m, which can achieve the best detection effect. Therefore, the distance threshold is determined as 3 m for all datasets.

Morphological Outlier Removal (MOR): MOR requires first transforming the point cloud into grids and then conducting the opening operation, a composite of two other fundamental morphological operations: erosion and dilation. The erosion operation essentially lowers the elevated points to match the lowest point within the grid defined by the structuring element. Conversely, the dilation operation raises the lower points to align with the highest point within the corresponding structuring element. After erosion and dilation, the transformed elevation of points lower than the predefined elevation threshold are considered noisy points, otherwise, the points are considered valid points. We define the grid in this method as a cylinder and select neighbor points located within the cylinder. The radius of the cylinder is set to 1 m for all four LiDAR datasets [22]. The mean elevation of the extracted ground point cloud is set as the predefined elevation threshold for this denoising method.

IV. Quantitative and Qualitative Results

Our proposed MEOR, together with the SFOR, SOR, SHDR, and MOR methods, were applied to these four datasets, respectively. Visual interpretation of the point cloud denoising results and quantitative evaluations both provide useful insights to assess denoising performance. Four point cloud datasets are located inside the red box at the top of Figs. 4(a), 6(a), 8(a), and 10(a), respectively. There are two major categories of noisy points, i.e., clustered noisy points and scattered noisy points in white as shown at the bottom of Figs. 4(a), 6(a), 8(a), and 10(a), respectively. Our proposed MEOR method aims to look for an elevation threshold for the point cloud to distinguish noisy points from valid points. Here, we use the histogram of global elevation difference to illustrate MEOR's effectiveness in discriminating between noisy and valid points. Figs. 3, 5, 7, and 9 show the histograms of the global elevation difference of four datasets and the parts that cannot be clearly shown in the whole histogram are further enlarged at the top of the histogram. The global elevation thresholds T' estimated by the maximum entropy principle are 457.559 m, 116.260 m, 66.720 m, and 42.453 m for the Aluto, Santorini, Vancouver, and Scarborough datasets, respectively. Those points with elevation differences larger than the estimated thresholds are regarded as noisy points (as emphasized by a red ellipse).

A. Experiments in Dataset Aluto

Compared with SFOR, SOR, SHDR, and MOR, our proposed MEOR demonstrates an outstanding performance in eliminating noisy points while preserving valid points. Table II shows the results of four point cloud denoising methods. MEOR-G accurately identifies 99.972% of noisy points, meanwhile, no valid point is incorrectly removed. The subsequent MEOR-L further removes noisy points for individual local clusters, leading to 99.996% of noisy points being removed and 242 valid points being wrongly identified as noisy points. Among the four compared methods, the performance of SHDR is comparable with our MEOR and all valid points completely remain. However, our method still generates a better denoising result with 99.996% noisy points removed compared with 99.855% noisy points detected by SHDR. Although the performance of MOR is much worse than our proposed method, it still slightly surpasses SOR and SFOR. 23.858% of noisy points are removed and only 327 valid points are misidentified as noisy points. SOR and SFOR can only identify a small portion of noisy points and most noisy points remain, i.e., only 1.407% and 3.594% of noisy points are removed, respectively. In terms of preserving valid details of the point cloud, SOR and SFOR generate results by wrongly removing 0.086% and 0.746% of valid points, respectively. Our proposed method consistently achieves higher recall, precision, accuracy, and F1 score compared with other methods. The MEOR-L obtains a slight improvement of 0.024% and 0.010% in terms of recall and F1 score compared with MEOR-G. However, the precision is slightly reduced by 0.005%. SHDR produces a higher precision (i.e., 100%) than that of our method, while the recall, accuracy, and F1 score of SHDR are slightly lower than our method with 99.855%, 99.994%, and 99.927%. MOR can also generate a high precision and accuracy of 99.972% and 96.972%, respectively. On the contrary, the recall and F1 score are low with 23.858% and 38.522%, respectively. SOR and SFOR can achieve 95.997% and 95.460% in terms of accuracy. Nevertheless, the recall and F1 score are found in between 1.407% to 5.922%, indicating a very poor performance. Precisions of SOR and SFOR are 40.329% and 16.815%, respectively, which are much lower than those of SHDR, MOR, and our proposed MEOR.

The denoising results of these four methods are shown in Fig. 4(b)–(f) to compare the results of different OR methods. Our proposed MEOR can remove all the clustered noisy points and only a small number of scattered noisy points close to the ground may be missed. On the other hand, SHDR is useful to remove clustered noisy points but invalid to remove scattered noisy points. The other three existing denoising methods show poor performance in removing isolated noisy points.

TABLE II
DENOISING PERFORMANCE OF OUR PROPOSED MEOR AND COMPARISON WITH SFOR, SOR, SHDR, AND MOR IN ALUTO DATASET

Methods	ONP	RNP	OVP	RVP	R (%)	P (%)	A (%)	F1 (%)
MEOR-G	4,921,256	4,919,854	118,849,388	0	99.972	100.00	99.999	99.986
MEOR-G + MEOR-L	4,921,256	4,921,059	118,849,388	242	99.996	99.995	99.999	99.996
SFOR [19]	4,921,256	176,858	118,849,388	874,936	3.594	16.815	95.460	5.922
SOR [20]	4,921,256	69,220	118,849,388	102,418	1.407	40.329	95.997	2.718
SHDR [21]	4,921,256	4,914,129	118,849,388	0	99.855	100.00	99.994	99.927
MOR [22]	4,921,256	1,174,105	118,849,388	327	23.858	99.972	96.972	38.522

*ONP = Original noisy points, RNP = Removed noisy points, OVP = Original valid points, RVP = Removed valid points , R = Recall, P = Precision, A = Accuracy, and F1 = F1 score.

TABLE III
DENOISING PERFORMANCE OF OUR PROPOSED MEOR AND COMPARISON WITH SFOR, SOR, SHDR, AND MOR IN SANTORINI DATASET

Methods	ONP	RNP	OVP	RVP	R (%)	P (%)	A (%)	F1 (%)
MEOR-G	282,080	281,697	35,209,796	1	99.864	99.999	99.999	99.932
MEOR-G + MEOR-L	282,080	281,949	35,209,796	60	99.954	99.980	99.999	99.967
SFOR [19]	282,080	392	35,209,796	1,884	0.139	17.223	99.201	0.276
SOR [20]	282,080	16,459	35,209,796	1,803,997	5.835	0.904	94.169	1.566
SHDR [21]	282,080	281,597	35,209,796	483	99.828	100.00	99.998	99.914
MOR [22]	282,080	89,143	35,209,796	352	31.602	99.607	99.455	47.981

*ONP = Original noisy points, RNP = Removed noisy points, OVP = Original valid points, RVP = Removed valid points , R = Recall, P = Precision, A = Accuracy, and F1 = F1 score.

B. Experiments in Dataset Santorini

Table III presents and compares the denoising results of the dataset Santorini derived by five methods. Our proposed MEOR removes most noisy points, leaving a tiny portion of valid points being misidentified and removed. MEOR-G removes 99.864% of noisy points and one valid point. MEOR-L further removes noisy points, leading to a final result of 99.954% of noisy points and 60 valid points being removed. SHDR shows a slightly worse result than our method by removing 99.828% noisy points and wrongly removing 483 valid points. By contrast, a vast majority of noisy points cannot be precisely identified, i.e., only 0.139%, 5.835%, and 31.602% of noisy points are removed by SFOR, SOR, and MOR, respectively. A total of 0.005% and 0.001% of valid points are incorrectly removed by SFOR and MOR, and 5.124% of valid points are misclassified into noisy points and removed when we exploit the SOR method. MEOR-L achieves high scores in all four metrics and generates slight improvements of 0.09% in recall and 0.035% in F1 score based on the coarse denoising results by MEOR-G. A slight reduction of 0.019% in precision is found compared with the coarse denoising result. SFOR, SOR, SHDR, and MOR can exhibit robust performance in terms of accuracy with 99.201%, 94.169%, 99.998%, and 99.455%, respectively. SHDR and MOR are able to produce high precision of 100% and 99.607% compared with low precision of 17.223% and 0.904% generated by SFOR and SOR. Except for SHDR producing high recall and F1 score with 99.828% and 99.914%, all three other methods come with a poor performance in recall and F1 scores, ranging from 0.139% to 47.981%.

The denoising results of the Santorini dataset based on five methods are presented in Fig. 6(b)–(f) for visual comparison. Similar to the results for the Aluto dataset, our proposed method surpasses four other existing methods in removing clustered noisy points. The subsequent MEOR-L is able to further remove isolated noisy points close to the ground point cloud on the basis of coarse denoising result; however, SFOR, SOR, SHDR, and MOR all retain more noisy points close to the ground than MEOR.

C. Experiments in Dataset Vancouver

As shown in Fig. 7(a), the noisy points of this dataset are distributed as a string of points, which may be induced by ambiguous range measurements. The number of noisy points is 995, which only accounts for 0.009% of the entire dataset. The clustered noisy points in this dataset are close to the ground point cloud presented in white at the bottom of Fig. 8(a), which are different from those located far from the ground point cloud in the datasets of Aluto and Santorini.

The denoising results of five methods are presented in Table IV. MEOR-G applied to this dataset does not perform well compared to the scenarios of Aluto and Santorini, i.e., nearly one-third of noisy points are retained after implementing MEOR-G. MEOR-L stage subsequently solves the problem and removes 98.593% of noisy points, leaving only 14 noisy points retained. Our method performs well in preserving valid data points and only 44 valid points are removed in the final result. SOR outperforms SFOR, SHDR, and MOR in removing noisy points, i.e., 89.849%, 50.754%, 58.492%, and 66.633% of noisy points are identified, respectively. However, SOR generates the worst performance in preserving valid points. A total of 8.455% of valid points are misidentified and removed by SOR, which is much higher than 0.404%, 0.004%, and 0.009% of valid points being removed by SFOR, SHDR, and MOR. There are

TABLE IV
DENOISING PERFORMANCE OF OUR PROPOSED MEOR AND COMPARISON WITH SFOR, SOR, SHDR, AND MOR IN VANCOUVER DATASET

Methods	ONP	RNP	OVP	RVP	R (%)	P (%)	A (%)	F1 (%)
MEOR-G	995	689	11,089,296	26	69.246	96.364	99.997	80.585
MEOR-G + MEOR-L	995	981	11,089,296	44	98.593	95.707	99.999	97.129
SFOR [19]	995	505	11,089,296	44,764	50.754	1.116	99.592	2.183
SOR [20]	995	894	11,089,296	937,569	89.849	0.095	91.545	0.190
SHDR [21]	995	582	11,089,296	413	58.492	100.00	99.996	73.811
MOR [22]	995	663	11,089,296	1,021	66.633	39.371	99.988	49.496

*ONP = Original noisy points, RNP = Removed noisy points, OVP = Original valid points, RVP = Removed valid points , R = Recall, P = Precision, A = Accuracy, and F1 = F1 score.

TABLE V
DENOISING PERFORMANCE OF OUR PROPOSED MEOR AND COMPARISON WITH SFOR, SOR, SHDR, AND MOR IN SCARBOROUGH DATASET

Methods	ONP	RNP	OVP	RVP	R (%)	P (%)	A (%)	F1 (%)
MEOR-G	527	189	13,255,470	0	35.863	100.00	99.997	52.793
MEOR-G + MEOR-L	527	463	13,255,470	0	87.856	100.00	99.999	93.535
SFOR [19]	527	130	13,255,470	1,378	24.668	8.621	99.987	12.776
SOR [20]	527	319	13,255,470	413,421	60.531	0.077	96.880	0.154
SHDR [21]	995	0	11,089,296	0	0.000	0.000	99.996	0.000
MOR [22]	527	220	13,255,470	0	41.746	100.00	99.998	58.902

*ONP = Original noisy points, RNP = Removed noisy points, OVP = Original valid points, RVP = Removed valid points , R = Recall, P = Precision, A = Accuracy, and F1 = F1 score.

significant improvements, i.e., 29.347% and 16.544% of recall and F1 score, respectively, by conducting MEOR-L compared to the intermediate results of MEOR-G. MEOR-L produces a slight reduction of 0.657% in precision and slightly rises a 0.002% of accuracy on the result derived by the MEOR-G stage. SFOR, SHDR, and MOR yield an accuracy of 99.592%, 99.996%, and 99.988%, while SOR registered a notably lower accuracy with a value of 91.545%. SFOR, SOR, SHDR, and MOR present acceptable recalls of 50.754%, 89.849%, 58.492%, and 66.633%, respectively, whereas a poor performance occurs in precision and F1 scores ranging from 0.095% to 49.496% in these methods except for SHDR with higher precision and F1 score of 100% and 73.811%, respectively.

The denoising results of four methods are presented in Fig. 8(b)–(f). All the methods cannot completely remove such clustered stripped noisy points, except our proposed method. MEOR-L can further identify noisy points located close to the ground point cloud represented as valid points.

D. Experiments in Dataset Scarborough

The dataset reported has no clustered noisy points are found in this dataset, but only scattered noisy points exist, in which they are located close to or far away from the ground valid point cloud.

The quantitative OR results are shown in Table V. The proportion of noisy points removed by MEOR-G, MOR, SOR, and, SFOR are similar, which are 35.863%, 41.746%, 60.531%, and 24.668%, respectively. SHDR yields the worst result, which cannot identify any noisy points. The subsequent MEOR-L further enhances the performance of MEOR-G by further removing 51.993% of noisy points. Our method, SHDR, and MOR preserve the integrity of valid point cloud, whereas SOR and SFOR remove 3.119% and 0.010% of valid points. All the methods achieve a high accuracy ranging from 96.880% to 99.999%. Our proposed MEOR and MOR, with a precision of 100%, far surpass SFOR, SOR, and SHDR with a precision of 8.621%, 0.077%, and 0.000%. MEOR comes with the best performance in recall of 87.856% and F1 score of 93.535%, whereas corresponding recall and F1 score of other methods range from 0.000% to 60.531% and 0.000% to 58.902%.

For intuitively observing and comparing the results among all the methods, the OR results of the Scarborough dataset are shown in Fig. 10. The best performance can be found in our proposed MEOR, which produces the desired result. Conversely, the worst performance is achieved by SHDR, where all of the scattered outliers cannot be detected. SFOR shows a slightly better performance than SHDR, but a number of noisy points, both far away and close to the ground, are not identified. Also, SOR and MOR are inefficient in removing the noisy points close to the ground point cloud.

V. DISCUSSION

A. Importance of Global–Local OR

The proposed MEOR includes two stages. The MEOR-G stage is able to look for a global elevation threshold to distinguish between valid points and noisy points that are found far from the ground point cloud. As the terrain of the dataset varies, the value of the elevation threshold calculated by MEOR-G is significantly higher than the valid points on the ground so as to preserve the integrity and details of the valid point cloud. As such, the valid points on the ground are less likely misidentified as noisy points that may be wrongly removed. It thus retains a number of scattered noisy points that may still exist close to the ground

TABLE VI
COMPUTATIONAL TIME OF OUR PROPOSED MEOR AND COMPARISON WITH SFOR, SOR, SHDR, AND MOR IN FOUR DATASETS

Methods	Aluto	Santorini	Vancouver	Scarborough
MEOR-G	5,435 s	1,552 s	409 s	525 s
MEOR-G + MEOR-L	47,552 s	10,440 s	2,748 s	3,663 s
SFOR [19]	20,128 s	3,924 s	1,344 s	2,289 s
SOR [20]	12,032 s	3,900 s	1,224 s	1,644 s
SHDR [21]	2,681,662 s	304,384 s	42,803 s	34,661 s
MOR [22]	22,960 s	4,332 s	1,518 s	2,391 s

point cloud. Accordingly, MEOR-L divides the denoising result of MEOR-G into small local parts to generate nearly flat terrain local point clouds. Within individual local point cloud, the MEOR-L stage estimates a local elevation threshold and further identifies noisy points based on the maximum entropy principle. The experiments show that the subsequent refinement achieved by the MEOR-L is necessary, leading to a notable improvement ranging from 0.024% to 51.993%, 0% to 0.002%, and 0.010% to 40.742% in recall, accuracy, and F1 score, respectively, compared to the results generated by MEOR-G.

B. Proposed MEOR Versus Four Classic or Methods

Our proposed MEOR method consistently outperforms four other classic OR methods in removing noisy points and preserving valid points according to the four experiments. The recall, precision, accuracy, and F1 score improvements of our method are found to be 47.839–99.815%, 82.757–94.591%, 0.012–4.539%, and 80.759–99.691%, respectively, compared to SFOR method. The improvements of our method in recall, precision, accuracy, and F1 score range from 8.744% to 98.589%, 59.666% to 99.923%, 3.119% to 8.454%, and 93.381% to 98.401%, respectively, compared to the results derived by SOR. Our method MEOR yields a slight improvement of accuracy ranging from 0.001% to 0.005% and a remarkable rise in recall and F1 score by 0.126–87.856% and 0.053–93.535%, which are calculated from Tables II to V. Compared to MOR, our proposed MEOR achieves an improved recall, precision, accuracy, and F1 score by 31.96–76.138%, 0–56.336%, 0.001–3.027%, and 34.633–61.474%, respectively. Our proposed MEOR significantly improves recall, precision, and F1 score and slightly improves accuracy as the other four classic methods still produce comparably high accuracy. SHDR can yield comparable results to our method since it considers detecting both clustered and scattered noisy points. Although SFOR and SOR are not comparable with our proposed MEOR, the MOR method still slightly performs better than SFOR and SOR. This can be explained by the similar mechanism between MOR and MEOR, in which both rely on the use of elevation to distinguish the noisy and valid points. The better performances achieved by MEOR and MOR than SOR and SFOR imply that point cloud elevation information-based denoising methods are much more effective than density-based denoising methods, especially for the datasets possessing clustered noisy points.

Table VI demonstrates the running time of our method and the other four methods on four datasets, respectively. The computational time of our proposed MEOR is approximately two to three times that of SFOR, SOR, and MOR methods. However, the time efficiency of our method is still much better than SHDR, which takes dozens of times more time than our method. Although the running time of our method is slightly longer than SFOR, SOR, and MOR, the point cloud denoising performance of our method is remarkably better than the other three methods. SHDR method generates a comparable denoising performance of our method; nevertheless, it is inefficient to remove noisy points and the running time grows exponentially with the increase of the total number of points. Overall, our method presents a better performance than the other four methods by taking both denoising results and time efficiency into account.

C. Comparison to Deep Learning-Based Method

To further validate the superiority of our proposed method, we exploit the classic Pointnet++ [23] as a benchmark deep learning-based method to compare with MEOR. In the experiment, we regard noisy points and valid points as two classes so that we can employ the classification function of Pointnet++ to identify noisy points. This experiment is conducted based on datasets Aluto, Santorini, Vancouver, and Scarborough, respectively, by using 10% of individual datasets as training data for Pointnet++ [23]. The denoising results shown in Table VII present that Pointnet++ [23] removes most part of the noisy points for dataset Aluto, Santorini, and Vancouver, whereas only a small number of noisy points can be removed for dataset Scarborough. The produced accuracy, F1 score, and computational time among four datasets range from 99.977% to 99.999%, 28.972% to 99.916%, and 41 322 to 288 635s. On the other hand, our proposed MEOR outperforms Pointnet++ [23] with an accuracy of 99.999% for all datasets, F1 score ranging from 93.535% to 99.99%, time cost ranging from 2748 to 47 552s (refer to Tables II to VI). Though the novel deep learning-based method is relatively effective in removing noisy points, it requires sufficient training data and consumes a large amount of computational time. Our proposed method demonstrates the advantages of generating more accurate denoising results and taking less computational time without the need of training data.

D. Ablation Study

Since the proposed MEOR-G and MEOR-L include a certain number of parameters, we intentionally conduct a series of ablation studies to examine the sensitivity of these parameters with respect to the denoising results and computational time.

Ablation study on parameter l in MEOR: In our proposed MEOR, we calculate the absolute elevation difference of point cloud and divide the elevation difference into l levels, the value of which is predefined manually, so that we can find elevation threshold based on maximum entropy to detect noisy points. Accordingly, we conduct an experiment based on part of the Aluto dataset to test the impact of l on the point cloud denoising performance. The value of l is set from 10 to 160 and increases sequentially by 10. We calculate the overall accuracy and computational time of the corresponding value of l. The result in Fig. 11 presents that the accuracy appears to be a sharp increase when the value of l equals 40 and slightly increases

TABLE VII
DENOISING PERFORMANCE OF POINTNET++ [23] IN ALUTO, SANTORINI, VANCOUVER, AND SCARBOROUGH DATASETS

Datasets	ONP	RNP	OVP	RVP	R (%)	P (%)	A (%)	F1 (%)	Time (s)
Aluto	4,921,256	3,654,415	118,849,388	0	74.258	100.00	98.977	85.228	288,635
Santorini	282,080	281,605	35,209,796	0	99.832	100.00	99.999	99.916	123,392
Vancouver	995	536	11,089,296	0	87.725	100.00	99.999	93.461	26,337
Scarborough	527	93	13,255,470	22	17.647	80.869	99.996	28.972	41,322

*ONP = Original noisy points, RNP = Removed noisy points, OVP = Original valid points, RVP = Removed valid points , R = Recall, P = Precision, A = Accuracy, and F1 = F1 score.

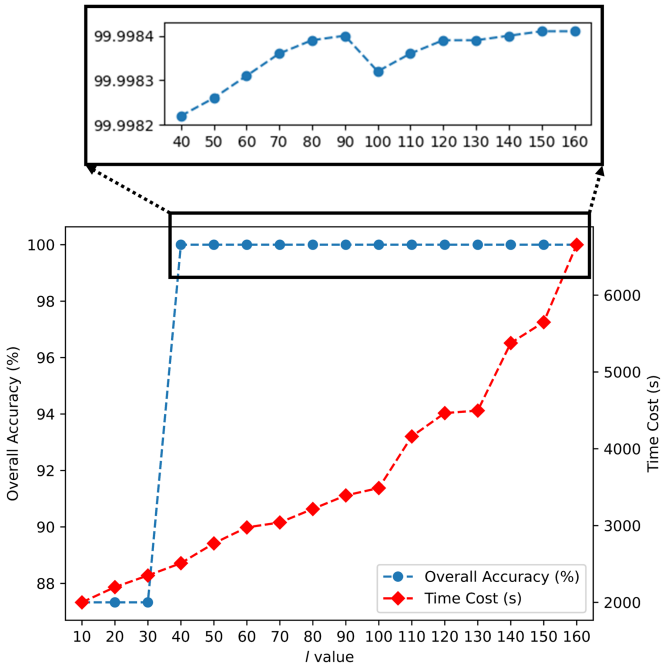


Fig. 11. Denoising performance of different l values.

with l from 40 to 100. Then the accuracy goes down when $l = 100$ and slightly goes up with l larger than 100. Whereas, the computational time increases gradually and appears to greatly rise when l is larger than 100. Therefore, we set the value of parameter l as 90, which can produce the most satisfied denoising accuracy and take relatively less time to implement the algorithm.

Ablation study on parameters η, θ_{th}, c_{th}, and c_{max} in MEOR-L: In our proposed MEOR-L, the values of four parameters, i.e., the number of searched neighbor points η, normal angle difference threshold θ_{th}, curvature threshold c_{th}, and maximum number of cluster c_{max}, are set to segment the point cloud dataset into clusters with minimum topographic variation. We conducted experiments based on the three datasets (dataset 1: 57 noisy points and 2 693 390 valid points; dataset 2: 45 noisy points and 2 258 472 valid points; dataset 3: 82 noisy points and 2 557 061 valid points) with great terrain relief to explore the impact of these four parameters on the point cloud denoising result. The experimental results based on three datasets generate similar conclusions and consistent trends about the parameters. Accordingly, we only present the results of dataset 1, which is shown in Fig. 2, to analyze the impact of the parameters on the clustering and denoising result.

Table VIII shows the denoising performance of our proposed MEOR with different values of parameters in the clustering as mentioned in Section II-E. The first row of parameters setting in the table presents the optimal values of four parameters, which generate the best performance of point cloud denoising. When $\eta = 5$, $\theta_{\text{th}} = 2.0$, $c_{\text{th}} = 0.00001$, and $c_{\text{max}} = 100$, the MEOR-L generates a total of 160 clusters, removes all of the 57 noisy points, and preserves all valid points, which realizes 100% in terms of recall, precision, accuracy, and F1 score, respectively. Subsequently, we test the impact of four parameters in sequence by changing the values of the concerned parameters and keeping the values of the other three parameters consistent with the parameters setting in the first row.

First, we focus on the impact of the first parameter η on denoising performance by setting the value of η as 10 and 15, respectively, and the other three parameters remain the same with the parameters setting in the first row. It is found that a larger number of searched neighbor points induces less number of clusters and causes worse point cloud denoising results with only 29 and 25 noisy points removed, respectively. In terms of parameter θ_{th}, we set it as 1.5, 2.5, and 3.0, respectively to analyze the sensitivity of normal angle difference θ_{th} on denoising performance. The result shows that the smaller value of θ_{th} produces more clusters and better denoising results. As for the effect of parameter c_{th}, we set the values of the curvature threshold as 0.001 and 0.1 to compare with the setting in the first row. The smaller curvature threshold generates more clusters, leading to a satisfied performance. Finally, we set the values of maximum number of clusters c_{max} as 500, 1500, and 2000, respectively. It is found that a larger value of maximum number of points to form a cluster usually produces fewer clusters and leads to better denoising performance. In summary, it is recommended to generate more clusters to achieve satisfied denoising results and a smaller value of these four parameters leads to produce more clusters. In the practical application of our proposed method, one can realize the requirement by tuning parameters depending on the terrain complexity of the given dataset. The far right column in Table VIII presents the time consumed to accomplish the clustering for MEOR-L with respect to the corresponding setting. It is recommended to set smaller values for the parameters to generate more clusters, especially for the point clouds collected on relief and complex terrain. However, more clusters thus require a longer computational time that reduces the efficiency. Therefore, one can make a tradeoff to determine the values of the parameters.

TABLE VIII
DENOISING PERFORMANCE OF OUR PROPOSED MEOR-L WITH DIFFERENT VALUES OF PARAMETERS IN THE CLUSTERING STEP

η	θ_{th}	c_{th}	c_{max}	NC	RNP	RVP	R (%)	P (%)	A (%)	F1 (%)	Time (s)
5	**2**	**0.00001**	**1,000**	**160**	**57**	**0**	**100.00**	**100.00**	**100.00**	**100.00**	466
10	2.0	0.00001	1,000	12	29	**0**	50.877	100.00	99.999	67.442	589
15	2.0	0.00001	1,000	7	25	**0**	43.860	100.00	99.999	60.976	506
5	1.5	0.00001	1,000	152	54	2	94.737	96.429	99.999	95.575	394
5	2.5	0.00001	1,000	146	50	6	87.719	89.286	99.999	88.500	571
5	3.0	0.00001	1,000	131	47	4	82.456	92.157	99.999	87.037	473
5	2.0	0.001	1,000	127	48	6	84.211	88.889	99.999	86.486	541
5	2.0	0.1	1,000	99	41	20	71.930	67.213	99.999	69.492	570
5	2.0	0.00001	500	166	55	4	96.491	93.220	99.999	94.828	404
5	2.0	0.00001	1,500	153	52	3	91.228	94.545	99.999	92.857	472
5	2.0	0.00001	2,000	126	46	6	80.702	88.462	99.999	84.404	447

*η = the number of searched neighbor points, θ_{th} = normal angle difference threshold, c_{th} = curvature threshold, c_{max} = maximum number of clusters, NC = Number of Clusters, RNP = Removed noisy points, RVP = Removed valid points, R = Recall, P = Precision, A = Accuracy, and F1 = F1 score.

VI. CONCLUSION

OR or denoising is a significant prerequisite for various applications of 3-D point cloud analysis. The complexity of noisy points dispersed throughout an airborne LiDAR point cloud presents significant challenges for effective identification and removal. Traditional classic noisy point removing methods lack the ability to simultaneously remove both clustered and scattered noisy points, especially the noisy points that are located close to the ground point cloud. This study proposed a global-to-local OR method based on the maximum entropy principle, named MEOR, to effectively discriminate clustered and scattered noisy points with valid point clouds. As the maximum entropy principle is proven that the largest entropy's summation of the probability distribution of valid and noisy points generates a maximum distinction between them. The maximum entropy principle can be exploited to estimate an elevation threshold for the given dataset to distinguish noisy points and valid points. The proposed method is built upon the two-stage strategy to calculate global and local elevation thresholds sequentially. The points with elevation differences smaller than the predicted elevation threshold are treated as valid points and preserved. Otherwise, the points are identified as noisy points and removed. The abundant experimental results sufficiently demonstrate our proposed MEOR consistently outperforms the existing classic SFOR, SOR, SHDR, and MOR methods in terms of removing the clustered and scattered noisy points and preserving the valid points with recall, precision, accuracy, and F1 score improvements by 0.126–99.815%, 0–100%, 0.001–8.454%, and 0.053–99.691%, respectively. Although our method MEOR produces a slight reduction of precision by 0.005–4.293% compared with SHDR in Aluto, Santorini, and Vancouver datasets, MEOR can realize time efficiency improvement by 89.432–98.227% compared with SHDR. Our proposed MEOR is proven to be more appropriate to efficiently denoise airborne LiDAR point clouds with multiple types of noisy points and preserve the integrity and details of ground and object point clouds while demanding a rather reasonable computational time.

ACKNOWLEDGMENT

The authors would like to thank Prof. Ayman Habib of Purdue University, Prof. Ahmed Shaker of Toronto Metropolitan University, McElhanney Consulting Services Ltd., and Teledyne Optech for providing the LiDAR dataset of Vancouver and Scarborough to support our experimental work, and the constructive comments provided by the reviewer and editor.

REFERENCES

[1] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, and M. Bennamoun, "Deep learning for 3D point clouds: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4338–4364, Dec. 2021.

[2] K. Mirzaei, M. Arashpour, E. Asadi, H. Masoumi, Y. Bai, and A. Behnood, "3D point cloud data processing with machine learning for construction and infrastructure applications: A comprehensive review," *Adv. Eng. Inform.*, vol. 51, 2022, Art. no. 101501.

[3] B. Fei et al., "Comprehensive review of deep learning-based 3D point cloud completion processing and analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 22862–22883, Dec. 2022.

[4] W. Liu, J. Sun, W. Li, T. Hu, and P. Wang, "Deep learning on point clouds and its application: A survey," *Sensors*, vol. 19, 19, 2019, Art. no. 4188.

[5] J. Zhang, X. Zhao, Z. Chen, and Z. Lu, "A review of deep learning-based semantic segmentation for point cloud," *IEEE Access*, vol. 7, pp. 179118–179133, 2019.

[6] S. A. Bello, S. Yu, C. Wang, J. M. Adam, and J. Li, "Deep learning on 3D point clouds," *Remote Sens.*, vol. 12, no. 11, 2020, Art. no. 1729.

[7] Y. Cui et al., "Deep learning for image and point cloud fusion in autonomous driving: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 722–739, Feb. 2022.

[8] A. F. Guarda, N. M. Rodrigues, and F. Pereira, "Point cloud coding: Adopting a deep learning-based approach," in *Proc. Picture Coding Symp.*, 2019, pp. 1–5.

[9] X. Li, C. Liu, Z. Wang, X. Xie, D. Li, and L. Xu, "Airborne LiDAR: State-of-the-art of system design, technology and application," *Meas. Sci. Technol.*, vol. 32, no. 3, 2020, Art. no. 032002.

[10] U. Okyay, J. Telling, C. L. Glennie, and W. E. Dietrich, "Airborne lidar change detection: An overview of earth sciences applications," *Earth- Sci. Rev.*, vol. 198, 2019, Art. no. 102929.

[11] L. Tian, Y. Qu, and J. Qi, "Estimation of forest LAI using discrete airborne LiDAR: A review," *Remote Sens.*, vol. 13, no. 12, 2021, Art. no. 2408.

[12] M. Michałowska and J. Rapiński, "A review of tree species classification based on airborne LiDAR data and applied classifiers," *Remote Sens.*, vol. 13, no. 3, 2021, Art. no. 353.

[13] W. Y. Yan, "Airborne LiDAR data artifacts: What we know thus far," *IEEE Geosci. Remote Sens. Mag.*, vol. 11, no. 3, pp. 21–45, Sep. 2023.

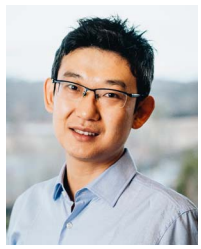
[14] Z. Hui, P. Cheng, Y. Guan, and Y. Nie, "Review on airborne LiDAR point cloud filtering," *Laser Optoelectron. Prog.*, vol. 55, no. 6, 2019, Art. no. 060001.
[15] C. Chen, J. Guo, H. Wu, Y. Li, and B. Shi, "Performance comparison of filtering algorithms for high-density airborne LiDAR point clouds over complex LandScapes," *Remote Sens.*, vol. 13, no. 14, 2021, Art. no. 2663.
[16] S. Jin, Y. Su, X. Zhao, T. Hu, and Q. Guo, "A point-based fully convolutional neural network for airborne LiDAR ground point filtering in forested environments," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3958–3974, 2020.
[17] M. J. Rakotosaona, V. La Barbera, P. Guerrero, N. J. Mitra, and M. Ovsjanikov, "PointCleanNet: Learning to denoise and remove outliers from dense point clouds," *Comput. Graph. Forum*, vol. 39, no. 1, pp. 185–203, 2020.
[18] Y. Regaya, F. Fadli, and A. Amira, "Point-denoise: Unsupervised outlier detection for 3D point clouds enhancement," *Multimedia Tools Appl.*, vol. 80, no. 18, pp. 28161–28177, 2021.
[19] M. Isenburg, "LAStools: Converting, filtering, viewing, processing, and compressing LiDAR data in LAS format," Accessed on: Apr. 23, 2023. [Online]. Available: https://rapidlasso.de/product-overview/
[20] A. C. Carrilho, M. Galo, and R. C. Santos, "Statistical outlier detection method for airborne LiDAR data," *Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 42, pp. 87–92, 2018.
[21] Y. Lin et al., "Noise point detection from airborne LiDAR point cloud based on spatial hierarchical directional relationship," *IEEE Access*, vol. 10, pp. 82076–82091, 2022.
[22] J. Zhang, J. Wang, and G. Liu, "Vertical structure classification of a forest sample plot based on point cloud data," *J. Indian Soc. Remote Sens.*, vol. 48, pp. 1215–1222, 2020.
[23] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," in *Proc. 31st Int. Conf. Adv. Neural Inf. Process. Syst.*, 2017, vol. 30, pp. 5105–5114.
[24] F. Pistilli, G. Fracastoro, D. Valsesia, and E. Magli, "Learning graph-convolutional representations for point cloud denoising," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 2, pp. 402–414, Feb. 2021.
[25] Z. Hui, P. Cheng, L. Wang, Y. Xia, H. Hu, and X. Li, "A novel denoising algorithm for airborne LiDAR point cloud based on empirical mode decomposition," *Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 42, pp. 1021–1025, 2019.
[26] Y. Duan, C. Yang, H. Chen, W. Yan, and H. Li, "Low-complexity point cloud denoising for LiDAR by PCA-based dimension reduction," *Opt. Commun.*, vol. 482, 2021, Art. no. 126567.
[27] S. Si et al., "Multiscale feature fusion for the multistage denoising of airborne single photon LiDAR," *Remote Sens.*, vol. 15, no. 1, 2023, Art. no. 269.
[28] A. Yang et al., "Filtering of airborne LiDAR bathymetry based on bidirectional cloth simulation," *ISPRS J. Photogrammetry Remote Sens.*, vol. 163, pp. 49–61, 2020.
[29] C. Chen, B. Chang, Y. Li, and B. Shi, "Filtering airborne LiDAR point clouds based on a scale-irrelevant and terrain-adaptive approach," *Measurement*, vol. 171, 2021, Art. no. 108756.
[30] S. Cai et al., "Filtering airborne LiDAR data through complementary cloth simulation and progressive TIN densification filters," *Remote Sens.*, vol. 11, no. 9, 2019, Art. no. 1037.
[31] A. Varlik and F. Uray, "Filtering airborne LiDAR data by using fully convolutional networks," *Surv. Rev.*, vol. 55, no. 388, pp. 21–31, 2023.
[32] E. T. Jaynes, "On the rationale of maximum-entropy methods," *Proc. IEEE*, vol. 70, no. 9, pp. 939–952, Sep. 1982.
[33] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
[34] S. Guiasu and A. Shenitzer, "The principle of maximum entropy," *Math. Intelligencer*, vol. 7, pp. 42–48, 1985.
[35] W. Hutchison, T. A. Mather, D. M. Pyle, J. Biggs, and G. Yirgu, "Structural controls on fluid pathways in an active rift system: A case study of the Aluto volcanic complex," *Geosphere*, vol. 11, no. 3, pp. 542–562, 2015.
[36] P. Nomikou et al., "The emergence and growth of a submarine volcano: The Kameni islands, Santorini (Greece)," *GeoResJ*, vol. 1, pp. 8–18, 2014.
[37] P. Rieger, "Range ambiguity resolution technique applying pulse-position modulation in time-of-flight scanning LiDAR applications," *Opt. Eng.*, vol. 53, no. 6, 2014, Art. no. 061614.

Ge Jiang (Student Member, IEEE) received the B.S. degree in automation engineering and the M.S. degree in control theory and control engineering from Hohai University, Nanjing, China, in 2016 and 2019, respectively. She received the Ph.D. degree in land surveying and geo-informatics from The Hong Kong Polytechnic University, Hong Kong, in 2024.

Her research interests include multispectral airborne LiDAR point cloud processing and land cover classification and segmentation.

Derek D. Lichti received the bachelor's degree in survey engineering from Ryerson University, Toronto, ON, Canada, in 1993, and the M.Sc. and Ph.D. degrees in geomatics engineering from the University of Calgary, Calgary, AB, Canada, in 1996 and 1999, respectively.

He is currently a Professor with the Department of Geomatics Engineering, University of Calgary, which he joined in 2008. He was the Head of the Department for the period 2013 to 2018. Following the completion of his Ph.D. degree in 1999, he worked as an Academic with Curtin University, Perth, Australia, until 2007. His research program focuses on precision 3-D reality capture from imaging sensors, principally terrestrial laser scanners, and digital cameras.



Tiangang Yin received the B.Sc. degree in physics and the M.Eng. degree in computer engineering from National University of Singapore, Singapore, in 2008 and 2011, respectively, and the Ph.D. degree in remote sensing and geoscience from the Centre d'Etudes Spatiales de la BIOsphère, Paul Sabatier University, Toulouse, France, in 2015.

He is currently an Assistant Professor with the Department of Land Surveying and Geo-Informatics, the Hong Kong Polytechnic University, Hong Kong. Prior to joining LSGI, he was a Research Scientist jointly at NASA Goddard Space Flight Center and the Earth System Science Interdisciplinary Center, University of Maryland. Since 2012, he has been leading research and development of several modules (LiDAR, photogrammetry, and urban atmosphere) for the discrete anisotropic radiative transfer model.

Wai Yeung Yan (Senior Member, IEEE) received the Ph.D. degree in civil engineering from Ryerson University, Toronto, ON, Canada, in 2012.

He was an Assistant Professor with the Department of Land Surveying and Geo-Informatics, the Hong Kong Polytechnic University and is now an Adjunct Professor with the Department of Civil Engineering, Ryerson University. His research interests include improving LiDAR data quality, point cloud processing, laser scanning, and remote sensing.