Research Article
Observation Selection, Total Variation, and L-Curve Methods for LiDAR Data Denoising

Jie Zhang,1 Pak Wai Chan,2 and Michael K. Ng1

1Department of Mathematics, The University of Hong Kong, Hong Kong, China
2Aviation Weather Services, Hong Kong Observatory, Hong Kong, China

Correspondence should be addressed to Michael K. Ng; mng@maths.hku.hk

Received 6 October 2021; Accepted 27 December 2021; Published 22 January 2022

Academic Editor: Hiroyuki Hashiguchi

Copyright © 2022 Jie Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this article, we propose a light detection and ranging (LiDAR) data denoising scheme for wind profile observation as a part of quality control procedure for wind velocity monitoring and windshear detection. The proposed denoising scheme consists of several components. (i) It selects LiDAR observations according to their SNR values so that serious noisy data can be removed. (ii) A polar-based total variation smoothing term is employed to regularize LiDAR observations. (iii) The regularization parameters are automatically determined to balance the data-fitting term and the total variation smoothing term. Numerical results for LiDAR data collected at the Hong Kong International Airport are reported to demonstrate that the denoising performance of the proposed method is better than that of the testing LiDAR data denoising schemes in the literature.

1. Introduction

Light detection and ranging (LiDAR) technique [1, 2] is a remote sensing tool that plays a significant role in environmental monitoring sciences. It is widely used in meteorological data observing. Generally, LiDAR emits a beam of light to the observation region. Some of this light would be backscattered towards the LiDAR receiver since it interacts with the medium or particles under observation [3]. The backscattered light captured by the LiDAR receiver is used to determine the characteristics of the observation area, e.g., the velocity of wind. Due to the impact of measurement environments and some other reasons, there would be some observation errors and very noisy observations in the observational LiDAR data as the range of observation increases [4]. It can have a serious effect in different LiDAR data applications such as windshear detection. Therefore, it is indispensable to develop an effective denoising method as a part of quality control for LiDAR observational data to improve the data quality and remove bad observations.

Several denoising methods have been developed to improve the quality of LiDAR data. There are mainly two different types of methods for LiDAR data denoising. One is for the time-varying but location-fixed LiDAR data. For example, a stationary wavelet domain spatial filtering-based denoising method was proposed by Yin et al. [5]. The method can effectively remove noise and detect the sudden change of LiDAR signal. Hassanpour [6, 7] proposed a singular value decomposition-based Savitzky-Golay approach for signal denoising. Similarly, Azadbakht et al. [8] employed the Savitzky-Golay method for full-waveform LiDAR data denoising. The second one is for the time-varying and distance-varying LiDAR data. For instance, Wu et al. [9] proposed a biorthogonal discrete wavelet transform (DWT) with a distance-dependent threshold algorithm to do the line-of-sight wind velocity denoising. In [10], Wu et al. studied the empirical mode decomposition (EMD) method [11] to analyze LiDAR data. Liu et al. applied the EMD method to a Doppler wind LiDAR acquisition system and got much better denoising results than the original method in [12]. In [13], Zhang et al. combined the EMD method with the Savitzky-Golay filtering algorithm, which can retain more features of LiDAR signal. Li et al. proposed a LiDAR denoising method based on ensemble empirical mode decomposition in [14]. This method can overcome the
2. Methods

In this section, the information about LiDAR data we investigate in this paper is given in Subsection 2.1. Next, we introduce the proposed model in Subsection 2.2. Also, the algorithm and parameter selection method are given in Subsections 2.3 and 2.4, respectively.

2.1. Data Sets. The Hong Kong International Airport (HKIA) is located at the place lying to the north of Lantau Island that is quite mountainous with heights ranging from 300 m to 900 m. Due to the complex terrain near the airport, it is necessary to collect LiDAR data of wind velocities and observe any windshear phenomena appearing over the flight paths of the airport. To provide timely windshear alerting, the Hong Kong Observatory devised a Doppler LiDAR system (see [20, 21] for more details). Due to the highly cluttered environment around HKIA such as vehicles, derricks, barges, windmills, and cable cars, there are lots of noise and bad observations in the observational LiDAR data. For example, we show in Figure 1(a) the LiDAR radial velocity data of conical scan, where the radius and the polar angle of the scan refer to the slant range and the azimuth angle of LiDAR beam, respectively. In Figure 1(b), we show the signal-to-noise ratio (SNR) of the measured wind velocities corresponding to Figure 1(a). It is obvious that there are noise and some outliers whose SNR values are from $-10$ to $-60$ in the observational data. Therefore, it is significant to develop an efficient data denoising method to remove bad observations and improve the quality of observational LiDAR data.

The set of data used in this study was collected at HKIA from 1 March to 31 March 2015, including the 147 windshear cases that reported in the pilot report and several nonwindshear cases. Each scan of windshear cases took $y_i \in \mathbb{R}^n$ about 25 seconds (see one example in Figure 1).

2.2. The Proposed Model. Let be the LiDAR data observed at azimuth angle $\theta_j$. For simplicity, we assume that there are $m$ azimuth angles ($\theta_1 < \theta_2 < \cdots < \theta_m$) to be recorded in between 0° and 359°, and also there are $n$ range values to be recorded, i.e.,

$$y_i = [y_{i,1}, y_{i,2}, \ldots, y_{i,n}], \quad 1 \leq i \leq m. \quad (1)$$

Note that the locations of such range values are not necessary to be uniform, and the distance between the LiDAR centre and the observed value $y_{i,j}$ is equal to $r_j$. We are interested to compute the denoising LiDAR data as follows:

$$x_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,n}], \quad 1 \leq i \leq m, \quad (2)$$

according to the given SNR values:

$$s_i = [s_{i,1}, s_{i,2}, \ldots, s_{i,n}], \quad 1 \leq i \leq m, \quad (3)$$

from the observed LiDAR data. When $y_{i,j}$ is a missing LiDAR observation, the corresponding $s_{i,j}$ can be set to be $-\infty$. In total, there are $mn$ observations and SNR values and $mn$ unknowns covered in the conical scan in a two-dimensional plane. In the proposed minimization model, there are three components in the objective function.
The first term is the data-fitting term between the observed LiDAR data $y_i$ and the denoised data $x_i$ ($1 \leq i \leq m$). In order to determine whether $y_{i,j}$ at the azimuth angle $\theta_i$ and range value $r_j$ to be used in the model, we incorporate a nonnegative weight $w_{i,j}$ for $y_{i,j}$. When the value of $w_{i,j}$ is equal to zero, the LiDAR observation $y_{i,j}$ is not used. The resulting data-fitting term is given by

$$\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j}^2 (y_{i,j} - x_{i,j})^2.$$  \hspace{1cm} (4)

It is clear that when the LiDAR data observation is removed ($w_{i,j} = 0$), the corresponding data-fitting term does not appear.

In the model, we would like to smooth LiDAR observations over all possible range and azimuth values together. In image processing, total variation regularization techniques based on rectangular pixel-based form were shown to be a very useful denoising method (see, for instance, [23–25]). Here, the polar-based total variation regularization term is proposed and employed as follows:

$$TV(x_{i,j}) = \frac{|x_{i,j} - x_{i-1,j}|}{d(x_{i,j}, x_{i,j-1})} + \frac{|x_{i,j} - x_{i-1,j}|}{d(x_{i,j}, x_{i-1,j})},$$  \hspace{1cm} (5)

where $d(x_{i,j}, x_{i,j-1})$ refers to the distance between the two observed range values at the same azimuth angle, i.e.,

$$d(x_{i,j}, x_{i,j-1}) = r_j - r_{j-1},$$  \hspace{1cm} (6)

and $d(x_{i,j}, x_{i,j-1})$ refers to the distance between the same observed range value at the two adjacent azimuth angles $\theta_i$ and $\theta_{i-1}$, i.e.,

$$d(x_{i,j}, x_{i-1,j}) = 2r_j \sin \left( \frac{\theta_i - \theta_{i-1}}{2} \right).$$  \hspace{1cm} (7)

In (7), the first and the second terms are the one-dimensional total variation regularization. However, we apply the one-dimensional total variation regularization to the polar-based LiDAR data observations by using their actual distances in the formula. The combined total variation regularization term is given as follows:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j}^2 TV(x_{i,j}).$$  \hspace{1cm} (8)

We see from (7) and (8) that the denoised values $x_{i,j}$ are coupled into the regularization formula, and therefore the denoising procedure is adapted to the whole polar-based LiDAR data observations instead of observations in range values only such as the methods we introduce in Section 1.

In the model, we should include SNR values $s_{i,j}$ to determine whether the LiDAR observation $y_{i,j}$ is used in the denoising procedure. Here, we propose the following term in the model:

$$\lambda_1 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(-s_{i,j})w_{i,j}^2$$

$$\lambda_2 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(s_{i,j})(w_{i,j} - 1)^2,$$  \hspace{1cm} (9)

where $\lambda_1$ and $\lambda_2$ are the positive numbers to control the balance between the two terms in (9). We note in (9) that when $s_{i,j}$ is too small (a negative number), the LiDAR observation is very noisy, and therefore $\exp(-s_{i,j})$ is large and $\exp(s_{i,j})$ is small. The optimization process drives $w_{i,j}$ to be zero. Similarly, when $s_{i,j}$ is not small (the LiDAR observation is not noisy), $\exp(-s_{i,j})$ is small and $\exp(s_{i,j})$ is large and therefore the optimization process drives $w_{i,j}$ to be one. Finally, we also incorporate nonnegativity constraint on $w_{i,j}$. In the optimization model:
\[ w_{i,j} \geq 0, \quad 1 \leq i \leq m, \ 1 \leq j \leq n. \]  

(10)

For the ease of presentation, we denote \( x = \{x_{i,j}\} \) and \( w = \{w_{i,j}\} \). According to (4), (6), (9), and (10), the combined optimization model is given as follows:

\[
(x, w) = \arg \min_{(x, w)} \left\{ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} (y_{i,j} - x_{i,j})^2 + \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} TV(x_{i,j}) + \chi(w) 
\right. 
\]
\[
+ \lambda_1 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(-s_{i,j}) w_{i,j}^2 + \lambda_2 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(s_{i,j})(w_{i,j} - 1)^2 \right\},
\]

(11)

where \( \alpha \) is a positive number to balance the contribution of the total variation regularization term and the other terms and \( \chi_W(w) \) is the characteristic function of the non-negativity constraint set \( W = \{w: w_{i,j} \geq 0, \ 1 \leq i \leq m, \ 1 \leq j \leq n\} \)

\[
\chi_W(w) = \begin{cases} 
0, & w \in W, \\
+\infty, & \text{otherwise}.
\end{cases}
\]

(12)

2.3. The Algorithm. In this subsection, we will introduce the algorithm for the proposed model. Since both \( d(x_{i,j}, x_{i-1,j}) \) and \( d(x_{i,j}, x_{i-1,j}) \) are fixed in LiDAR observations, for simplicity, we use \( \tau_{i,j} \) and \( \mu_{i,j} \) to represent them, respectively, in the following discussion. Let \( p_{i,j} = x_{i,j} - x_{i-1,j}, q_{i,j} = x_{i,j} - x_{i-1,j} \). We can rewrite (11) as follows:

\[
(x, w, p, q) = \arg \min_{(x, w, p, q)} \left\{ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} (y_{i,j} - x_{i,j})^2 + \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \left| p_{i,j} \right| \frac{\left| p_{i,j} \right|}{\tau_{i,j}} + \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \frac{\left| q_{i,j} \right|}{\mu_{i,j}} 
\right. 
\]
\[
+ \chi_W(w) + \lambda_1 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(-s_{i,j}) w_{i,j}^2 + \lambda_2 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(s_{i,j})(w_{i,j} - 1)^2 \right\},
\]

\[
\text{s.t. } p_{i,j} = x_{i,j} - x_{i-1,j}, \ q_{i,j} = x_{i,j} - x_{i-1,j}.
\]

The augmented Lagrangian of the above-constrained optimization problem is given as follows:

\[
L_{\gamma_1, \gamma_2}(x, w, p, q; \Lambda, \Gamma) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} (y_{i,j} - x_{i,j})^2 + \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \left| p_{i,j} \right| \frac{\left| p_{i,j} \right|}{\tau_{i,j}} 
\]
\[
+ \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \frac{\left| q_{i,j} \right|}{\mu_{i,j}} + \chi_W(w) + \lambda_1 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(-s_{i,j}) w_{i,j}^2 
\]
\[
+ \lambda_2 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(s_{i,j})(w_{i,j} - 1)^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} \Lambda_{i,j} \left| p_{i,j} \right| - \left| (x_{i,j} - x_{i-1,j}) \right| 
\]
\[
+ \sum_{i=1}^{m} \sum_{j=1}^{n} \Gamma_{i,j} \left| q_{i,j} \right| - \left| (x_{i,j} - x_{i-1,j}) \right| + \frac{\gamma_1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \left| p_{i,j} \right| - \left| (x_{i,j} - x_{i-1,j}) \right|^2 
\]
\[
+ \frac{\gamma_2}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \left| q_{i,j} \right| - \left| (x_{i,j} - x_{i-1,j}) \right|^2,
\]

(14)
where \( \gamma_1, \gamma_2 > 0 \) are penalization parameters and \( \Lambda \) and \( \Gamma \) are Lagrange multipliers.

With an initial guess of \( (x^0, w^0, p^0, q^0, \Lambda^0, \Gamma^0) \), the iterations of the alternating direction method of multipliers [26] are given as follows:

\[
\begin{align*}
x^{k+1} & = \arg \min_x L_{y_i}(x, w^k, p^k, q^k; x^k, i^k), \\
w^{k+1} & = \arg \min_w L_{y_i}(x^{k+1}, w, p^k, q^k; x^k, i^k), \\
p^{k+1} & = \arg \min_p L_{y_i}(x^{k+1}, w^{k+1}, p, q^k; x^k, i^k), \\
q^{k+1} & = \arg \min_q L_{y_i}(x^{k+1}, w^{k+1}, p^{k+1}, q; x^k, i^k), \\
\Lambda_{i_{k+1}} & = \Lambda_{i_k} + \gamma_1 \left[ \frac{p^{k+1}}{\gamma_1} - (x^{k+1}_{i_{k+1}} - x^{k+1}_{i_{k+1}-1}) \right], \\
\Gamma_{i_{k+1}} & = \gamma_2 \left[ \frac{q^{k+1}}{\gamma_2} - (x^{k+1}_{i_{k+1}} - x^{k+1}_{i_{k+1}-1}) \right].
\end{align*}
\]

Next, we demonstrate how to solve the abovementioned subproblems (15) with respect to \((x, w, p, q)\).

Let us consider the \( x \)-subproblem in (15):

\[
x^{k+1} = \arg \min_x \left\{ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j}(y_{i,j} - x_{i,j})^2 + \frac{\gamma_1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[ \frac{p_{i,j}^k + \Lambda_{i,j}^k}{\gamma_1} - (x_{i,j} - x_{i,j-1}) \right]^2 \\
+ \frac{\gamma_2}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[ \frac{q_{i,j}^k}{\gamma_2} - (x_{i,j} - x_{i-1,j}) \right]^2 \right\}.
\]

Noting that the data-fitting term will vanish if there is no observation data \( y_{i,j} \). The optimality condition of \( x \)-subproblem is given by

\[
\begin{align*}
\left( u_{i,j}^k + 2\gamma_1 + 2\gamma_2 \right)x_{i,j} - y_1 \left( x_{i,j+1} + x_{i,j-1} \right) - y_2 \left( x_{i-1,j} + x_{i+1,j} \right) \\
= u_{i,j}^k x_{i,j} + y_1 \left( p_{i,j}^k - p_{i-1,j}^k \right) + y_2 \left( q_{i,j}^k - q_{i+1,j}^k \right) + \Lambda_{i,j}^k - \Lambda_{i,j+1}^k + \Gamma_{i,j}^k - \Gamma_{i+1,j}^k, \\
2 \left( y_1 + y_2 \right)x_{i,j} - y_1 \left( x_{i,j+1} + x_{i,j-1} \right) - y_2 \left( x_{i-1,j} + x_{i+1,j} \right) \\
= y_1 \left( p_{i,j}^k - p_{i+1,j}^k \right) + y_2 \left( q_{i,j}^k - q_{i+1,j}^k \right) + \Lambda_{i,j}^k - \Lambda_{i,j+1}^k + \Gamma_{i,j}^k - \Gamma_{i+1,j}^k, \\
\end{align*}
\]

We can solve the above set of equations by using the conjugate gradient (CG) method [27, 28] for \( w \)-subproblem, we have

\[
\begin{align*}
\arg \min_{w \geq 0} \left\{ \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j}(y_{i,j} - x_{i,j}^k)^2 + \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left| p_{i,j}^k \right|}{\mu_{i,j}} \\
+ \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \left[ q_{i,j}^k \right] \mu_{i,j} + \lambda_1 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(-s_{i,j})w^2_{i,j} \\
+ \lambda_2 \sum_{i=1}^{m} \sum_{j=1}^{n} \exp(s_{i,j})(w_{i,j} - 1)^2 \right\}.
\end{align*}
\]

We note that \( w_{i,j} \) are decoupled, and we only need to deal with scalar optimization problems. More precisely, if \( y_{i,j} \) exists, the scalar optimization problem is given by

\[
\begin{align*}
\arg \min_{w_{i,j} \geq 0} \left\{ \frac{1}{2} w_{i,j}(y_{i,j} - x_{i,j}^k)^2 + \alpha w_{i,j} \left| \frac{p_{i,j}^k}{\mu_{i,j}} \right| \\
+ \alpha w_{i,j} \left[ q_{i,j}^k \right] \mu_{i,j} + \lambda_1 \exp(-s_{i,j})w^2_{i,j} \\
+ \lambda_2 \exp(s_{i,j})(w_{i,j} - 1)^2 \right\}.
\end{align*}
\]

If \( y_{i,j} \) does not exist, it is equal to
\[ w_{i,j}^{k+1} = \arg\min_{w \in \mathbb{R}} \left\{ \alpha w_{i,j} \frac{|p_{i,j}^k|}{\tau_{i,j}} + \alpha w_{i,j} \frac{|q_{i,j}^k|}{\mu_{i,j}} + \lambda_1 \exp(-s_{i,j}) w_{i,j}^2 + \lambda_2 \exp(s_{i,j}) (w_{i,j} - 1)^2 \right\}. \]  

(20)

Both the above two scalar optimization problems are convex. And one can readily get that

\[ w^{k+1} = \max\{w^{k+1}, 0\}, \]

where

\[ w_{i,j}^{k+1} = \begin{cases} \frac{2\lambda_2 \exp(s_{i,j}) - 1/2(y_{i,j} - x_{i,j}^{k+1})^2 - \alpha |p_{i,j}^k|/\tau_{i,j} - \alpha |q_{i,j}^k|/\mu_{i,j}}{2\lambda_1 \exp(-s_{i,j}) + 2\lambda_2 \exp(s_{i,j})}, & \text{if } y_{i,j} \text{ exists}, \\ \frac{2\lambda_2 \exp(s_{i,j}) - \alpha |p_{i,j}^k|/\tau_{i,j} - \alpha |q_{i,j}^k|/\mu_{i,j}}{2\lambda_1 \exp(-s_{i,j}) + 2\lambda_2 \exp(s_{i,j})}, & \text{otherwise}. \end{cases} \]  

(22)

For \( p \)-subproblem and \( q \)-subproblem, they are given by

\[ p^{k+1} = \arg\min_p \left\{ \alpha m \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \frac{|p_{i,j}|}{\tau_{i,j}} + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[ p_{i,j} + \frac{\Lambda_{i,j}^k}{\gamma_1} - (x_{i,j}^{k+1} - x_{i,j}^{k+1-1}) \right]^2 \right\}, \]  

(23)

and

\[ q^{k+1} = \arg\min_q \left\{ \alpha m \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \frac{|q_{i,j}|}{\mu_{i,j}} + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[ q_{i,j} + \frac{1_{i,j}^k}{\gamma_2} - (x_{i,j}^{k+1} - x_{i,j}^{k+1-1}) \right]^2 \right\}, \]  

(24)

respectively. These two subproblems can be solved by using the soft-thresholding technique [29], and their solutions are given by

\[ p_{i,j}^{k+1} = \max\left( \left(x_{i,j}^{k+1} - x_{i,j}^{k+1-1}\right) - \frac{\alpha w_{i,j}^{k+1}}{\gamma_1} \frac{\Lambda_{i,j}^k}{\gamma_1}, 0 \right), \]

\[ q_{i,j}^{k+1} = \max\left( \left(x_{i,j}^{k+1} - x_{i,j}^{k+1-1}\right) - \frac{\alpha w_{i,j}^{k+1}}{\gamma_2} \frac{1_{i,j}^k}{\gamma_2}, 0 \right), \]  

where \( \text{sign}(p) = \alpha |p| \).

Finally, the overall algorithm is summarized in Algorithm 1.

2.4. The Calculation of Parameters. In the proposed denoising scheme, there are several parameters to be determined. In this subsection, we incorporate the L-curve method to determine the values of parameters. The L-curve is a tradeoff curve between two quantities that both need to be controlled and balanced [30, 31]. It is widely used in the engineering and applied mathematics field.

First, the regularization parameter \( \alpha \) in (14) is crucial in our proposed method since it controls the ratio between the data-fitting term and the total variation regularization term. When the value of \( \alpha \) is large (small), the importance of the
In order to compare the degrees of data fitting and smoothness with respect to different numbers of denoising results are more (less) regularized. Moreover, there are two parameters $\lambda_1$ and $\lambda_2$ in (14), and they can affect the values of $w_{i,j}$ according to the SNR values of LiDAR observations. When the ratio $\lambda_1/\lambda_2$ is large (i.e., $\lambda_1$ is larger than $\lambda_2$), the term $\lambda_1 \sum_{i=1}^m \sum_{j=1}^n \exp(-s_{i,j})w_{i,j}^2$ is more important than the term $\lambda_2 \sum_{i=1}^m \sum_{j=1}^n \exp(-s_{i,j}) (w_{i,j} - 1)^2$. It follows that when LiDAR observations have small SNR values, there would be more zero values of $w_{i,j}$ in the optimization procedure. In contrast, when the ratio $\lambda_1/\lambda_2$ is small and LiDAR observations have large SNR values, there would be more values of $w_{i,j}$ to be one in the optimization procedure.

In the L-curve method, we consider the following two quantities:

$$D(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2)) = - \log_{10} \left( \frac{\sum_{i=1}^m \sum_{j=1}^n w_{i,j}(\mathbf{a},\lambda_1,\lambda_2)(y_{i,j} - \mathbf{x}_{i,j}(\mathbf{a},\lambda_1,\lambda_2))^2}{\sum_{i=1}^m \sum_{j=1}^n w_{i,j}(\mathbf{a},\lambda_1,\lambda_2)} \right).$$ (26)

and

$$S(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2)) = - \log_{10} \left( \frac{\sum_{i=1}^m \sum_{j=1}^n w_{i,j}(\mathbf{a},\lambda_1,\lambda_2)TV(x_{i,j}(\mathbf{a},\lambda_1,\lambda_2))}{\sum_{i=1}^m \sum_{j=1}^n w_{i,j}(\mathbf{a},\lambda_1,\lambda_2)} \right).$$ (27)

The first quantity $D(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2))$ refers to the degree of data fitting when $\mathbf{x}$ and $\mathbf{w}$ are computed by Algorithm 1 for fixed $a$, $\lambda_1$, and $\lambda_2$. Similarly, the second quantity $S(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2))$ refers to the degree of smoothness in the regularization for the computed $\mathbf{x}$ and $\mathbf{w}$. In order to compare the degrees of data fitting and smoothness with respect to different numbers of denoising values $w_{i,j}$ across different values of $\lambda_1$ and $\lambda_2$, the weighted average of data fitting and smoothness are accounted in the two quantities. The main idea of the L-curve method is to select the suitable values of $a$, $\lambda_1$, and $\lambda_2$ such that both $D(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2))$ and $S(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2))$ are balanced.

As an illustration, we apply Algorithm 1 to the LiDAR observations in Figure 1 for different values of $a$, $\lambda_1$, and $\lambda_2$. More precisely, in Figure 2(a), we fix $\lambda_2$ (100) and $\lambda_1$ (0.5, 5, 50, 500) and compute $\mathbf{x}$ and $\mathbf{w}$ for different values of $a$ (0.01, 0.05, 0.1, 0.3, 0.5, 0.8, 1, 5). In Figure 2(a), we plot $D(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2))$ and $S(\mathbf{x}(a,\lambda_1,\lambda_2), \mathbf{w}(a,\lambda_1,\lambda_2))$. The lines of points of $D$ and $S$ with respect to $a$ are generated for each fixed $\lambda_1$ and $\lambda_2$. In total, there are four lines of points in Figure 2(a). We observe in each line that when $a$ is small, $D$ is large and $S$ is small, and thus the points appear in the left hand side of the line. When $a$ is large, $D$ is small and $S$ is large, and thus the points appear in the right hand side of the line. We find that there are several corner points in the line, and they refer to several large rates of change of data fitting with respect to the change of smoothness. Here, we can pick up the corner point with the largest rate of change and employ the corresponding value of $a$ for regularization. Next, we select the values of $\lambda_1$ and $\lambda_2$ by comparing the selected corner points of different lines. Here, we can pick up the selected corner point with the largest $D$ (the data fitting is good) and the largest $S$ (the smoothness is large). According to the plot in Figure 2(a), we select $a^* = 0.5$, $\lambda_1^* = 100$, and $\lambda_2^* = 500$.

Based on the above idea, we summarize our L-curve method with automatic selection of $a$, $\lambda_1$, and $\lambda_2$ in Algorithm 2.

From Figure 2, it is obvious that the model is very sensitive for different values of the regularization parameter $a$ due to its balance effect for data fitting and smoothing. However, from Figure 2(a), for fixed values of parameters $a$
3. Results and Discussion

In this section, to illustrate the effectiveness of the proposed scheme (denoted as the “TV” method), we will show some denoising results by this scheme and compare them with the results generated by other denoising methods for the LiDAR data collected at HKIA. Clearly, the LiDAR data we study in this paper are time-varying and distance-varying, so that we compare the results by the DWT method [9] and the EMD-CIIT method [16]. They are denoted as “DWT” and “EMD-CIIT.” Moreover, we consider the method used by Baranov et al. [18] (denoted as “Baranov” in the rest of this article) and Newsom et al. [19] (denoted as “Newsom” in the rest of this article) as well as the method used by Hong Kong Observatory (denoted as the HKO method in the rest of this article) as well as the method used by Hong Kong Observatory for the LiDAR data observed at HKIA. The effects of averaging smoothing algorithm and median smoothing algorithm are shown to be about the same in [18], so that we apply a five-point-based averaging smoothing algorithm to the data selected by the SNR threshold level. Refer to Algorithm 4 for more detailed description.

3.1. Comparison Results with the “DWT” and “EMD-CIIT” Methods. Figures 3 and 4 show the results of the “DWT” method, the “EMD-CIIT” method, and the proposed TV method for LiDAR data collected on 5 March 2015 from 04:16:02 to 04:16:25 with a fixed azimuth angle 224° and 4 March 2015 from 21:26:16 to 21:46:39 with a fixed azimuth angle 314°, respectively. According to the record from Hong Kong International Airport, these two wind profiles from LiDAR data correspond to two windshear cases. According to Figures 3 and 4, one can readily find that

The “DWT” method and the “EMD-CIIT” method cannot remove very noisy observations, but our proposed scheme can remove noisy observations with very low SNR values effectively.

For example, there is a high peak appearing at the range of around 8200 m in the LiDAR data observation in Figure 3. The DWT denoising method ($x_{DWT}$) and the EMD-CIIT denoising method...
**Input:** LiDAR observations $y_{i,j}$ and their SNR values $s_{i,j}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$

1. $i = 1$ (for each azimuth angle)
2. for $j = 1$ to $n$ (along each range) do
3. If $s_{i,j} < -5$, then $y_{i,j} = NA$
4. end for
5. $x_{1} = y_{1,1}$
6. for $j = 2$ to $n$ do
7. If $|s_{i,j}^{(HKO)} - y_{i,j}| > 16$, then $x_{i,j}^{(HKO)} = s_{i,j}^{(HKO)} \times 0.95 + y_{i,j} \times 0.05$
8. If $x_{i,j-1} = NA$ and $y_{i,j} = NA$, then $x_{i,j} = NA$
9. end for
10. Set $i = i + 1$ and Goto Step 2

**Algorithm 3:** The Denoising Method used by Hong Kong Observatory.

**Input:** LiDAR observations $y_{i,j}$ and their SNR values $s_{i,j}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$

1. Set $i = 1$ (for each azimuth angle)
2. for $j = 1$ to $n$ (along each range) do
3. If $s_{i,j} < -5$, then $y_{i,j} = NA$
4. end for
5. $x_{i,1}^{Baranov} = y_{i,1}$
6. $x_{i,2}^{Baranov} = y_{i,1} + y_{i,2}/2$
7. for $j = 3$ to $n - 2$ do
8. $x_{i,j}^{Baranov} = y_{i,j-2} + y_{i,j-1} + y_{i,j} + y_{i,j+1} + y_{i,j+2}/5$
9. end for
10. $x_{i,n-1}^{Baranov} = y_{i,n-3} + y_{i,n-2} + y_{i,n-1} + y_{i,n}/4$
11. $x_{i,n}^{Baranov} = y_{i,n-2} + y_{i,n-1} + y_{i,n}/3$
12. Set $i = i + 1$ and Goto Step 2

**Algorithm 4:** The Denoising Method used by Baranov et al.

---

Figure 3: The denoising results by the "DWT" (a), "EMD-CIIT" (b), and our proposed scheme (c) for the LiDAR data collected on 5 March 2015 from 04:16:02 to 04:16:25 with a fixed azimuth angle (224.0°). The SNR values of the corresponding observed LiDAR data are shown in (d).
EMD−CIIT) cannot remove this very noisy observation. We see from Figure 3(d) that the corresponding SNR value is about $-8$ at the range of around 8200 m. Indeed, it is quite acceptable to remove such noisy observation compared with neighborhood observations. Similarly, in Figure 4, there is a rock bottom appearing at the range of around 5000 m in the LiDAR data observation. From Figure 4(d), one can readily get that the corresponding SNR value is about $-7$ at the range of around 5000 m, which indicates the high noise intensity at this slant range. We see from Figures 4(a) and 4(b) that the “DWT” method and the “EMD-CIIT” method cannot remove such noisy observations. However, the proposed scheme (Figures 3 and 4(c)), according to the corresponding SNR values, filters both the high peak and rock bottom as well as the noisy data around them out.

(1) The Baranov method might restore few missing points as well as some data points that are removed by the SNR threshold level by the averaging smooth algorithm. However, the values of them seem to be quite the same as the previous observed one, which play a tiny role in the following data applications such like windshear detection.

For example, in Figure 5(a), there is a rock bottom appearing at the range of around 5000 m in the LiDAR data observation. From Figure 5(b), the corresponding SNR value is about $-7$, which indicates the data point is removed by the SNR threshold level. However, this data point is restored by the averaging smoothing algorithm in the results by the “Baranov” method ($x_{Baranov}$). It is obvious that the velocity values of data points between the range of 4800 m and 5800 m are quite the same in $x_{Baranov}$. Similarly, there are some missing values at the range of around 9000 m and 10800 m. Some missing data points are restored with almost same values in $x_{Baranov}$.

(2) Due to the effect of the averaging smoothing algorithm, the “Baranov” method might mis-shift the velocities of some data points.

For example, there are some mis-shifted data points in the result by the “Baranov” method at the range of around 4500 m in Figure 5(a). From 5(b), the SNR values around 4500 m are around 0, which are still acceptable. Similarly, the SNR values of data points at

---

3.2. Comparison Results with the “Baranov” Method.

Figures 5(a) and 5(c) show the results by the “Baranov” method and the proposed scheme for LiDAR data collected on 4 March 2015 from 21:26:16 to 21:46:39 with a fixed azimuth angle (314.0°). The SNR values of the corresponding observed LiDAR data are shown in the (d).
3.3. Comparison Results with HKO Method. In Figures 6(a) and 6(c), we consider LiDAR data collected on 4 March 2015 from 21:26:16 to 21:46:39 with a fixed azimuth angle 314° and the denoising results. (b) The SNR values of the LiDAR data in the 1st row. (c) The LiDAR data were on 5 March 2015 from 04:16:02 to 04:16:25 with a fixed azimuth angle 277° and the denoising results. (d) The SNR values of the LiDAR data in the 3rd row.

(1) There are some mis-shifted results at the range values around 4000 m–4200 m in Figure 6(a) and around 4700 m–5000 m in Figure 6(c). For reference, the SNR values corresponding to LiDAR data in Figures 6(a) and 6(c) are given in Figures 6(b) and 6(d), respectively. We note in Figure 6(b) that the SNR values at 4000 m–4200 m are 2–10. Similarly, the SNR values at 4700 m–5000 m are 13–18 in Figure 6(d). In these two cases, SNR values can still be in an acceptable range. However, the method by Hong Kong observatory cannot provide a reasonable denoising results in these two range values, and the detected wind profiles may be affected.

(2) In comparison, the results by the proposed method $x_{TV}$ of LiDAR data by the proposed method in Figures 6(a) and 6(c). We find that the proposed method fits the LiDAR data quite well with as much smoothness as possible based on the SNR values of the given observations.

3.4. Comparison Results with the “Newsom” Method. Note that the “Newsom” method just removes the bad observations based on the SNR values (the SNR threshold level is set to be 0.02 in [19]) without any other operations for LiDAR data, so that there are no value differences between the data after processing and the observed data. For LiDAR data in Figure 3, all the observed data after the range 4400 m are removed because their SNR values are less than 0.02. It seems that it is not reasonable to remove all these observed data points. In Figure 4, the observed data points near the rock bottom are removed. In Figures 5 and 6, the processed data by the “Newsom” method are about the same as the original observed data.

3.5. Comparison Results in Conical Scans. In Figures 7 and 8, we plot the conical scans of LiDAR data collected on 4 March 2015 from 21:26:16 to 21:46:39 and 5 March 2015 from 04:16:02 to 04:16:25 and the results by the “DWT” method, the “EMD-CIIT” method, the “Newsom” method, the “Baranov” method, the HKO method, and the proposed
method. The computational time for these methods is also given. From these figures, we can readily find that

1. There are some spikes in the denoising results of the “DWT” method and the “EMD-CIIT” method. For example, in Figures 7(e) and 7(f), there are some spikes in the range of azimuth angles from 206° to 250°.

2. The over-smoothness of the “EMD-CIIT” method is also overcome by the proposed scheme.

3. There are some outliers that provide little information of the wind profile in the denoising results generated by the “Baranov” method. In comparison, the denoising results of our proposed method are much clear.

We note that Figures 7 and 8 correspond to two windshear cases in March 2015. Although the existence of outliers does not have an effect on visual judgement for windshear analysis, it can have a strong effect on machine learning methods in windshear detection. For instance, the wind profile features suggested by the machine learning method in [32] are calculated based on the maximal difference in wind velocities within a range of azimuth angles, and it is clear that the outliers can change the resulting wind profile features.

4. Since the SNR threshold level (SNR > 0.02) used by “Newsom” method is high, there are only few data points retained in the results. And the proposed method, without any fixed SNR threshold level, retained more data points which contain useful information of wind profiles.

5. Comparing with the results by HKO method, the results of our proposed scheme is much smoother.

6. The proposed scheme is faster than the EMD-CIIT method, but it is slower than the other methods.

For example, there are some fluctuations in the range of azimuth angles from 257° to 359° in Figure 8(g), but the data are much smoother in Figure 8(h).

To further evaluate the performance of the proposed method, we show a nonwindshear case in Figure 9. In the figure, we plot the conical scans of the LiDAR data and the denoising results by the abovementioned six methods. The LiDAR data were collected on 4 March 2015 from 00:00:16 to 00:00:39. Similarly, the denoising results given by the proposed scheme do not contain any outliers and retain more data points with good smooth effect.

3.6. Comparison Results in Noisy Observations Removal.

Finally, we would like to demonstrate the capability of the proposed method in removing noisy observations compared with the “Newsom” method, the “Baranov” method, as well as
the HKO method. As a matter of fact, when SNR values are too small, the LiDAR observations may not be suitable for data analysis usage. In Table 1, we show the noisy observation removal results for a set of LiDAR observations for windshear cases at Hong Kong International Airport. It here was collected from 1 March to 31 March 2015. It here were 147 windshear cases reported in the pilot report. Each scan of windshear cases took about 25 seconds (see the two examples in Figures 7 and 8). We see from Table 1 that

(1) Since the SNR threshold level of the "Newsom" method is 0.02, all the noisy observations are removed in the results of the "Newsom" method, but the number of the retained data points are really small.

(2) The number \( \sum_{i=1}^{m} \sum_{j=1}^{n} w_{i,j} \) of denoising values by the proposed method is about the same as the number of original observations \( y_{i,j} \) and the number of denoising values \( x_{i,j}^{(\text{Baranov})} \) by the "Baranov" method and the number of denoising values \( x_{i,j}^{(\text{HKO})} \) by HKO method in the slant range values from 359 m to 3509 m. Note that the SNR values of LiDAR observations in the slant range values from 359 m to
3509 m are usually high, which means LiDAR observations are quite accurate.

(3) There are usually more noisy observations in large slant range values. We see from Table 1 that there are only 281927 LiDAR observations in the slant ranges from 6659 m to 9809 m and 793282 LiDAR observations in the slant ranges from 3509 m to 6659 m compared with 1033268 LiDAR observations in the slant ranges from 359 m to 3509 m.

(4) The number $\sum_{m}^{M} \sum_{j=1}^{J} w_{i,j}$ of denoising values by the proposed method is less than the number of original observations $y_{i,j}$ and the number of denoising values $x_{i,j}^{(Baranov)}$ by the “Baranov” method and the number of denoising values $x_{i,j}^{(HKO)}$ by the HKO method when there are many noisy observations in the slant ranges from 6659 m to 9809 m. However, the proposed method preserves more data points with high SNR values in the slant ranges from 6659 m to 9809 m.

Figure 8: (a) The LiDAR data and (b) SNR values collected on 5 March 2015 from 04:16:02 to 04:16:25. The denoising results by (c) the “DWT” method, (d) the “EMD-CIIT” method, (e) the “Newsom” method, (f) the “Baranov” method, (g) the HKO method, and (h) the proposed scheme (denoted by TV).
Figure 9: (a) The LiDAR data and (b) SNR values collected on 4 March 2015 from 00:00:16 to 00:00:39. The denoising results by (c) “DWT” method, (d) “EMD-CIIT” method, (e) “Newsom” method, (f) “Baranov” method, (g) HKO method, and (h) the proposed scheme (denoted by TV).

Table 1: The comparison of noisy observations removal by the “Newsom” method, “Baranov” method, HKO method, and proposed method.

<table>
<thead>
<tr>
<th>Range</th>
<th>Data</th>
<th>( \text{SNR} &gt; -5 )</th>
<th>( \text{SNR} &gt; -3 )</th>
<th>( \text{SNR} &gt; 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiDAR observations</td>
<td>1033268</td>
<td>1032885</td>
<td>1031484</td>
<td></td>
</tr>
<tr>
<td>Denoising values by the “Newsom” method</td>
<td>1031472</td>
<td>1031472</td>
<td>1031472</td>
<td></td>
</tr>
<tr>
<td>359 m–3509 m (total number: 1036067)</td>
<td>Denoising values by the “Baranov” method</td>
<td>1033268</td>
<td>1032885</td>
<td>1031484</td>
</tr>
<tr>
<td>Denoising values by the HKO method</td>
<td>1033212</td>
<td>1032874</td>
<td>1031484</td>
<td></td>
</tr>
<tr>
<td>Denoising values by the proposed method</td>
<td>1032987</td>
<td>1032793</td>
<td>1031394</td>
<td></td>
</tr>
</tbody>
</table>
4. Conclusion

In this article, we have proposed a LiDAR data denoising scheme for wind profile observations to improve data quality and remove bad observations. By introducing a weight in each observation, the scheme can filter out very noisy observations according to their SNR values and neighborhood observations. Combining with the data-fitting term and the polar-based total variation smoothing term, a global-based denoising model is built. The alternating direction method of multipliers is applied to find a solution of the proposed model. To find suitable regularization parameters of the proposed scheme, we consider an L-curve based parameter selection method that can select parameters automatically via the balance between the data-fitting term and the polar-based total variation regularization term. By applying the proposed scheme to the LiDAR data collected at the Hong Kong International Airport, we find that (i) our proposed scheme performs better than the denoising methods such as DWT and EMD-CIIT, where they cannot handle very noisy observations and they can conduct a denoising procedure along each slant range; (ii) our proposed method can handle noisy observations quite well, and its performance is better than that of HKO method and “Newsom” method as well as “Baranov” method; (iii) our proposed scheme can balance the data-fitting and the smoothing regularization via suitable parameter selection. However, the proposed scheme is slower than the comparing methods except the EMD-CIIT method. As a future research work, we will investigate a much faster algorithm for the proposed model. And we also need to consider an online denoising scheme such that it can handle LiDAR data sets in a continuous-time manner and study three-dimensional LiDAR data observations for denoising.

Data Availability

The request for data used to support the findings of this study could be addressed to the Hong Kong Observatory. The data provision would be considered on a case-by-case basis.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The research was supported in part by HKRGC GRF 17201020 and 17300021 and UGC-RMGS 207300829.

References


