COMPARISON OF WAVEFORM DECOMPOSITION METHODS FOR AIRBORNE FULL-WAVEFORM LIDAR

Yu-Shan Lin[1] and Tee-Ann Teo [2]
[1] Graduate student, Department of Civil Engineering, National Chiao Tung University,
1001, University Road, Hsinchu, Taiwan 30010; Tel: + 886-3-5712121#54982;
E-mail: pitfall.cv95@g2.nctu.edu.tw

[2] Assistant Professor, Department of Civil Engineering, National Chiao Tung University,
1001, University Road, Hsinchu, Taiwan 30010; Tel: + 886-3-5712121#54929;
E-mail: tateo@mail.nctu.edu.tw

KEY WORDS: Full-waveform, LIDAR, multi-echo, Gaussian distribution, Weibull distribution

ABSTRACT: A full-waveform lidar system is able to record received signal continually for analysis. As it provides more information than the conventional discreted lidar, the waveform analysis is a key process to extract the implied information. The objective of this study is to analyze the received waveform as well as its parameter extraction. The waveform parameters include position, amplitude, and echo width. This study selects Gaussian distribution as a symmetric function and Weibull distribution as an asymmetric function to decompose the return waveform into a series of components. In data preprocessing, we use Gaussian smoothing to reduce the noise effect and differentiation of waveform to extract the initial echo peak. Then, we employ Trust Region algorithm to solve the non-linear optimization problem for Gaussian and Weibull distributions. The waveform attributes such as peak, amplitude, and echo width are further extracted by the extracted components. Finally, we calculate the residuals, RMSE, R-square between raw waveform data and fitting function for quality assessment. The test site is a mountain area located in the central part of Taiwan. The experimental data was acquired by Leica ALS60 system and was recorded in the format of LAS1.3. This study compares two different functions in waveform decomposition. The result of Weibull distribution is better than Gaussian distribution as this asymmetry distribution is more suitable for lidar waveform. However, the latter is relatively simple and easy to implement in the waveform analysis.

1. INTRODUCTION

Full-waveform (FWF) lidar is a new generation of airborne laser scanner which receives one dimensional continuous signal. It offers useful information about the structure of the target. Therefore, the analysis of return signal of FWF lidar and obtaining the implicit information is a very important issue. The discrete lidar only contains three dimensional coordinates and intensity of point. On the other hand, full-waveform lidar record continuous waveform for each pulse. Each returned signal represents the vertical characteristics of surface features. These returned signals may obtain several coefficients such as echo width, amplitude and etc. Hence, full waveform lidar may increase the capability of landcover classification as it contains more information than discrete lidar.

Waveform decomposition is used to extract the peak and parameters of full waveform lidar. Several papers have been reported on this issue. The major steps of waveform decomposition include mathematic model selection, models fitting, parameters determination and evaluation. The mathematic model selection usually contains symmetric and asymmetric models. Gaussian distribution is the most common and easiest model (Wagner et al., 2006). However, the FWF lidar is a natural asymmetric model due to the change of surface. Figure 1 illustrates the waveform pattern in different shape of surface (Jutzi and Stilla, 2006). Several asymmetrical models have been discussed (Chauve et al., 2007; Mallet et al., 2009), for example General Gaussian distribution, Log Normal distribution, Weibull distribution, Nakagami distribution and Burr distribution. Weibull distribution is the most common and relatively simple asymmetrical model (Coops and Hilker, 2007). Many studies have focused on the comparison of the decomposition results, relatively few studies have discussed about extracted parameters. Hence, the objective of this study is to compare the decomposition results as well as the waveform parameters.

This study uses both symmetric and asymmetric functions to decompose full waveform lidar and extracts their waveform parameters. We select Gaussian distribution as a symmetric function and Weibull distribution as an asymmetric function. The quantitative evaluation analyzes the extracted parameters and fitting accuracies. Besides, we also compare the peak location between the proposed methods and original data.

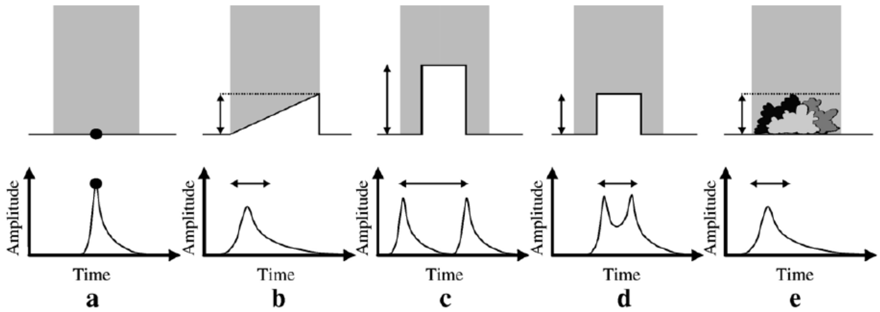

Figure 1. Surface characteristic and pulse form. (Jutzi and Stilla, 2006)

2. METHODOLOGY

Figure 2 is the workflow of this study. There are four main parts. In data preprocessing, we use Gaussian smoothing to reduce the noise effect. Then, we extract the initial echo peak to determine the initial parameters. Then, we employ Trust Region algorithm to solve the fitting curve for Gaussian and Weibull distributions and obtain the waveform attributes such as peak, amplitude, echo width, scale factor and shape factor. Finally, we perform accuracy analysis between raw waveform data and fitting function for quality assessment.

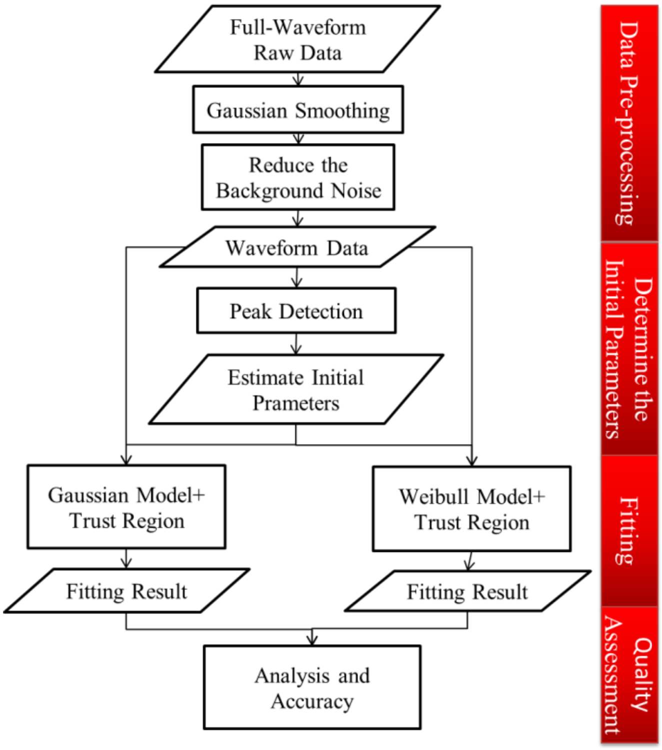

Figure2. Workflow

2.1 Data Preprocessing

There are two major steps in data preprocessing. First, we use Gaussian smoothing to remove the high frequency noise in the waveform (Hofton et al., 2000), as the serrated original data will affect the fitting results. Second, we eliminate the background noise which is smaller than the signal-to-noise level (Chauve, etc., 2007).

2.2 Determine the Initial Parameters

As waveform decomposition is a non-linear fitting problem, initial parameters may affect the results. The initial parameters include number of peak, initial peak and initial echo width. This study uses different waveforms to find the place of zero crossing (Wagner, etc., 2006). Then, we can obtain the number of peak and peak positions. For the echo width, it is the twice of the distance between the minimum and maximum points. Both symmetric and asymmetric functions use the same initial parameters.

2.3 Model Fitting

We select Gaussian distribution as a symmetric function (Chauve, etc., 2007) and Weibull distribution as an asymmetric function (Coops and Hilker, 2007) to fit the return waveform. The followings explain the two models in detail. Figure 3 compares the symmetric and asymmetric functions. The red line represents the symmetric function (Gaussian distribution) while the blue line represents the asymmetric function (Weibull distribution). This figure shows that the asymmetric function is more suitable for FWF lidar in some cases.

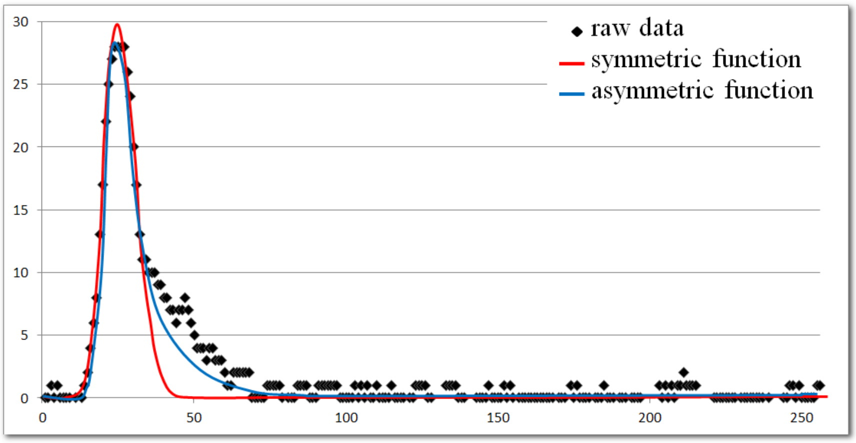

Figure 3. Comparison of symmetric function and asymmetric function

The Gaussian model is shown as equation 1. The waveform may decompose into many Gaussion models in different parameters. The additional amplitude is added for model fitting as in equation 2. Equation 3 is Weibull function. The extended model for decomposition is shown as equation 3. Notice that the time *x* are positive value in FWF lidar.

$$y = e^{-\left(\frac{x}{2\sigma^2}\right)} \tag{1}$$

$$y = a \times e^{-\left(\frac{x}{2\sigma^2}\right)} \tag{2}$$

$$\begin{cases} y = \frac{\kappa}{\lambda}\left(\frac{\kappa}{\lambda}\right)^{\kappa-1} e^{-\left(\frac{x}{\lambda}\right)^{\kappa}} & , x \geq 0 \\ y = 0 & , x < 0 \end{cases} \tag{3}$$

$$y = a \times \frac{\kappa}{\lambda}\left(\frac{\kappa}{\lambda}\right)^{\kappa-1} e^{-\left(\frac{x}{\lambda}\right)^{\kappa}} \tag{4}$$

Where,
- x: time in ns
- y: receipted signal in V
- a: amplitude
- σ: echo width
- κ: scale parameter
- λ: shape parameter

Trust region algorithm is a method to deal with general bound-constrained optimization problems. We use this algorithm to calculate the coefficients of fitting models by minimizing the summed square of residuals (Lin et al., 2010). At each iteration for minimizing $f(\mathrm{w})$, the algorithm has an iterate w^k, a size of the trust region ∇_k, and a quadratic model (Lin et al., 2008).

$$q_k(\mathrm{s}) = \nabla f(\mathrm{w}^k)^T\mathrm{s} + \frac{1}{2}\mathrm{s}^T\nabla^2 f(\mathrm{w}^k)\mathrm{s} \tag{5}$$

2.4 Analysis and Quality Assessment

We use three indexes to analyze the results. There are sum of squares due to error (SSE), root mean squared error (RMSE) and R-square. SSE is the sum of squares error due to the difference between the raw data and the fitting data. Equation (6) shows the detail of SSE. RMSE is the root mean squared error between raw data and fitting data. RMSE is shown as equation (7). R-Square is the coefficient of determination as shown as equation (8). The ideal value of SSE and RMSE is close to zero. R-square is between zero and one, the ideal value of R-square is close to one.

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6}$$

$$RMSE = \sqrt{\frac{SSE}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}} \tag{7}$$

$$R^2 = 1 - \frac{SSE}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2} \tag{8}$$

Where,

y_i: observed data
$\hat{y}_i$: reference value
n: number of the observed data
$\bar{y}_i$: mean of the observed data

3. EXPERIMENTAL RESULTS

3.1 Test Data

The experimental data was acquired by Leica ALS60 system and was recorded in the format of LAS1.3. The test site is a mountain area located in the central part of Taiwan. We subset a small region contain ground and trees as shown in Figure4. Table 1 shows the numbers of return. We compare two different models in waveform decomposition and analyze the relationship between parameters.

Table 1. Numbers of Return

Numbers of Return	Number
Total returns	1981
One return	1906
Two returns	74
Three returns	1

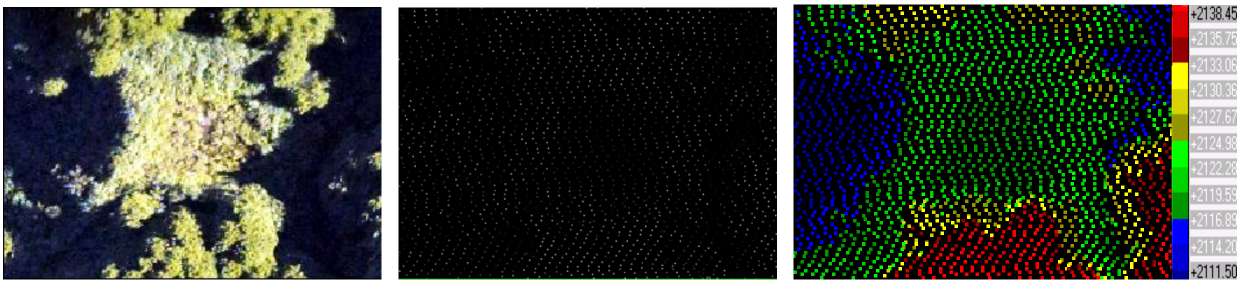

Figure 4. Test Area (Left: Aerial Image, Middle: LIDAR point cloud with intensity, Right: LIDAR point cloud with elevation.)

3.2 Waveform Parameters

The empirical thresholds in data preprocessing are window size at 9, σ at 5 and background noise at 13. Table 2 shows the difference between the 3D coordinates from original data and the proposed methods. The extracted 3-D position is similar to the original data. The airborne lidar acquired data vertically. Hence, the mean difference is in z direction is larger than X and Y directions. Moreover, the mean difference in vertical direction is about 0.3m as the time resolution is about 1ns.

Table2. The difference between the raw data position and the extracted position (unit: meter)

	Mean(X)	Std.(X)	Mean(Y)	Std.(Y)	Mean(Z)	Std.(Z)
Gaussian Model	0.0138	0.0116	0.0312	0.0244	0.3004	0.2338
Weibull Model	0.0175	0.0085	0.0396	0.0154	0.3805	0.1478

Figure 5 compares the waveform parameters. We find a high correlation between amplitude and intensity. Echo width is related to land cover. The greater the echo width, the smaller the amplitude, when the energy is dispersed. Dense trees are hard to penetrate and they usually have only one echo, so the echo width is smaller. In summary, land cover which is rough and complex has high echo width, and negative correlation with amplitude. In Weibull Model, the meaning of scale parameter is similar to the echo width in Gaussian Model. Shape parameter determines the asymmetry of the fitting curve. The value of shape parameter that is less than 3.44 is skewed to the left, and the one that is more than 3.44 is to the right. It is related to the slope of terrain.

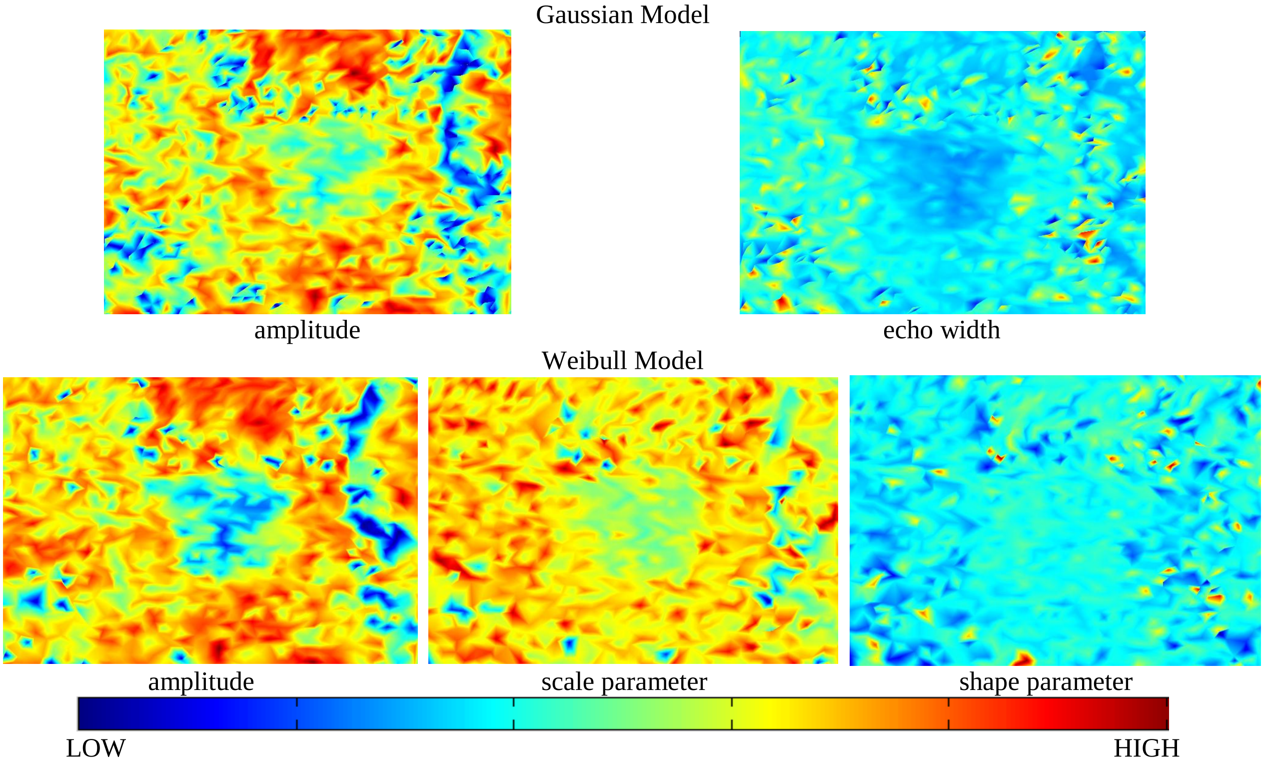

Figure 5. Waveform Parameters

3.3 Analysis and Quality Assessment

Table 3 summaries the results of quality assessment. It shows that Gaussian Model extracts more peaks but Weibull Model's fitting result is better. Figure 6 shows the histogram of RMSE. RMSE of Weibull model is smaller than Gaussian model. In other word, the fitting of Weibull Model is better than Gaussian Model. To compare their computation time, Gaussian is faster than Weibull because of the less unknown parameters.

Table 3. Quality Assessment

	Gaussian Model	Weibull Model
Raw data peaks	1923	1832
Increased peaks	137	87
SSE	255.0697	151.9003
RMSE	0.9982	0.7703
R-Square	0.9879	0.9927

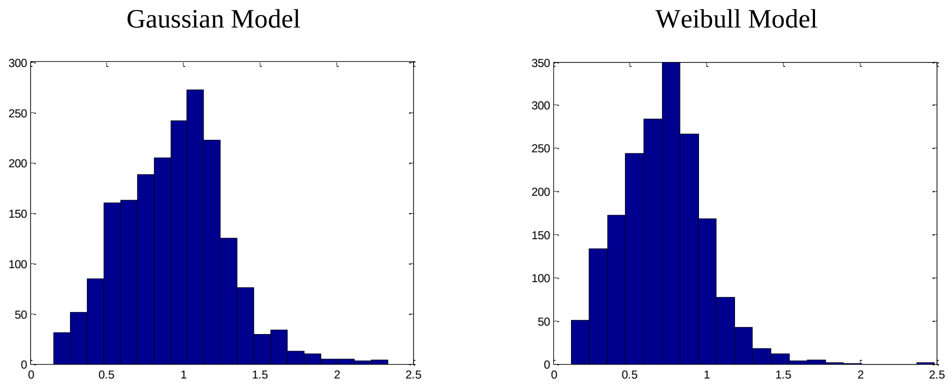

Figure 6. Histogram of RMSE (Left: Gaussian, Right: Weibull)

4. CONCLUSIONS

This study presented a procedure of waveform decomposition which includes data pre-processing, initial value determination, waveform fitting, analyzing and quality assessment. We have extracted the peak position as well as waveform parameters like amplitude, echo width, scale parameter, shape parameter, and XYZ location. The experimental results indicated that we can extract more peaks and get more information when compared to the original data. Moreover, the asymmetry distribution has better fitting results than symmetry distribution. Gaussian distribution need to extracts more peaks for returned signal. In brief, the result of Weibull distribution is better than Gaussian distribution as this asymmetry distribution is more suitable for lidar waveform. However, the latter is relatively simple and easy to implement in the waveform analysis.

ACKNOWLEDGEMENT

The authors would like to thank the LIDAR Technology Company (LIDAR-Tech) in Taiwan for providing the test data sets.

REFERENCES

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Deseilligny, M.P. and Puech, W., 2007.Processing full-waveform lidar data: modelling raw signals. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/W52): 102-107.

Coops, N. C., T. Hilker, 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. Trees - Structure and Function 21(3): 295-310.

Hofton, M., Minster, J. and Blair, J. B., 2000. Decomposition of laser altimeter waveforms, IEEE Transactions on Geoscience and Remote Sensing, 38(4): 1989- 1996.

Jutzi, B. and Stilla, U., 2006. Range determination with waveform recording laser systems using a Wiener Filter. ISPRS Journal of Photogrammetry and Remote Sensing, 61: 95-107.

Lin C. J., Weng R. C. and Keerthi S. S., 2007. Trust region Newton methods for large-scale logistic regression, Proceedings of the 24th international conference on Machine learning, 561-568.

Lin, Y. C., Mills, J. P. and Smith-Voysey, S., 2010. Rigorous pulse detection from full-waveform airborne laser scanning data. International Journal of Remote Sensing, 31(5):1303-1324.

Mallet, C., Lafarge, F., Bretar, F., Roux, M., Soergel, U. and Heipke, C., 2009. A Stochastic Approach for Modelling Airborne Lidar Waveforms. Laser scanning 2009, ISPRS, Vol. XXXVIII, Part 3/W8 – Paris, France, September 1-2, 2009, 201-206.

Stilla, U. and Jutzi, B., 2009. Waveform analysis of small-footprint pulsed laser systems. In:Shan, J., Toth, C.K. (Eds.), Topographic Laser Ranging and Scanning: Principles and Processing, Taylor & Francis Group, London, 215-234.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T. and Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry & Remote Sensing, 60: 100-112.