Bathymetric LiDAR Waveform Decomposition with Temporal Attentive Encoder-Decoders

Mathilde Letard, Thomas Corpetti, Dimitri Lague

BATHYMETRIC LIDAR WAVEFORM DECOMPOSITION WITH TEMPORAL ATTENTIVE ENCODER-DECODERS

Mathilde Letard[1], Thomas Corpetti[2] and Dimitri Lague[1]

[1] CNRS, Univ Rennes, UMR 6118 Geosciences, Rennes, France
[2] CNRS, UMR 6554 LETG, F-35000 Rennes, France

ABSTRACT

This paper is concerned with the decomposition of bathymetric lidar waveforms. Because of the presence of water, processing such data remains a challenge since water impacts their shape and signal-to-noise ratio, depending in particular on the associated turbidity. In this paper, we explore the use of attentive autoencoders to decompose bathymetric waveforms and recover their air/water interface, water column, and water bottom components simultaneously, without relying on assumptions about the impulse or target surface nature. On simulated waveforms, the method achieves lower decomposition error than existing approaches, handling overlapping echoes of very shallow waters and weak returns in deeper water. This opens to attractive strategies to process real bathymetric waveforms.

Index Terms— full-waveform lidar, bathymetric lidar, signal decomposition, waveform decomposition, attention UTime

1. INTRODUCTION

Though airborne LiDAR data are mainly used as 2D products or 3D point clouds, their origin lies in the processing of backscattered laser signals and their conversion into spatial information. The precise extraction of the various signal components linked to the hit targets is thus crucial to interpreting with precision the observed scene. However, this extraction entirely relies on the ability both to process the noise efficiently and to remain sensitive enough to subtle variations denoting low-reflective or vertically close targets. Research on full-waveform lidar data processing revealed the potential of finer analysis to retrieve additional targets and increase the resulting point clouds' density [1]–[3]. However, handling large datasets while minimizing the level of false detections – which produce noise in the point clouds and complexify their interpretation – is still a challenge, namely for airborne bathymetric lidar (ALB). ALB uses green lasers to penetrate water and produce bathymetric data. The resulting waveforms have two main peaks: one generated by the water surface and another occurring when, and if, the water bottom is hit [4], [5]. ALB signals have to compose with additional sources of noise due to the optical scattering of the laser beams in the water column, which 1) exponentially attenuates the reflected intensity of the water bottom and 2) reflects a portion of the signal towards the receptor without signifying a solid target, resulting in a third waveform component that elongates the surface peak [4], [5]. In addition, extracting meaningful information from waveforms is not trivial because of the noise that disrupts the signal and gets mixed up with weak information, which is often lost during processing. In this paper, instead of a deconvolution by inversion of the pulse response function that can be challenging due to the noise in the signal, we explore the use of deep neural networks to tackle these problems and perform bathymetric waveform decomposition. We use a temporal attentive autoencoder to reconstruct the three main components of the bathymetric lidar signal simultaneously: the water surface, the water column, and the bottom. Our method targets three main limitations of existing waveform processing methods: (1) it does not rely on an iterative process, and (2) it simultaneously recovers the three main bathymetric waveform components.

2. WAVEFORM PROCESSING METHODS

The collected waveform is a sum of sub-signals that each result from a convolution between the emitted pulse and the target surface function [6]. Traditionally, waveform processing is thus made either with **deconvolution** methods [7], [8] or decomposition procedures relying on the **fitting of mathematical functions** to previously detect peaks in the signal [6], [9]. These approaches rely on various assumptions about the components' impulse function or nature. In [10] and [11], the authros show that waveforms can be considered as a sum of Gaussians corresponding to the objects in the illuminated cone. The emitted pulse and the target functions are thus often approximated with Gaussian functions of fixed width or more asymmetric shapes to fit the empirical sensor impulses [6], [12], [13]. When using the decomposition approach, Gaussian mixture models are often used to estimate iteratively the parameters of these Gaussian or other predefined functions, which are then used as proxies on the nature of the targets [14]. Deconvolution procedures include, among others, the Richardson-Lucy deconvolution and the Wiener filter deconvolution. Several comparative studies [8], [15], [16] show the superiority of Richardson-Lucy

deconvolution in retrieving weak returns and recovering the amplitude of the different components but underline its costly iterative process. Several limitations appear with these classical approaches:

(1) Decomposition methods relying on the fitting of mathematical functions are not always adapted to the presence of the water column component, which has an entirely different shape;
(2) Target surface functions may not always be similar/gaussian, and returns may be poorly approximated or mixed;
(3) The denoising methods or post-processing procedures used with these approaches to avoid false detections may lead to loss of information.

3. PROPOSED DECOMPOSITION NETWORK

We adapt U-Time, the 1D version of U-Net introduced in [17] and applied to bathymetric lidar waveforms in [18]. We use the resulting encoder-decoder structure to output three waveform components, as illustrated in Figure 1.

The encoder is a stack of 4 identical blocks comprising two series of 1D convolution, rectified linear unit activation, and batch normalization. Each block is followed by a max pooling operation of size 2, thereby reducing by half the length of the feature vector passed on to the next step. In addition, at each step, the number of convolution filters, initially set to 16, is doubled to increase the network's abstraction ability.

The decoder is a stack of four blocks with a different structure. Each block starts with a cross-attention mechanism, followed by up-sampling of size 2, a typical U-Net skip connection, and two series of 1D convolution, rectified linear unit activation, and batch normalization. Contrary to the encoder, the number of filters at each block is reduced by half, thus returning to 16 in the last convolution layers.

Additive attention, as introduced in [19], is used to better pass information about the waveform across the network and allow the model to focus on features from different representation spaces in the decoder, thereby reconstructing waveform components more effectively. Using this setup, each sample of the decoder layer attends to each sample of the corresponding encoder layer.

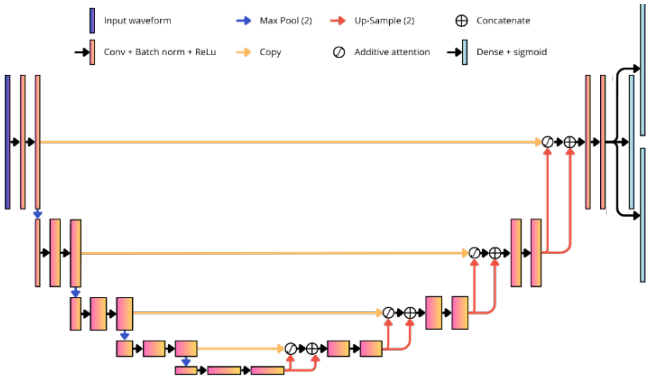

Figure 1. Illustration of the proposed waveform decomposition network architecture.

Between the encoder and the decoder, two series of 128 convolution filters with ReLu activation and batch normalization are applied to the 16 samples-long embedding of the input waveform.

At the end of the decoder, three dense layers with sigmoid activation allow the network to output three sequences corresponding to the three waveform components.

To constrain the network to learn both amplitude and shape relevance when reconstructing the waveform components, we designed a custom loss function, relying on the Mean Absolute Error (MAE) and the Bhattacharyya distance.

$$loss = \left(\frac{1}{N}\sum_{i=0}^{N} abs(pred - true) \times (1 + true)\right) + \left(1 - \sum\sqrt{\frac{pred}{\int pred} \times \frac{true}{\int true}}\right) \quad (1)$$

where N is the number of samples in the waveform, $pred$ is the predicted sequence, and $true$ is the true sequence. Roughly, the idea of the first term is to ensure global consistency in the estimations, this consistency being balanced by the waveform intensity (more weight is applied for locations with high intensities); the second term rather focuses on shapes of the decomposed waveforms. Training is performed on batches of 100 waveforms using an Adam optimizer with a constant learning rate of 0.0001. The model was trained for 500 epochs. One training epoch took, on average, 32 seconds on an Nvidia Quadro RTX 5000 GPU.

4. BATHYMETRIC WAVEFORMS SIMULATION

To simplify the training process (that requires pairs of input waveforms and associated decompositions), to evaluate the approach, and to fully control the physical conditions and their output, we use simulated data obtained with a bathymetric waveform simulator [20]. For training, validation and testing, we simulated three sets of 15000, 3700, and 10000 waveforms, respectively. In each dataset, five types of environments were simulated by varying the simulation parameters (the water body's depth and turbidity). Different impulse functions were used to generalize the method to different sensors and a systematic noise is added. The parameters used are summarized in Table 1.

Parameter	Value(s)	Parameter	Value(s)
Depth (m)	[0.15, 20]	Pulse amplitude	[0.05, 1]
Kd (m^{-1})	[0.08, 0.85]	Pulse width	[0.1, 1]
Bottom reflectance	[0.03, 0.85]	Pulse type	(Generalized) extreme value
Incidence on surface (rad)	[0, 0.8]	Incidence on bottom (rad)	[0, 1.35]
Vertical resolution (m)	0.0626		

Table 1. Parameters used to simulate lidar waveforms.

5. EVALUATION METRICS AND COMPARISON

To evaluate the performance of our method, we compare each retrieved component to its simulated equivalent.
For each of the three components, three metrics are computed. They are defined in equations (2) – (4), where N is the waveform length, $pred$ the predicted sequence, $true$ the true sequence, and KL the Kullback-Leibler divergence.

The weighted MAE, quantifies the mean element-wise amplitude error:

$$wMAE = \frac{1}{N}\sum_{i=0}^{N} abs(pred - true) \times (1 + true) \qquad (2)$$

The Kullback-Leibler divergence measures the relative entropy, or difference in information that two distributions contain. As it is not symmetric, we use a custom metric:

$$mKL = \frac{KL\left(\frac{pred}{\int pred}, \frac{true}{\int true}\right) + KL\left(\frac{true}{\int true}, \frac{pred}{\int pred}\right)}{2} \qquad (3)$$

The Bhattacharyya distance quantifies how far two samples of two distributions are from each other in time and amplitude. For consistency purposes, we used the metric B':

$$B' = 1 - \sum \sqrt{\frac{pred}{\int pred} \times \frac{true}{\int true}} \qquad (4)$$

These three metrics range between 0 and 1, where 0 indicates high similarity and 1 indicates high dissimilarity.

We compare our results to Gaussian Mixture Model (GMM) decompositions with two components – as the water column is not Gaussian, and as in [15] – and a Lucy-Richardson (RL) deconvolution. Other methods will be tested in the future.

Waveform component	Median wMAE	Median mKL	Median B'
Total	0.00333	0.00056	0.00028
Surface	0.00420	0.00040	0.00020
Column	0.00302	0.00056	0.00028
Bottom	0.00333	0.00074	0.00037

Table 2. Results obtained with the proposed approach.

Waveform component	GMM	RL
Total	0.235	0.122
Surface	0.147	/
Column	/	/
Bottom	0.098	/

Table 3. Results obtained with state-of-the-art approaches

5. RESULTS

Results obtained with the best model and with the classical approaches on the test data are presented in Tables 2 and 3.

Overall, the autoencoder reaches lower wMAEs than the GMM and RL methods, although it has higher amplitude errors than distribution distances, which shows that it tends to estimate component shapes better than their amplitude. Analyzing the metrics more deeply shows that there is no relationship between the value of wMAE, mKL and B' and the turbidity of the simulated water body. It also highlights the higher decomposition errors when processing waveforms at depth below 0.5 m. Qualitative results, presented in Figure 2 illustrate these observations: in Figure 2.a), which illustrates very shallow waters, the gap between the prediction and the simulation is broader than in Figures 2.b) and 2.c), and the amplitude error is noticeable, while the peaks appear close to their true locations. Figure 2.c) shows the ability of the network to recover very weak returns in deeper waters. Visual analysis of the corresponding GMM and RL results show that RL misses the bottom return in the cases presented in Figure 2.a) and mistakes noise for returns in the 2.c) one, while GMM with two components retrieves two returns in the first case, and falsely decomposes the surface return into two echoes in the second.

6. DISCUSSION AND CONCLUSION

Our results outline the potential of temporal attentive convolutional neural networks for lidar waveform

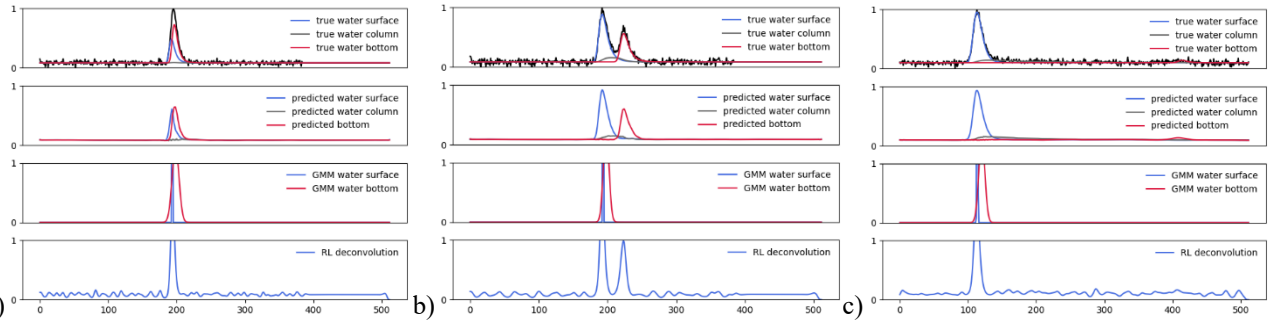

Figure 2. Examples of waveform decomposition results in a) very shallow, b) shallow and c) deep water.

decomposition. Using deep neural networks has the advantage of not necessitating setting fixed parameters. Compared to GMM and RL it is possible to reconstruct the three components of bathymetric waveforms simultaneously. Amplitude errors are lower than with these two traditional methods, and overlapping or weak echoes not captured with GMM and RL are detected.

Considering the correlation between the physical characteristics of the targets and their response functions, decomposing lidar waveforms into a sum of surface functions is of great potential to derive information on the nature or the structure of the objects detected. In the case of bathymetric lidar, it offers the potential to retrieve more information about the bathymetry. In addition, it may pave the way for water column turbidity assessments relying on remotely sensed rather than field measurements. Further developments to add physical constraints in the loss function and better estimate the maximum signal-to-noise ratio supported by the approach are being conducted. Application to real data should also give further information on the usability of attentive autoencoders to process large sets of bathymetric waveforms.

8. ACKNOWLEDGEMENTS

The authors are grateful to the Saur Group, the Région Bretagne and the Nantes-Rennes Topo-Bathymetric LiDAR platform for their financial support.

11. REFERENCES

[1] C. Mallet and F. Bretar, "Full-waveform topographic lidar: State-of-the-art," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 1, pp. 1–16, Aug. 2009, doi: 10.1016/j.isprsjprs.2008.09.007.

[2] C. Mallet, F. Bretar, M. Roux, U. Soergel, and C. Heipke, "Relevance assessment of full-waveform lidar data for urban area classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 6, Supplement, pp. S71–S84, Mar. 2011, doi: 10.1016/j.isprsjprs.2011.09.008.

[3] A. Chauve *et al.*, "Advanced full-waveform lidar data echo detection: Assessing quality of derived terrain and tree height models in an alpine coniferous forest," *Int J Remote Sens*, vol. 30, no. 19, pp. 5211–5228, 2009, doi: 10.1080/01431160903023009.

[4] W. Philpot, *Airborne Laser Hydrography II*. 2019. [Online]. Available: https://ecommons.cornell.edu/handle/1813/66666

[5] D. Lague and B. Feldmann, "Topo-bathymetric airborne LiDAR for fluvial-geomorphology analysis," in *Developments in Earth Surface Processes*, P. Tarolli and S. M. Mudd, Eds., in Remote Sensing of Geomorphology, vol. 23. Elsevier, 2020, pp. 25–54. doi: 10.1016/B978-0-444-64177-9.00002-3.

[6] A. Chauve, C. Mallet, F. Bretar, S. Durrieu, M. Pierrot-Deseilligny, and W. Puech, "Processing Full-Waveform Lidar Data: Modelling Raw Signals," vol. 36 (Part 3/W52), pp. 102–107, Sep. 2007, Accessed: Jan. 03, 2023. [Online]. Available: https://hal-lirmm.ccsd.cnrs.fr/lirmm-00293129

[7] T. Zhou, S. C. Popescu, K. Krause, R. D. Sheridan, and E. Putman, "Gold – A novel deconvolution algorithm with optimization for waveform LiDAR processing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 129, pp. 131–150, Jul. 2017, doi: 10.1016/j.isprsjprs.2017.04.021.

[8] J. Wu, J. A. N. Van Aardt, and G. P. Asner, "A comparison of signal deconvolution algorithms based on small-footprint LiDAR waveform simulation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6 PART 2, pp. 2402–2414, Jun. 2011, doi: 10.1109/TGRS.2010.2103080.

[9] G. Zhou *et al.*, "Comparison Analysis of Five Waveform Decomposition Algorithms for the Airborne LiDAR Echo Signal," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 14, pp. 7869–7880, 2021, doi: 10.1109/JSTARS.2021.3096197.

[10] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka, "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 2, pp. 100–112, Apr. 2006, doi: 10.1016/J.ISPRSJPRS.2005.12.001.

[11] H. J. Zwally *et al.*, "ICESat's laser measurements of polar ice, atmosphere, ocean, and land," *J Geodyn*, vol. 34, no. 3–4, pp. 405–445, Oct. 2002, doi: 10.1016/S0264-3707(02)00042-X.

[12] M. A. Hofton, J. B. Minster, and J. B. Blair, "Decomposition of laser altimeter waveforms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 4 II, pp. 1989–1996, Jul. 2000, doi: 10.1109/36.851780.

[13] B. Jutzi and U. Stilla, "Range determination with waveform recording laser systems using a Wiener Filter," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, no. 2, pp. 95–107, Nov. 2006, doi: 10.1016/J.ISPRSJPRS.2006.09.001.

[14] J. Niemeyer, J. D. Wegner, C. Mallet, F. Rottensteiner, and U. Soergel, "Conditional random fields for urban scene classification with full waveform LiDAR data," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6952 LNCS, pp. 233–244, 2011, doi: 10.1007/978-3-642-24393-6_20/COVER.

[15] C. Wang, Q. Li, Y. Liu, G. Wu, P. Liu, and X. Ding, "A comparison of waveform processing algorithms for single-wavelength LiDAR bathymetry," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 101, pp. 22–35, Mar. 2015, doi: 10.1016/J.ISPRSJPRS.2014.11.005.

[16] C. E. Parrish, I. Jeong, R. D. Nowak, and R. Brent Smith, "Empirical comparison of full-waveform lidar algorithms: Range extraction and discrimination performance," *Photogramm Eng Remote Sensing*, vol. 77, no. 8, pp. 825–838, 2011, doi: 10.14358/PERS.77.8.825.

[17] M. Perslev, M. Jensen, S. Darkner, P. J. Jennum, and C. Igel, "U-Time: A Fully Convolutional Network for Time Series Segmentation Applied to Sleep Staging," *Adv Neural Inf Process Syst*, vol. 32, 2019.

[18] M. Letard *et al.*, "Classification of coastal and estuarine ecosystems using full-waveform topo-bathymetric lidar data and artificial intelligence," *Oceans Conference Record (IEEE)*, vol. 2021-September, 2021, doi: 10.23919/OCEANS44145.2021.9705797.

[19] D. Bahdanau, K. Cho, and Y. Bengio, "NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE".

[20] H. Abdallah, N. Baghdadi, J. S. Bailly, Y. Pastol, and F. Fabre, "Wa-LiD: A new LiDAR simulator for waters," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 4, pp. 744–748, 2012, doi: 10.1109/LGRS.2011.2180506.