

# Exponential Decomposition with Implicit Deconvolution of Lidar Backscatter from the Water Column

**Type:** Peer reviewed journal article.

**Status:** Accepted April 26, 2017, published July 6, 2017.

**Version:** Manuscript accepted for publication.

**Copyright:** Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e.V.

**Bibliography:** Roland Schwarz et al. “Exponential Decomposition with Implicit Deconvolution of Lidar Backscatter from the Water Column.” In: *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 85.3 (July 6, 2017), pp. 159–167. ISSN: 2512-2789, 2512-2819. DOI: [10.1007/s41064-017-0018-z](https://doi.org/10.1007/s41064-017-0018-z)



# Exponential Decomposition with Implicit Deconvolution of Lidar Backscatter from the Water Column

Roland Schwarz<sup>1</sup> · Norbert Pfeifer<sup>2</sup> · Martin Pfennigbauer<sup>1</sup> · Andreas Ullrich<sup>1</sup>

Received: 3 January 2017 / Accepted: 26 May 2017

**Abstract** Bathymetric laser scanning is a powerful tool to obtain information about the morphology of coastal, river, and inland waters. Laser scanning in general is a method to sense the shape of remote objects by sweeping a laser beam across the objects while measuring the distance to every surface point. In bathymetric applications the electromagnetic light wave also needs to penetrate the water column resulting in a spread reflection from below the surface of the water body complicating the interpretation of the received wave. As the signal seen by the sensor's receiver is the result of a convolution of the system waveform with the differential backscatter cross-section, one approach is to use a deconvolution method to recover the object shape. An alternative approach is to fit a parametrised model to the measured receiver signal. While deconvolution methods are not capable to directly deliver object parameters such as distance to water surface or bottom, modelling methods suffer from neglecting the system waveform. We present a new waveform decomposition method that avoids current shortcomings. The

proposed method uses a model composed of segments of exponential functions, which is motivated by the physics of the backscatter process in the water column, and a record of the system waveform which is stored as part of the sensor's calibration data. The method further consists of an algorithm which evaluates the parameters of the exponential model while at the same time performing a deconvolution from the system waveform in an implicit manner. The effectiveness of the method is exemplified using real data from a near-shore airborne LIDAR data acquisition.

**Keywords** Laser radar, Sea surface, Bathymetry, Transient response, Deconvolution

**Zusammenfassung** *Zerlegung von rückgestreuten Lasersignalen mittels Exponentialfunktionen mit impliziter Dekonvolution.* Bathymetrisches Laserscanning dient zur Erfassung der Morphologie küstennaher Gewässer, von Binnengewässern und von Flüssen. Laser Scanning im Allgemeinen ist eine Methode zur Erfassung von Objekten mit Hilfe eines abtastenden Laserstrahls bei gleichzeitiger Distanzmessung zu den Oberflächenpunkten. In der Bathymetrie wird die Auswertung erschwert, da die elektromagnetische Lichtwelle auch (obzwar erwünscht) in den Wasserkörper eindringt und dadurch eine auseinanderlaufende Rückstreuung von Bereichen unter der Wasseroberfläche verursacht. Da das vom Sensor gemessene Signal der Faltung der System-Wellenform mit dem differentiellen Rückstreuquerschnitt entspricht, kann die Objektform mittels Entfaltung (Dekonvolution) erhalten werden. Eine alternative Methode ist es, ein parametrisches Modell an die Empfängerdaten anzupassen. Während die Entfaltungsmethode nicht direkt Parameter wie die Distanz zur Wasseroberfläche oder zum Grund liefert, vernachlässigen parametri-

---

✉ Roland Schwarz  
rschwarz@riegl.com

Norbert Pfeifer  
norbert.pfeifer@geo.tuwien.ac.at

Martin Pfennigbauer  
mpfennigbauer@riegl.com

Andreas Ullrich  
aullrich@riegl.com

<sup>1</sup> Riegl Research Forschungs Gesellschaft mbH, 3580 Horn, Austria

<sup>2</sup> Department of Geodesy and Geoinformation, Technische Universität Wien, Gußhausstr. 27-29, 1040, Vienna, Austria

sche Modelle den Einfluss der System-Wellenform. Wir präsentieren eine verbesserte Methode zur Zerlegung für Wellenformen, die diese Nachteile vermeidet. Motiviert durch die Physik der Rückstreuung in der Wassersäule kommt ein Modell aus Segmenten von Exponentialfunktionen ebenso zur Anwendung, wie die System-Wellenform, die aus einer Kalibrierung bekannt ist. Weiters beschreiben wir einen Algorithmus, der die Parameter des Exponentialmodells ermittelt sowie zugleich eine implizite Entfaltung der System-Wellenform bewirkt. Die Wirksamkeit der Methode wird exemplarisch anhand von Daten veranschaulicht, die aus einer luftgestützten LIDAR Aufnahme stammen.

## 1 Introduction

Airborne remote sensing of the ocean has been a well established technology for several decades. In particular, LIDAR (Light Detection and Ranging) systems with green lasers for shallow water bathymetry in coastal areas have been in use since 1980 (Gordon 1980; Guenther 1985). Bathymetric laser scanning is also applied over inland waters, e.g. rivers, to improve the understanding of hydraulic processes comprising sediment transport, river morphology, and ecology (Steinbacher and Aufleger 2013).

Laser scanning in general is an application of RADAR (Radio Detection and Ranging) techniques at the frequencies of light, commonly termed LIDAR. A laser beam composed of a sequence of pulse shaped waveforms is emitted from an airborne platform and swept across a remote surface. The time a single pulse requires to travel from the platform to the surface (target) and back is a measure of the distance by virtue of the known speed of light. Current measurement systems do not only deliver the distances but also store a recording of the sampled full waveform (FW) as seen by the sensor's receiver. Storing and giving access to the FW allows for more control by the end user (Hug et al. 2004) and enhances the possibilities for the interpretation of the physical measurement process (Mallet and Bretar 2008). Together with the exterior orientation of the measurement platform, determined by satellite and inertial navigation systems (Skaloud et al. 2010), a geo-referenced point cloud with additional attributes is computed (Otepka et al. 2013).

At a wavelength range of 400 nm to 500 nm, corresponding to green visible light, lasers are able to penetrate water (Mobley 1994), which can be used for the remote ranging of submerged objects. A green laser impulse shaped waveform, emitted from an airborne platform, will be partly reflected at the air-water boundary,

whereas another part will be transmitted into the water volume. Depending on the surface roughness and the angle of incidence, the back-scatter from the air-water boundary can be observed as a sharp peak in the received waveform. The part of the impulse that penetrates the water surface is exponentially damped during propagation within the water body. If enough signal energy reaches the bottom of the water column, some of the energy is reflected back towards the sensor, causing a second sharp peak. If the impulse is too weak when reaching the bottom, no reflection from the ground will be received. Within the water column, small fluctuations in density, particles and turbulence cause volume scattering, observable as an exponentially decreasing slope of the received waveform (Petzold 1972; Guenther 1985; Mobley 1994). Although there is a body of literature in photogrammetry and remote sensing on LIDAR signal processing, we identified a lack of physically interpretable and effective methods for analysis of bathymetric LIDAR signals, which we address in this paper.

In general, the shape of any received waveform is determined by the dBCS, i.e. the differential back-scatter cross-section of the target, the properties of the laser source, and the receiver electronics (Wagner et al. 2006). Under the condition that the mean distance to the scatterers is much larger than their spatial spread, the received signal can be approximated as the convolution between the system waveform and the dBCS. The system waveform is a property of the measurement system alone and the dBCS resembles the target shape that is independent of the sensor. Interpreting the dBCS as an input signal to a linear time invariant (LTI) system, defined by the system waveform, the signal after the receiver can be understood as a filtered version of the dBCS. From the theory of linear systems it is known (see e.g. Papoulis 1962) that the effect of such a filter is a smoothing and widening of the signal, an introduction of advanced and retarded echoes, and a delay of the signal. These properties explain the delay, the broadening, and the change of relative peak positions with respect to the dBCS observed in received laser waveforms.

The expected outcome of waveform processing is the extraction of certain features, such as range, reflectance, surface roughness, turbidity, and object classification. One of the expected benefits from processing the FW is to improve target discrimination (Jutzi and Stilla 2006), which is of particular interest for shallow water bathymetry. FW algorithms can be categorized into deconvolution and modelling methods. Deconvolution is the method to invert the effects of filtering i.e. the interaction between laser pulse, dBCS, and receiver

characteristics. Modelling, on the other hand, aims to approximate the received waveform using a parameterized function by finding the optimum parameters that typically minimize a least squares error criterion. Decomposition can be understood as a special form of modelling where the parts of the model are modelled by prototype functions. Exponential Decomposition makes use of segments of exponential functions.

Since deconvolution generally tends to suffer from noise amplification, special precautions need to be taken to make the method usable. The result of deconvolution is a waveform that is only dependent on the dBCS, within the bandwidth constraints of the system waveform, meaning the result cannot be arbitrarily sharp. While the outcome of modelling methods is a discrete set of values, such as range, deconvolution methods need a subsequent step to convert the deconvolved waveform to a discrete set of ranges. Examples of explicit deconvolution methods using a Wiener filter or B-Splines can be found in (Jutzi and Stilla 2006) and (Roncat et al. 2011), respectively. It shall be noted that decomposition based on Gaussian decomposition (Wagner et al. 2006) can be interpreted in a deconvolution context as well, but requires that the assumption of a Gaussian shaped system waveform is valid (Wagner et al. 2006; Roncat 2014).

The prototypic shape of a water body consists of three main contributions: the air-water boundary, the water column, and the bottom, also called the benthic-layer. The dBCS of a homogeneous water column alone follows an exponential law where the factor governing decay is related to the effective attenuation coefficient (Guenther 1985; Phillips and Koerber 1984). It may be obvious that only the air-water and the bottom layers can successfully be described by using a Dirac shaped model for the dBCS, while the water column cannot. Indeed, already the first airborne bathymetric system LARSEN 500 (Wong and Antoniou 1991) used a convolution of a Gaussian with an exponential function to model the returned waveform.

Other modelling functions, taking care of asymmetric behaviour, have been used with a stochastic algorithm by Mallet et al. (2009). Their finite-mixture model is able to obtain very good results in the topographic case using a range of model functions, but it is not appropriate for bathymetric waveforms because it has no provision for the exponential decay of the dBCS that can be observed in the water column. They conclude that the set of models should be extended with an exponential decaying function when used for bathymetry.

A comparison of range discrimination performance for three FW algorithms, Gaussian decomposition (GD), EM deconvolution, and a hybrid deconvolution GD approach has been done by Parrish et al. (2011). Their results indicate that “there is no single best waveform strategy”, but they found that an expectation-maximization (EM) based deconvolution algorithm provides the best target separation if the system and noise parameters are known and stable. The authors did not explicitly perform their study with bathymetry signals, but the assumed target model, a sequence of Dirac spikes, is likely to be insufficient for modelling bathymetric waveforms.

Pan et al. (2015) conducted a comprehensive performance assessment of three algorithms for shallow river bathymetry and contributed an algorithm based on the continuous wavelet transform (CWT). The authors conclude that the CWT performed better than Gaussian decomposition or a method using the empirical system response (ESR) described by Hartzell et al. (2015). However, they found it is difficult to obtain good results with any method. The authors attribute a part of these problems to not having modelled the influence of the water column scattering and suggest focussing on this problem in future work. They found a bias of water surface determination with respect to Real-Time Kinematic (RTK) GPS observations in the range from 0.63 m to 0.82 m. A closer investigation of the waveforms indicated that the actual position of the water surface could be assigned to the leading edge of the observed waveforms.

The work of Abady et al. (2014) explicitly assesses the contribution of the water column using a six-parametric quadrilateral model function leading to a significant reduction of estimation bias on a simulated data set.

The influence of the water bottom geometry has been studied by Bouhdaoui et al. (2014). Using a simulated data set, they demonstrate a systematic underestimation of depth for a water bottom which exhibits roughness which is comparable in scale to the laser beam footprint. They conclude that in that case the range measurement of the bottom target should not be based on the peak of the corresponding waveform.

Current methods to process laser waveforms either are based on a deconvolution method followed by a discrete point scatter model, or they are based on a modelling approach, but neglect the influence of the system waveform. Both approaches are insufficient to properly describe waveforms resulting from the backscatter of the water body. We present a new method that uses a model of the dBCS that is motivated by physics while performing an implicit deconvolution from the system

waveform simultaneously. We do so by introducing a continuous model for the received waveform which considers the effect of the system waveform, and we use an exponentially damped function to model the dBCS of the water column. Furthermore, we present an efficient method to find the optimum model parameters based on a hierarchical recursive decomposition with a constrained non-linear least-square optimization in the basic step.

## 2 Modeling the Received Waveform

The received waveform  $p(t)$  at the receiver can be written as a convolution integral

$$p(t) = \int_{-\infty}^{+\infty} h(t - \tau)\sigma(\tau) d\tau = h(t) * \sigma(t) \quad (1)$$

where  $\sigma(t)$  is the dBCS and  $h(t)$  is the system waveform. For simplification the dBCS is dependent on a *virtual* time equivalent of range instead of true range, see (Wagner et al. 2006). The system waveform  $h(t)$  (the blurring function) is the convolution of the outgoing laser pulse shape with the impulse response of the receiver electronics. Both  $h(t)$  and  $\sigma(\tau)$  are causal functions of time i.e.  $h(t) = 0$  and  $\sigma(t) = 0$  for  $t < 0$ .

### 2.1 System Waveform

The system waveform (system response) often is modelled as a function of Gaussian shape. When the dBCS also can be modelled by Gaussian shapes, the calculation of the convolution becomes trivial, because the convolution of two Gaussians again is a Gaussian. On the other hand, if the instrument's system waveform is asymmetrical and the dBCS is not properly modelled by a Gaussian, as is the case for bathymetric waveforms where exponential function segments are needed, a different approach is required.

Although it is possible to derive a differentiable expression for the convolution of an exponential function with a Gaussian function, at least approximately (Wong and Antoniou 1991), we use a different approach that allows the system waveform to be of unsymmetrical shape, leads to a simple calculation of the convolution with an exponential function, and allows for storing the shape as a part of the calibration data of a particular instrument. We model the system waveform  $h(t)$  by a sum of damped exponentials

$$h(t) = t_+^0 \sum_{i=1}^I \alpha_i e^{\beta_i t} \quad (2)$$

where  $\alpha_i$  and  $\beta_i$  are the complex valued model parameters,  $I$  is the model order (the number of components) and  $t_+^0$  is the Heaviside step function which is 0 for  $t < 0$  and 1 elsewhere. This is essentially equivalent to modelling the system (transmitter and receiver) by a linear differential equation with constant coefficients. Based on the modified Prony algorithm (Osborne and Smyth 1995) it is possible to determine a low order model of order  $I$  for the system waveform by analysing the receiver output when pointing the sensor towards a flat extended target with a surface normal parallel to the laser beam axis.

Since the system waveform  $h(t)$  is a real valued function the coefficients  $\alpha_i$  and  $\beta_i$  necessarily appear in conjugate pairs, which means that Eq.(2) alternatively can be expressed as a sum of damped harmonic functions

$$h(t) = t_+^0 \sum_{k=1}^K \mathcal{A}_k e^{\gamma_k t} \cos(\omega_k t + \varphi_k) \quad (3)$$

with the real valued model parameters  $\mathcal{A}_k$ ,  $\gamma_k$ ,  $\omega_k$ ,  $\varphi_k$  and  $K$ , which is the number of components<sup>1</sup>. It shall be noted that in order to have the system response approach zero for large  $t$ , the values of  $\gamma_k$  or the real parts of  $\beta_k$ , respectively, have to be negative.

In Fig. 1 the result of modeling the system waveform of a RIEGL VQ-880-G with a two-component exponential sum is shown. The dashed lines are the exponential components and the solid line is the resulting system waveform. The total number of parameters for this particular example with  $K = 2$  is  $2 \times 4 = 8$ .

### 2.2 Backscatter Cross-Section

Since back-scattering from the water column exhibits exponential decay behavior (see e.g. Guenther 1985 or Phillips and Koerber 1984), a simple and physically motivated model for the dBCS is the exponential segment

$$\tilde{\sigma}(t, \tilde{\varphi}) = \begin{cases} E e^{-\gamma(t-\tau)} & \tau < t < \tau + T \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

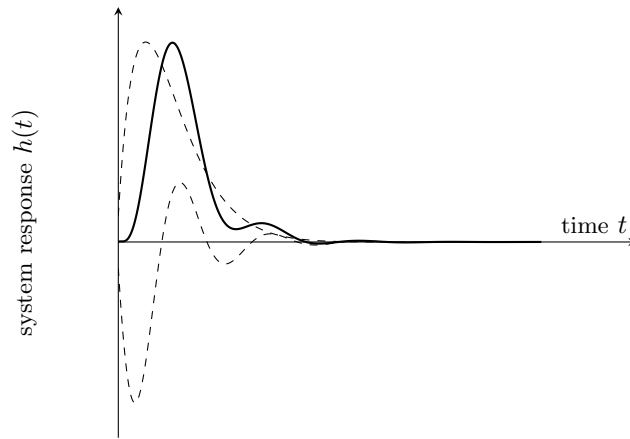
where the vector<sup>2</sup>

$$\tilde{\varphi} = (\tau, E, \gamma, T) \quad (5)$$

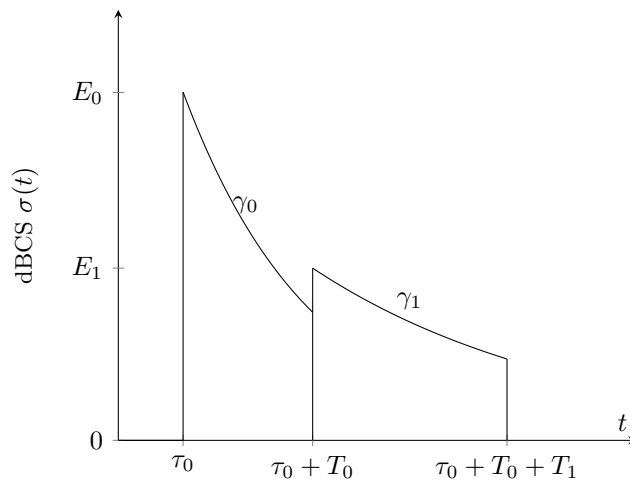
is used to group the parameters, namely the position  $\tau$ , the peak of the dBCS  $E$  (i.e. its maximum), the decay  $\gamma$  and the width of the segment  $T$ . The same expression

<sup>1</sup>  $K$  can be expected to equal  $I/2$ , but may be smaller if some  $\alpha_i, \beta_i$  are real valued.

<sup>2</sup> The tilde is used to indicate an individual component, whereas the composite of multiple individual components uses the same symbol, but without the tilde.



**Figure 1** Exponential-components (dashed) and sum of components of the sensor's system waveform (solid) as a function of time  $t$ .



**Figure 2** A model of the differential backscatter cross-section  $\sigma(t)$  consisting of 2 exponential segments. The particular example has a total of  $3 \times 2 + 1 = 7$  parameters:  $\tau_0, E_0, T_0, \gamma_0, E_1, T_1, \gamma_1$ .

can be used to describe the water surface at the water-air boundary, the bottom echo return, and also pulse spreading as a consequence of oblique angles of incidence of the laser beam. By proper selection of the parameters, the expression is able to model boxcar shaped, exponentially decaying, and almost<sup>3</sup> Dirac-shaped functions. To make up for more complex situations such as inhomogeneities in the water column we use a concatenation of such exponential segments. The parameter vector of the  $n$ -th sub segment

$$\tilde{\varphi}_n = (\tau_n, E_n, \gamma_n, T_n) \quad (6)$$

is indexed by  $n$ . For a set of  $N$  concatenated segments  $\tilde{\sigma}(t, \tilde{\varphi}_n)$  the number of independent parameters is  $3N + 1$ , because the starting and the lengths of the segments cannot be chosen independently. With the

<sup>3</sup> Almost Dirac shaped:  $E$  very large and either  $\gamma$  also very large or  $T$  very small.

four-component parameter vectors  $\tilde{\varphi}_n$  we construct the  $3N + 1$  parameter vector

$$\varphi = (\tau_0, E_0, \gamma_0, T_0, E_1, \gamma_1, T_1, \dots) \quad (7)$$

from which the parameters that are constrained due to concatenation,

$$\tau_n = \tau_0 + \sum_{j=0}^{n-1} T_j \quad , \quad n = 1 \dots N - 1 \quad (8)$$

have been removed. The composite dBCS  $\sigma(t, \varphi)$  finally is given by

$$\sigma(t, \varphi) = \sum_{n=0}^{N-1} \tilde{\sigma}(t, \tilde{\varphi}_n) \quad . \quad (9)$$

Fig. 2 depicts a model of the back scatter cross-section consisting of two exponential segments.

### 2.3 Received Waveform

The simplicity of the dBCS model according to Eq.(4) allows for an explicit evaluation of the convolution with the system waveform Eq.(2):

$$\begin{aligned}\tilde{p}(t, \tilde{\varphi}) &= h(t) * \tilde{\sigma}(t, \tilde{\varphi}) \\ &= Es(t - \tau, \gamma) \\ &\quad - Ee^{-\gamma T} s(t - \tau - T, \gamma)\end{aligned}\quad (10)$$

$$s(t, \gamma) = t_+^0 \sum_i \frac{\alpha_i}{\beta_i + \gamma} (e^{\beta_i t} - e^{-\gamma t})$$

Using the segment  $\tilde{p}(t, \tilde{\varphi})$ , the convolution of the concatenation of multiple exponential segments with the system waveform is simply given by the sum

$$p(t, \varphi) = \sum_{n=0}^{N-1} \tilde{p}_n(t, \tilde{\varphi}_n) = h(t) * \sigma(t, \varphi) \quad (11)$$

which is the received waveform for more complex target scenes. The components

$$\tilde{p}_n(t, \tilde{\varphi}_n) = h(t) * \tilde{\sigma}(t, \tilde{\varphi}_n) \quad (12)$$

are given by Eq.(10), substituting  $\tilde{\varphi}_n$  for  $\tilde{\varphi}$ .

Fig. 3 depicts an example of a synthetic target scenario  $\sigma(t, \varphi)$  with narrow echo contributions from the water surface and the bottom and an exponential decay from the water column, shown as a solid line. The trace of the modelled receive signal  $p(t, \varphi)$  after convolution with the system response is shown as a dashed line. The overall time lag of the received signal w.r.t. the cross-section model is a consequence of the convolution with the system waveform. It can be seen that the delay  $\Delta T_1$  of the peak position of the surface is larger than the delay  $\Delta T_2$  of the bottom peak. The reason for this behaviour is the influence of the exponential part caused by the water column which biases the peak towards the bottom.

### 3 Determination of Model Parameters and Echo Waveform Fitting

The objective is to find the optimum parameter vector  $\varphi = \varphi_{opt}$  that minimizes the target function

$$\chi^2(\varphi) = \sum_{m=0}^{M-1} (y_m - p(t_m, \varphi))^2 \quad (13)$$

under a non-negativity constraint  $(\varphi_{opt})_i > 0$  for all components  $\varphi_i$  of the optimum vector. The  $y_m$  are the sample values of the received waveform at sampling time instances  $t_m$  and  $M$  is the number of samples. In case the number of parameters is known - which is

usually not the case - this is a classical non-linear non-negative least squares problem that can be solved with a Levenberg-Marquardt (LM) algorithm. Non negativity of the parameters is satisfied by a simple parameter mapping  $\varphi(\mathbf{u})$

$$\varphi_i = \sqrt{u_i^2 + 1} - 1 \quad (14)$$

and solving for optimum  $\mathbf{u}$  instead. We note that it is possible to derive explicit expressions for the gradient and second derivatives of the model  $p(t, \varphi)$ , defined by Eq. (11), with respect to the parameters. This property is key to an efficient implementation of the algorithm.

As the number of parameters is not known beforehand, they are determined in an iterative manner by following the approach of Zhu et al. (2012). The authors of the paper find the number of models by successively removing fitted models from the waveform data. They use a peak detector to find the position, amplitude and inflection points of the global maximum peak and use them to seed the parameters which are fit with a non-linear least-squares (NLS) algorithm. We evaluate the first three sample moments of the global maximum peak and use them to calculate the initial parameters in each iteration. One advantage of using moments is that simple relations are available linking the moments of a signal at the input of a filter (the convolution with the system waveform) to the moments of the output signal. Another advantage is that moments are the result of an integrating process which reduces noise in contrast to peak finding, which being based on differentiation, is a noise amplifying process.

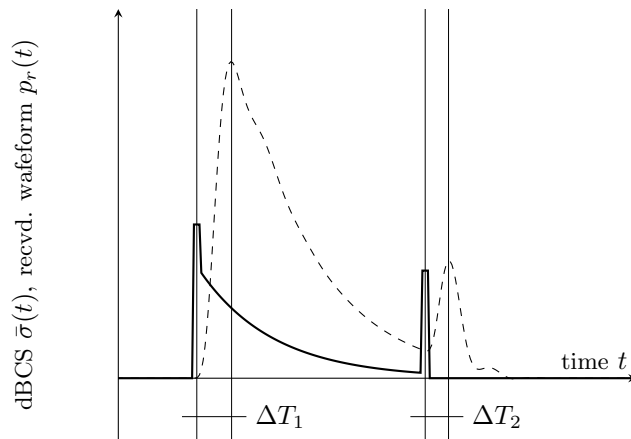
The algorithm starts with the minimum parameter count of four i.e.  $\tau_0, E_0, \gamma_0, T_0$ , by fitting the first model to the most prominent peak in the data. The resulting model is sampled at time instances  $t_m$  and subtracted from the data samples, giving a new sample set.

The reduced sample set again is scanned for the global maximum peak and the set of parameters is extended by three, i.e.  $E_1, \gamma_1, T_1$ . The new parameters are initialized from the samples surrounding the newly found peak. A new fitting process with the full set of parameters, i.e. a model defined by all components found so far, is performed with respect to the original data values.

The algorithm proceeds by repeatedly subtracting the refined model from the original data, identifying the most prominent peak still remaining, and fitting of an expanded model. For each iteration, initial values of (additional) model parameters have to be found. Since

<sup>5</sup> Note that the dBCS is converted from space to virtual time, to show both, received waveform and dBCS on one axis. The rise of the received waveform coincides with the start of the dBCS ( $\tau_0$ ).





**Figure 3** Solid: synthetic cross-section<sup>5</sup> consisting of one narrow echo for surface, one exponential for water column and one narrow echo for bottom; dashed: convolution of the cross-section with the system waveform shown in Fig. 1. Note that  $\Delta T_1 > \Delta T_2$  due to influence of the water column.

in every iteration we consider only the most prominent peak, we can estimate these parameters from a few samples surrounding the peak by calculating the first three sample moments.

The algorithm stops if at least one of several conditions is met: the remaining error is comparable to measurement noise, the improvement of the target  $\chi^2$  is below a threshold, the number of parameters exceeds a predefined value, or the LM does not converge after a maximum number of iterations.

## 4 Results and Discussion

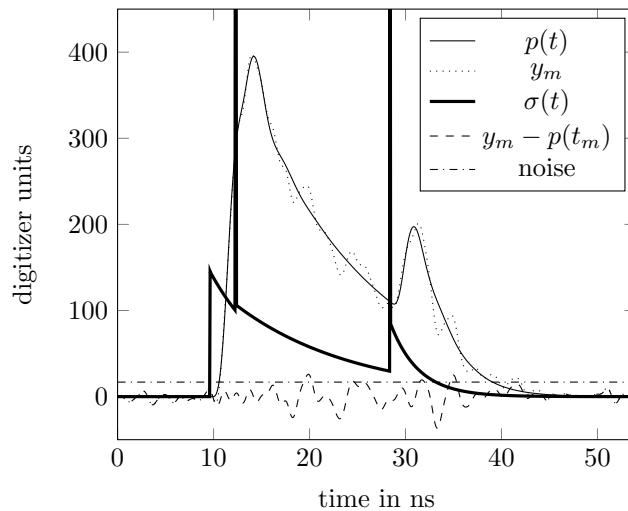
Several Full-waveform data sets of a coastal area acquired with the bathymetric airborne laser scanner RIEGL VQ-880-G (Riegl 2016) from a height above ground of about 600 m have been processed. The data used for this article are by courtesy of the *US National Oceanic and Atmospheric Administration* (NOAA). Due to the lack of reference data or water turbidity information and because the focus of this study was on the theoretical and algorithmic parts of the new decomposition algorithm, we have not carried out a full statistical analysis on the data set, but use only a few selected samples to explain certain features of the algorithm.

The nominal incidence angle of the laser beam was  $20^\circ$ . The system waveform was provided by Riegl as a separate data set. As can be seen in Fig. 1, the waveform shape was not a Gaussian. The waveform including a small trailing second peak could be modelled by a two component exponential sum with a maximum deviation of 1%.

Fig. 4 shows an almost "text book" example of a received waveform, obtained from the sample data, ex-

hibiting two peaks and an exponentially decaying slope. The dotted line follows the measured samples  $y_m$  of the waveform, and the dash-dotted horizontal line marks the 3-sigma noise level as read from the measurement meta data. The thick solid line, labelled  $\sigma(t)$ , represents the estimated model of the backscatter cross section and the reconstructed waveform is shown as solid line  $p(t)$ , the latter being the convolution of  $\sigma(t)$  with the system waveform. The dashed line corresponds to the residual between the measured data and the model. The RMS value of the residuals for this example is 9.47 and the one sigma noise level from the meta data is 5.63 digitizer units.

Since the incident angle of the laser beam was about  $20^\circ$ , no direct reflection from the surface entered the receiver optics, thereby having avoided the non-linear effects that such a strong reflection would have introduced into the receiving electronics. For this reason the first peak is mainly caused by the volume back-scatter alone. The rising edge of the peak can be seen to have a small additional tilt to the right. The algorithm chose two exponential segments and one Dirac for the first part of the reconstructed dBSCS. It can be seen that the start of the front of the dBSCS coincides with the start of the leading edge of the received waveform. The Dirac is most likely caused by some small submerged point scattering object, but without additional data this cannot be determined with certainty. The second Dirac, related to the second peak, can reasonably be attributed to the echo from the ground. It remains an interesting question why the second Dirac is followed by an exponential decay, a behaviour which is observable with quite a number of waveforms. A likely explanation is that the decay is caused by a mixture of volume scattering and the beam hitting the ground at an oblique angle.



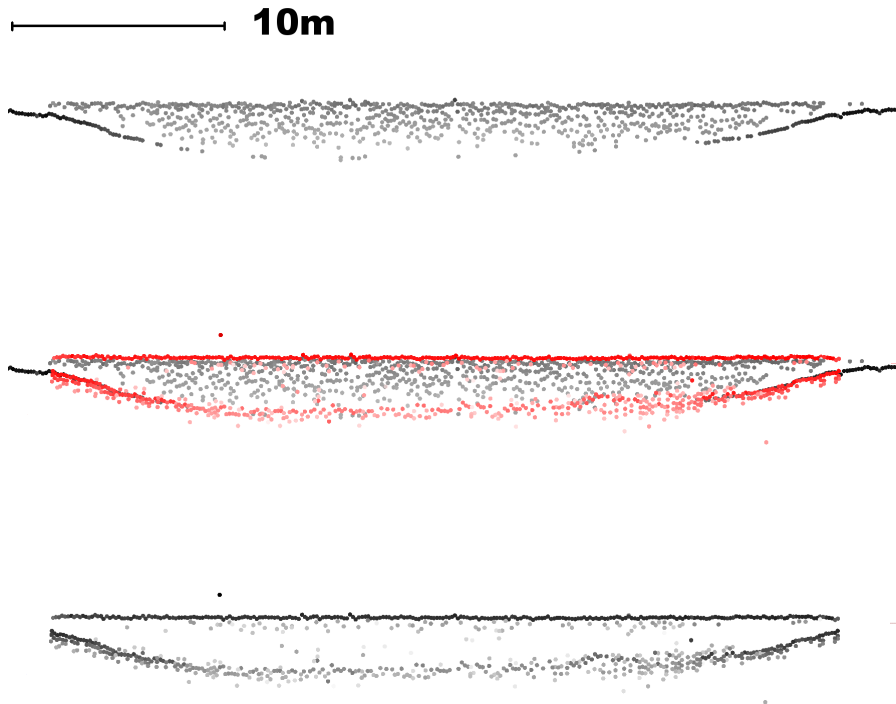
**Figure 4** Decomposed exponential segments, blurred version thereof, and measured data. Residual between data and model.

The particular example was chosen to highlight two important properties of the exponential decomposition algorithm, namely to provide a *good low-dimensional fit* to the measured data and be able to *explain* the data. The good low dimensional fit and the explanation of the data is similar to the approach of Gaussian decomposition of Wagner et al. (2006). They investigated LIDAR data with Gaussian system response over land surfaces, where the dBCS does not show exponential decay, but can be described well by Gaussians. In this case the Gaussian decomposition implicitly deconvolves the recorded waveform, providing position, peak and width of each scatterer. In our contribution, the system response and the dBCS are modelled by sums of (complex) exponential functions. Their convolution is likewise a sum of exponentials. Through implicit deconvolution the parameters position, peak, decay, and width are provided for each scattering element. However the exponential decomposition, having the additional decay parameter, is able to model spread scatterers such as the water column. Specifically it is possible to attribute the leading edge of the dBCS to the water surface and the decay to the effective attenuation coefficient (Phillips and Korber 1984).

We note that the statement of Pan et al. (2015) that “a simple leading edge detection method would be able to accurately estimate the actual water surface” seems to have some kind of resemblance to our findings above, considering that the “leading edge” or onset of the backscatter cross section of the water column accurately defines the position of the water surface, since the front positions of the exponential segments of our model coincide with the start of the leading edge of the received waveform reflected from the water surface (compare Fig. 3). This time instant refers to the

well-known signal-front delay (Papoulis 1962; Ye et al. 2015). However, we want to emphasize that there is a fundamental difference as dBCS describes the target itself whereas the received signal – being the convolution of dBCS and the system response – does not describe the target’s dBCS accurately. The detected time of arrival of the “leading edge” of the received signal relies strongly on the threshold value used. The lower the threshold, the better the time of arrival will resemble the position of the water surface, but at the cost of a poor noise rejection, as only a small fraction of the signal energy is utilized. We expect the implicit deconvolution of our method to be much less susceptible to noise and the front of the recovered dBCS  $\sigma(t)$  to be an unbiased estimate of the water surface. As no reference data was available, in particular of the water surface, which would be essential for a rigorous proof, we tried to assert plausibility of the method by comparing the exponential decomposition result against the on-line measurements of the laser sensor.

The findings were rather encouraging as can be seen in Fig. 5, which shows a single scan line, part of a point cloud of a canal profile, located near the shore. The canal extends from the waterfront into open sea and is entirely submerged i.e. also the banks of the canal are flooded. Although the water surface extends beyond the left and right side of the canal, no points from the water-air boundary are visible with either method. The reason for the lack of surface points is that the laser echoes from the ground become indiscernible from the echoes from the surface when they are spaced too closely. Although the total amount of reflected light is substantially higher for very shallow water than for deeper regions, only the online method is able to resolve target points from the ground. The exponential



**Figure 5** Near shore canal profile comparison of online method with exponential decomposition. Top: online method only - lot of *flying* points within the water column; middle: overlay of both methods - water surface points of decomposition method above online method; bottom: exponential decomposition only - almost no *flying* points.

decomposition, requiring linear behaviour, will not resolve points in these regions because the signals are simply too strong, causing the sensor's receiver electronics to operate in the non-linear regime. The three intensity-coded images, darker points meaning higher amplitude, compare the points provided by the instrument's online waveform processing unit with the points that were derived from the waveforms by the proposed algorithm. The uppermost image contains only points that were acquired with the online processing method of the scanner. The image exhibits a behaviour that is typical for any method relying on threshold detection, namely that on a slowly falling edge of a pulse the likelihood of triggering false targets by noise is increased substantially. The points obtained from the decomposition approach, visible in the lowest part of Fig. 5, do not suffer from such a problem because the exponential tail, the slowly falling edge, is part of a single target i.e. the water volume. The lowest part of Fig. 5 also shows a dense point cloud near the bottom. The sparse point density in the bottom region of the online method, visible in the uppermost image, however is not a fundamental limitation of the method but can be attributed to a configuration problem of the sensor which was limiting the amount of targets per laser shot, so no storage

was left for the points near the bottom. Another observation is that the water surface appears sharper in the decomposition case, which we explain by the fact that the surface is defined by the volume model which is supported by a larger number of points and suffers less from noise for that reason. Compared to the online method, the bottom appears to be not as sharp in the decomposition case. Whether this is caused by bottom roughness or a higher sensitivity to noise is an interesting question which unfortunately has to remain open due to lack of reference data. The middle image of Fig. 5 is an overlay of the top and bottom images. The images have been aligned along the ground-points that are common in both. One of the expectations of the exponential decomposition method is to deliver an unbiased estimate of the water surface, a result which can be seen from the composite image since the points from the exponential decomposition are offset above the points from the online method.

We found that the ability of the algorithm to obtain a good fit in the LM step critically depends on the quality of the initial estimates of the model parameters, which is in accordance with Wagner et al. (2006) and Pan et al. (2015). The use of the sample moments for computing the initial values proved to be a valid

choice, given the amount of successfully decomposed waveforms, although no quantitative assessment has been attempted yet. We have to mention that no pre-filtering step of the data was necessary as is common for methods relying on peak finding and thresholds, a fact that we attribute to the integrating behaviour of the moments calculation which has an implicit de-noising effect.

Finally it must be noted that the exponential decomposition algorithm provides two additional attributes per point: a *decay* and a *width* of the target object. We did not make use of these attributes for this article but we expect them to be a very useful input for classification algorithms.

## 5 Conclusion

In this article the feasibility and effectiveness of modelling the water column, surface, and bottom returns deduced from bathymetric LIDAR waveform data by means of a chain of exponential segments has been investigated and demonstrated.

In contrast to prior work our method is not only able to undo the *blurring* of the target dBCS that is caused by the laser sensor, but it also delivers a stable estimation of the water surface which is essential for a correction of the refraction below the water surface to calculate geometrically accurate target locations. The exponential decomposition in particular makes green-laser-only measurements practical, because it has the property to determine the water surface from the volume echo alone. This would be impossible with a peak detecting method because such a method suffers from a systematic bias. Another beneficial property of our algorithm is that substantially less false echoes from the water column will be triggered. This is a big saving because every detected point must be transformed into the project coordinate system, which is not a computationally cheap operation. Finally the exponential decomposition also delivers a decay parameter which may be physically interpreted and an object width that is likely to be useful as an additional input for classification algorithms.

**Acknowledgements** The work has been conducted in the context of the research project *Alpine Airborne Hydro Mapping* (AAHM). This project investigates potential of advanced LIDAR for surveying the underwater geometry of alpine rivers. AAHM was funded by the Austrian Research Promotion Agency (FFG) within the frame of the Austrian competence centre program COMET-K.

The authors would like to thank Dr. G. Mandlbürger from the Department of Geodesy and Geoinformation of TU Wien

for helpful discussions with respect to the methodological aspects.

## References

- Abady L, Bailly JS, Baghdadi N, Pastol Y, Abdallah H (2014) Assessment of quadrilateral fitting of the water column contribution in lidar waveforms on bathymetry estimates. *IEEE Geosci Remote Sens Lett* 11(4):813–817
- Bouhdaoui A, Bailly JS, Baghdadi N, Abady L (2014) Modeling the water bottom geometry effect on peak time shifting in lidar bathymetric waveforms. *IEEE Geosci Remote Sens Lett* 11(7):1285–1289
- Gordon HR (1980) Ocean remote sensing using lasers. Tech. Memo. ERL PMEL-18, National Oceanic and Atmospheric Administration, URL <http://www.pmel.noaa.gov/pubs/PDF/gord402/gord402.pdf>
- Guenther GC (1985) Airborne laser hydrography, system design and performance factors. NOAA professional paper series, National Oceanic and Atmospheric Administration, URL <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA488936>
- Hartzell P, Glennie C, Finnegan D (2015) Empirical waveform decomposition and radiometric calibration of a terrestrial full-waveform laser scanner. *IEEE Trans Geosci Remote Sens* 53(1):162–172
- Hug C, Ullrich A, Grimm A (2004) Litemapper-5600 – A Waveform-Digitizing Lidar Terrain and Vegetation Mapping System. *Int Arch of the Photogramm Remote Sens Spat Inf Sci*, vol XXXVI - 8/W2, pp 24–29, URL <http://www.isprs.org/proceedings/XXXVI/8-W2/HUG.pdf>
- Jutzi B, Stilla U (2006) Range determination with waveform recording laser systems using a wiener filter. *ISPRS J Photogramm Remote Sens* 61(2):95–107
- Mallet C, Bretar F (2008) Full-waveform topographic lidar: State-of-the-art. *ISPRS J Photogramm Remote Sens* 64(1):1–16
- Mallet C, Lafarge F, Bretar F, Roux M, Soergel U, Heipke C (2009) A Stochastic Approach for Modelling Lidar Waveforms. *Int Arch Photogramm Remote Sens Spat Inf Sci*, vol XXXVIII-3/W8, pp 201–206
- Mobley CD (1994) *Light and Water, Radiative Transfer in Natural Waters*. Academic Press, Inc., California
- Osborne MR, Smyth GK (1995) A modified prony algorithm for exponential function fitting. *SIAM J on Sci Comput* 16(1):119–138
- Otepka J, Ghuffar S, Waldhauser C, Hochreiter R, Pfeifer N (2013) Georeferenced Point Clouds: A Survey of Features and Point Cloud Management. *ISPRS Int J of Geo-Inf* 2:1038–1065
- Pan Z, Glennie C, Hartzell P, Fernandez-Diaz J, Legleiter C, Overstreet B (2015) Performance assessment of high resolution airborne full waveform lidar for shallow river bathymetry. *Remote Sens* 7(5):5133–5159
- Papoulis A (1962) *The Fourier integral and its applications*. McGraw-Hill, New York, USA
- Parrish CE, Jeong I, Nowak RD, Smith RB (2011) Empirical Comparison of Full-Waveform Lidar Algorithms. *Photogramm Eng Remote Sens* 77(8):825–838
- Petzold TJ (1972) Volume scattering functions for selected ocean waters. Technical report, Scripps Institution of Oceanography, URL <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=AD0753474>

- Phillips D, Koerber BW (1984) A theoretical study of an airborne laser technique for determining sea water turbidity. *Aust J Phys* 37(1):75–79
- Riegl (2016) Riegl VQ-880-G. URL [http://www.riegl.com/uploads/tx\\_pxpriegl/downloads/DataSheet\\_VQ-880-G\\_2016-09-16.pdf](http://www.riegl.com/uploads/tx_pxpriegl/downloads/DataSheet_VQ-880-G_2016-09-16.pdf), link checked: 2017-03-21
- Roncat A (2014) Backscatter signal analysis of small-footprint full-waveform lidar data. PhD thesis, TU Wien, Department of Geodesy and Geoinformation, Gußhausstr. 27-29, A-1040 Vienna, Austria
- Roncat A, Bergauer G, Pfeifer N (2011) B-spline deconvolution for differential target cross-section determination in full-waveform laser scanning. *ISPRS J Photogramm Remote Sens* 66(4):418–428
- Skaloud J, Schaer P, Stebler Y, Tomé P (2010) Real-time registration of airborne laser data with sub-decimeter accuracy. *ISPRS J Photogramm Remote Sens* 65(2):208–217, doi:10.1016/j.isprsjprs.2009.12.003
- Steinbacher F, Aufleger M (2013) Airborne hydromapping: a new approach to manage hydraulics, morphology and ecology. *Int J Hydropower Dams* 20(5):74–77
- Wagner W, Ullrich A, Ducic V, Melzer T, Studnicka N (2006) Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS J Photogramm Remote Sens* 60(2):100–112
- Wong H, Antoniou A (1991) Characterization and decomposition of waveforms for larsen 500 airborne system. *IEEE Trans Geosci Remote Sens* 29(6):912–921
- Ye D, Salamin Y, Huangfu J, Qiao S, Zheng G, Ran L (2015) Observation of Wave Packet Distortion during a Negative-Group-Velocity Transmission. *Sci Rep* 5:8100, doi:10.1038/srep08100
- Zhu J, Zhang Z, Hu X, Li Z (2012) Analysis and Application of Lidar Waveform Data Using a Progressive Waveform Decomposition Method. *Int Arch Photogramm Remote Sens Spat Inf Sci*, vol XXXVIII-5/W12, pp 31–36