Wa-LiD: A new LiDAR simulator for waters
H. Abdallah, N. Baghdadi, Jean-Stéphane Bailly, Y. Pastol, F. Fabre

To cite this version:

HAL Id: hal-00736576
https://hal.archives-ouvertes.fr/hal-00736576
Submitted on 28 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Wa-LiD: A new LiDAR simulator for waters

Hani Abdallah¹, Nicolas Baghdadi¹, Jean-Stéphane Bailly², Yves Pastol³, Frédéric Fabre⁴

¹ CEMAGREF, UMR TETIS, 500 rue François Breton, 34093 Montpellier Cedex 5, France.
² AgroParisTech, UMR TETIS, 500 rue François Breton, 34093 Montpellier Cedex 5, France.
³ SHOM, CS 92803, 29228 Brest Cedex 2, France
⁴ EADS – Astrium, 31 rue des cosmonautes, 31402 Toulouse Cedex 4, France

A simulator (Wa-LiD) was developed to simulate the reflection of LiDAR waveforms from water across visible wavelengths. The specific features of the simulator include (i) a geometrical representation of the water surface properties, (ii) the use of laws of radiative transfer in water adjusted for wavelength and the water’s physical properties, and (iii) modelling of detection noise and signal level due to solar radiation. A set of simulated waveforms was compared with observed LiDAR waveforms acquired by the HawkEye airborne and GLAS satellite systems in the near-infra red or green wavelengths and across inland or coastal waters. Signal-to-noise ratio (SNR) distributions for the water surface and bottom waveform peaks are compared with simulated and observed waveforms. For both systems (GLAS and HawkEye), Wa-LiD simulated SNR conform to the observed SNR distributions.

Index Terms— Waveform model, altimetry, bathymetry, SNR, GLAS, ICESat, HawkEye

I. INTRODUCTION

Water surface altimetry and bathymetry are key variables for many applications in hydrology, geomorphology and meteorology (e.g., [1], [2]). Several techniques for measuring water altimetry and bathymetry have been developed through the use of sonar, radar and optical imaging. Each of these techniques has drawbacks when used on inland or coastal waters [3], including (i) limited spatial coverage and use on navigable rivers (sonar), (ii) a large radar footprint when used for inland water, especially rivers and (iii) the lack of the use of optical imaging in unclear waters.

Airborne hydrographic LiDAR has proved to be a suitable sensor because of its accuracy and high spatial density features [4]. The potential for airborne hydrographic LiDAR is that it can be used for both altimetry and bathymetry because of the ability of LiDAR detectors to register returned signals from (i) the water surface for altimetry [5], (ii) the water bottom for bathymetry ([6], [7]), and (iii) the water column that allows some optical properties of water to be deduced [8]. The Airborne Hydrographic LiDAR technique is limited in terms of spatial coverage and because the data is expensive to gather.

ICESat has been the only altimeter LiDAR satellite up to now. The Geoscience Laser Altimeter System (GLAS) from ICESat provides accurate high resolution elevations for altimetry. The GLAS is promising for inland areas with widths greater than 1.5 km, even though the footprint is small (50 to 90 m), because of the time needed for detector gain adaptation when transitioning from land to water (0.25 s, approximately 1.5 km). The transition phase from land to water limits the use of GLAS in small, inland water areas [9].

To improve the prospective and the performance analysis in satellite LiDAR systems designed for water altimetry and bathymetry, a LiDAR signal database representative of the possible physical conditions of water using different LiDAR configurations is needed. The common procedure to obtain a database involves the development of a physical model simulating the signals. Such a model would allow the reproduction of the LiDAR signal regardless of the water and instrumental parameters.

Most hydrographic applications that use LiDAR signals are derived from the work of Guenther [5]. Before now, few studies have been published that model LiDAR waveforms on water; these studies include (i) the simulation model of Mclean [10], which generates LiDAR returned waveforms from a wind-roughened ocean, (ii) the work of Feigels [6] on the optimisation of airborne LiDAR’s parameters through modelling and analysing waveforms and (iii) work on modelling sea surface return and laser bathymetry from airborne LiDAR [7].

The objective of this study is to develop a new waveform simulation model available for any laser wavelength in the optical spectrum domain from 300 (Ultra Violet, UV) to 1500 nm (Near Infra Red, NIR). This model is similar to other Hydrographic LiDAR models (e.g., [5], [6], [7]) except it integrates radiative transfer laws governed by the physical properties of water for any wavelength. This simulation model uses a geometrical representation of the water surface with the geometric model of Cook and Torrance [11] and integrates both the characteristics of detection noise and signal level due to solar radiation.

The aim of this paper is (i) to present the water LiDAR waveform model (Wa-LiD) formulae regardless of LiDAR system parameters and water parameters, (ii) to exhibit a selection of simulated waveforms for certain wavelengths and the same water characteristics and (iii) to compare simulated waveforms to observed LiDAR waveforms from the satellite GLAS sensor and the airborne HawkEye sensor.
II. LiDAR MODELLING

A. The Water LiDAR (Wa-LiD) Simulation Model

1) Model description

The Wa-LiD model is able to simulate waveforms for any laser wavelength in the optical spectrum domain from 300 to 1500 nm regardless of the optical properties of water. These optical properties are determined by the attenuation (absorption + scattering) of optical light by pure water, yellow substances, phytoplankton and sediment (e.g., [12], [13]). In addition, this Wa-LiD model (i) takes into account any Bragg-scattering from the water surface using the geometric (facets) model of Cook and Torrance [11] and (ii) integrates the effects of detector noise and noise due to solar radiation.

The waveforms received by the LiDAR system (power as a function of time) are written as the sum of the echoes of multiple waves [7]:

\[ P(t) = P_{\text{o}}(t) + P_{\text{d}}(t) + P_{\text{b}}(t) + P_{\text{n}}(t) \]

Where \( P_{\text{o}}(t) \) is the total power received, \( P_{\text{d}}(t) \) is the power returned by the water surface, \( P_{\text{b}}(t) \) is the power returned by the water column, \( P_{\text{n}}(t) \) is the detector noise power and \( t \) is the time scale.

The emitted laser pulse from the LiDAR system is considered to be a Gaussian distribution [14]:

\[ w(t) = \frac{2}{T_{0}} \sqrt{\frac{\ln 2}{\pi}} \exp \left( -4 \ln 2 \frac{t^2}{T_{0}^2} \right) \]  

Where \( t \) is the two-way time delay of the emitted Gaussian pulse between the detector and the water surface \( t_{s} \), water column \( t_{c} \) or bottom \( t_{b} \). The returned waveforms from the water surface, water column and bottom correspond to the convolution of the Gaussian emitted pulse and the instant echoes from the water surface \( P_{s} \), water column \( P_{c} \) and bottom \( P_{b} \).

**Returned waveform from water surface:**

The returned waveform from the water surface received by the LiDAR detector \( P_{s}(t) \) is given by:

\[ P_{s}(t) = w(t) * P_{s} \]

\[ P_{s} = \frac{P T_{\text{2, atm}}^{2} A_{R} \eta_{R} \eta_{L} L_{S} \cos^{2}(\theta)}{\pi H^{2}} \]  

Where \( P_{s} \) is the emitted power by laser = \( E_{0} / T_{0} \), \( E_{0} \) is the emitted energy of the LiDAR system, \( T_{0} \) is the Gaussian FWHM (Full Width at Half Maximum), \( T_{\text{2, atm}} \) is the two way atmospheric loss, \( A_{R} \) is the receiver area, \( \eta_{R} \) and \( \eta_{L} \) are the optical efficiencies of emission and reception, \( L_{S} \) is the loss of transmission through the surface (surface albedo), \( \theta \) is the local incidence angle, \( H \) is the sensor altitude, \( t_{s} = \frac{2H}{c \cos(\theta)} \), and \( c \) is the velocity of light in air.

The loss factor through the surface \( L_{S} \) is calculated by the BRDF (Bidirectional Reflectance Distribution Function) of the water surface represented by specular microfacets using the geometric model of Cook and Torrance:

\[ L_{S} = \frac{k_{d} + k_{s} \text{DOF}}{\pi \cos^{2} \theta} \]  

Where \( k_{d} \) and \( k_{s} \) (\( k_{d} = 1 - k_{s} \)) determine the fraction of diffuse and specular components, respectively. \( D \) is the microfacet distribution function that models the surface as geometric facets and is described by the distribution function of Beckmann [15].

\[ D = \frac{1}{r^{2}} \cos\theta e^{-\frac{(\tan \theta/r)^{2}}{2}} \]  

\( r \) is the \( \text{rms} \) (root mean square) slope of the facets, which determines the surface roughness. Small values of \( r \) signify gentle facet slopes and give a distribution that is highly directional around the specular component while large values of \( r \) imply steep facet slopes and give a distribution that is spread out [11].

\( O \) is the geometric attenuation factor of the BRDF that takes into account the phenomenon of masking between facets (auto-shadowing). \( F \) is the function that describes the Fresnel reflection of light on each microfacet.

**Returned waveform from water column**

The returned waveform from the water column \( P_{c}(t) \) at a depth \( z \) is given by:

\[ P_{c}(t) = w(T \text{ )} \) * \( P_{c}(z) \) \]

\[ P_{c}(z) = \exp \left( \frac{\beta(z)}{\cos \theta} \right) \]  

\[ \theta_{w} \] is the local incidence angle in the water, \( F \) is a loss factor due to the field of view of the telescope, \( n_{w} \) is the water refractive index, \( \beta(\phi) \) is the volume scattering function, \( z \) is the column layer, \( t_{c} = t_{s} + 2z / c_{w} \) with \( c_{w} \) as the velocity of light in water, and \( k \) is the diffuse attenuation coefficient and is defined by the degradation rate of light with depth. An approximate formula was built by Guenther [5] to present the relationship between \( k \) and the optical properties of water:

\[ k = c(\lambda) \frac{0.19(1 - a(\lambda))}{a(\lambda)} \]  

Where \( a(\lambda) = b(\lambda) / c(\lambda) \) is the single scattering albedo, \( b(\lambda) \) is the scattering coefficient, \( c(\lambda) = a(\lambda) + b(\lambda) \) is the attenuation coefficient and \( a(\lambda) \) is the absorption coefficient.

The optical properties (\( a(\lambda), b(\lambda) \)) of turbid water are then expressed as the sum of contribution from different substances in water, such as yellow substances, phytoplankton and sediment [13], where:

\[ a(\lambda) = a_{w}(\lambda) + a_{b}(\lambda) + a_{ph}(\lambda) + a_{s}(\lambda) \]  

\[ b(\lambda) = b_{w}(\lambda) + b_{b}(\lambda) + b_{ph}(\lambda) + b_{s}(\lambda) \]

Where \( w, y, ph, S \) are the indices of pure water, yellow substances, phytoplankton and sediments, respectively.

**Returned waveform from water bottom**

The returned waveform from the water bottom received by the LiDAR detector \( P_{b}(t) \) is defined as:

\[ P_{b}(t) = w(t_{b}) * P_{b} \]

\[ P_{b} = \exp \left( \frac{\beta(z) / \cos \theta}{\pi \left( n_{w} + Z / c_{w} \right)^{2}} \right) \]  

Where \( \beta(z) \) is the angular distribution of the BRDF.
θ = \frac{R + 2t\sigma}{c_w\cos\theta_w}

Signal level due to solar radiation

The signal level due to solar radiation is defined as a Gaussian white noise with a mean of zero and a standard deviation equal to 1 convolved by the instant echo, \( P_{bg} \):

\[
P_{bg} = I_S A_R e^{-\gamma_r^2} \left( \frac{\pi \theta^2}{4} \right) A_\eta R
\]

Where \( I_s \) is the solar radiance reflected from the water column, \( A_\eta \) is the bandwidth of the optical filter in the receiver and \( \gamma_r \) is the receiver obscuration ratio.

Receiver internal noise

The detector internal noise is defined as a normal distribution with a zero mean and a standard deviation \( \sigma_N(t) \) that varies according to the signal level:

\[
\sigma_N(t) = \sqrt{2eB(P_{\text{out}}(t)G + I_d)}
\]

Where \( P_{\text{out}}(t) = P_{bg}(t) + P_\lambda(t) + P_\eta(t) + P_\phi(t) \), \( e \) is the electron charge (1.6 x 10⁻¹⁹ C), \( B \) is electrical bandwidth of the detector, \( G \) is the excess noise factor and \( R_1 \) is the responsivity.

Wa-LiD formulae presented above are programmed using a MATLAB code.

2) Wa-LiD simulation examples

Figure 1 shows simulated waveforms in the case of a satellite LiDAR sensor at an altitude of 500 km, emitting NIR (1064 nm), green (532 nm), and UV (350 nm) wavelengths. Water characteristics used in this simulation correspond to river water conditions with an average turbidity (average concentration of yellow substances \( a_\beta = 0.86 \) m⁻¹), phytoplankton \( C_{ph} = 4 \) mg/m³ and sediment \( S = 2.8 \) mg/l) and a 2 m depth. The NIR signal has the highest amplitude from the surface and penetrates a few cm into the water column. The green signal penetrates deeper into the water column and, depending on the water properties, is reflected from the bottom. The UV wavelength penetrates considerably through the water column and can reach the bottom, but the response is very low.

III. COMPARISON OF SIMULATED TO OBSERVED WAVEFORMS

Simulated waveforms are compared with observed LiDAR waveforms to identify the Wa-LiD model behaviour. The available observed data were provided from the satellite GLAS sensor and the airborne HawkEye sensor (© SHOM).

A. Observed waveforms Datasets

Two datasets were available. The first dataset contains waveforms acquired by the Geosciences Laser Altimeter System (GLAS) across the Lake Leman located in Switzerland and France (Lat. 46°26’N and Long. 6°33’E, Figure 2a). Lake Leman is one of the largest lakes in Western Europe with a surface area of 582 km². The maximum length and width of the lake are 73 km and 14 km, respectively. The relatively flat surfaces and the fact that the lake is not affected by tides make this a good location for validation studies. The GLAS data used in this study are GLA01. The GLA01 data include waveforms echoed from the water surface in NIR (1064 nm). Each waveform was sampled in 544 or 1000 bins of received power in volts at a sampling rate of 1 ns over land area [16]. In this study, two GLAS transects (set of successive shots or footprints) were selected. Waveforms from these two transects do not have saturation problems and are longer than other transects. The first transect was taken on March 4th 2005 with a length of 11.5 km and corresponds to 69 shots; the second transect was taken on June 7th 2006 with a length of 10.5 km and a corresponding 61 shots.

The second dataset contains waveforms registered by the airborne HawkEye system across a coastal area in the northern part of Reunion Island.

Fig. 2a. (a) GLAS transects over Leman Lake, (b) HawkEye points across a coastal area in the northern part of Reunion Island.

B. Simulated waveforms datasets

Table 2 shows the range of water parameters used for these simulations. A range of probable values of water parameters is
used for the simulation because the exact water properties were not precisely known at the time of the survey; in some cases, it was impossible to measure some of the properties, such as water surface roughness. The range of roughness, $r$, values used are from Beckman and Spizzichino [15]. The range of sediment concentration for coastal areas varies from 2.6 for clear water to 200 mg/l for very turbid water [18]. Here, we used a smaller range (2.6–10 mg/l) corresponding to clear water. In the same way, as sand in this area is black and formed as a result of volcanic activity [19], the selected range of values for bottom reflectance (Table 2) is the same used by Jun et al. [20].

### Table 1

<table>
<thead>
<tr>
<th>Parameters</th>
<th>GLAS</th>
<th>HawkEye</th>
<th>Parameters</th>
<th>GLAS</th>
<th>HawkEye</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$ (nm)</td>
<td>1064</td>
<td>532-1064</td>
<td>$\eta_r$</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>$H$</td>
<td>600 km</td>
<td>200 m</td>
<td>$\eta_b$</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$E_0$ (mJ)</td>
<td>75</td>
<td>3</td>
<td>$\gamma_r$</td>
<td>0.1</td>
<td>0.35</td>
</tr>
<tr>
<td>$T_0$ (ns)</td>
<td>5</td>
<td>7</td>
<td>$\Delta_{\eta_r}$ (nm)</td>
<td>1.2</td>
<td>1</td>
</tr>
<tr>
<td>$\theta$ ($^\circ$)</td>
<td>0</td>
<td>20</td>
<td>$\gamma_b$</td>
<td>0.3</td>
<td>0.35</td>
</tr>
<tr>
<td>$A_0$ (m$^2$)</td>
<td>0.8</td>
<td>0.025</td>
<td>B (MHz)</td>
<td>100</td>
<td>142</td>
</tr>
<tr>
<td>FOV (mrad)</td>
<td>5</td>
<td>30</td>
<td>G</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$R_c$ (A/W)</td>
<td>0.25</td>
<td>0.3</td>
<td>$\lambda_d$ (\AA)</td>
<td>$10^{-10}$</td>
<td>$10^4$</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Range</th>
<th>Parameters</th>
<th>GLAS</th>
<th>HawkEye</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>[0.1–0.5]</td>
<td>$T_{\text{atmos}}^2$</td>
<td>0.64</td>
<td>0.9</td>
</tr>
<tr>
<td>$S$ (mg/l)</td>
<td>[2.6–10]</td>
<td>$\beta (\mu \text{m}^2 \text{sr}^{-1})$</td>
<td>0.0014</td>
<td>0.0014</td>
</tr>
<tr>
<td>$R_c$</td>
<td>[0.05–0.2]</td>
<td>$L_{\text{w}}(\text{m}^2 \text{sr} \text{mm})$</td>
<td>0.025</td>
<td>0.025</td>
</tr>
<tr>
<td>$k$</td>
<td>0.9</td>
<td>$\rho_{\text{w}}$</td>
<td>1.33</td>
<td>1.33</td>
</tr>
</tbody>
</table>

C. Waveforms comparison

The amplitude of GLAS waveforms is in volts whereas the units for HawkEye are not known because of industrial trade secrets. For the Wa-LiD simulations, the waveforms are expressed in watts. For these reasons as well as the invariant shapes of waveforms and the fact that peaks in waveforms are the main information retrieved, the chosen method to compare simulated and observed waveforms was to compare the signal-to-noise ratio (SNR). Here, SNR is defined by the ratio between the waveform peak amplitude (water surface, bottom) and the standard deviation of detection noise (includes also the photon noise of the useful signal). The distributions of SNR for simulated and observed waveforms were computed. The 95%, 90%, 75% and 50% confidence intervals were thus calculated from SNR quantities for both observed waveforms (GLAS and HawkEye) and simulated ones. Next, the probabilities for an observed SNR to belong to a given simulated SNR confidence interval were computed. The number of simulated waveforms chosen was around two times higher than the number of observed waveforms to ensure a better simulation of extremes.

D. Results and discussion

1) Comparison with GLAS waveforms

Waveforms were simulated from Wa-LiD using the GLAS system parameters (Table 1) and the range of values for roughness $r$ (Table 2). In the infrared, the waveforms are influenced mainly by the water surface roughness because the penetration of this wavelength is very low in the water column and the LiDAR returns primarily from the water surface. In all, 260 waveforms were generated by random, and water surface properties were uniformly selected within the ranges given in Table 2. From the two GLAS transects selected, 130 GLAS waveforms across Leman Lake were used. Next, GLAS and simulated SNR distributions were computed and assigned confidence intervals. The probability of a GLAS SNR (SNR$_{\text{GLAS}}$) value belonging to the simulated SNR (SNR$_{\text{valid}}$) confidence intervals is higher (Figure 3a) (approximately 0.99 at a 95% confidence interval) than the opposite (Figure 3b) (approximately 0.57 at a 95% confidence interval). In the Wa-LiD model, the water surface albedo ($L_a$) is influenced by two water surface properties: the specular reflection coefficient and the surface roughness. The underestimation of the simulated SNR is probably due to the distribution of unknown water properties compared with simulations from distributions of uniform water properties.

2) Comparison with HawkEye waveforms

NIR channel

In all, 400 waveforms were simulated from Wa-LiD using the HawkEye system characteristics (Table 1) and a wide range of values for water surface roughness (Table 2). They were compared to 200 observed HawkEye waveforms. The probability of HawkEye SNR (SNR$_{\text{Hawk}}$) belonging to SNR$_{\text{valid}}$ confidence intervals are higher (approximately 0.98 at a 95% confidence interval) than the opposite (approximately 0.66 at a 95% confidence interval) (Figure 3). The weak underestimation of simulated SNR may result for the same reason as the underestimation for GLAS (i.e., unknown values of properties for actual surface water distribution), which may demonstrate the limitations of this comparison exercise. This result confirms the capability of the Wa-LiD model to simulate realistic LiDAR waveforms in the NIR.

Green channel

The simulated waveforms in the green were carried out using the same system parameters as those for the infrared, except for the wavelength value (532 nm). The water parameters (Table 2) were randomly selected as inputs for the simulator. The water surface roughness $r$ has a higher impact on the surface echo, and the sediment concentration $S$ has the higher impact on the water column echo for coastal areas [18].

![Figure 3](image-url)
In contrast, the bottom reflectance \( R_b \) has the higher impact on the bottom echo. The remaining water parameters were fixed for values that belong to clear coastal water (\( a_\phi=0.1 \text{ m}^{-1}, C_{\text{ph}}=1 \text{ mg/m}^3 \)). The comparison between HawkEye and simulated waveforms were carried out for 4 water depths: 3, 6, 10, and 15 m. For each depth, 400 simulations were generated. For comparison, 94 HawkEye waveforms were used: 21 at 3 m, 20 at 6 m, 20 at 10 m, and 18 at 15 m depths. HawkEye and simulated data were compared using the SNR for both the surface and bottom returns. The same confidence intervals were calculated for each considered bathymetry. Figure 4 presents the boxplots of SNR for surface and bottom returns from simulated and HawkEye green waveforms at a 3 m depth. There is a high probability (higher than 0.8) that the HawkEye surface SNR (SNR\(_{\text{Hawk}}\)) belong to the simulated surface SNR 95% confidence interval for all water depths (Figure 3a). These probabilities were higher than those of a simulated surface SNR to belong to the HawkEye surface SNR confidence intervals. At the bottom, the probability of HawkEye SNR (SNR\(_{\text{Hawk}}\)) to belong to the simulated bottom SNR confidence intervals were slightly higher than the opposite (Figure 3). This weak underestimation of simulated SNR (surface and bottom) is thought to be a result of the unknown actual distribution of 95% at the surface (\( k, r \)), the column (S) and the bottom (\( R_b \)), which does not fully match the distribution of uniform water properties for the simulation.

In future works, experiments and research are necessary for a proper characterisation of water surface roughness. Moreover, an analysis of the model sensitivity to different media characteristics and sensor parameters will be carried out to analyse the influential parameters and to assess the performance of future LiDAR systems for the water surface altimetry and bathymetry especially for what concerns the optimisation of the ground processing algorithms.

**ACKNOWLEDGMENT**

The authors would like to thank the NSIDC (National Snow and Ice Data Centre) and the SHOM (Service Hydrographique et Oceanographique de la Marine) for providing GLAS and HawkEye data respectively. The authors are very grateful to EADS (European Aeronautic Defence and Space Company) and CNES (Centre National d’Etudes Spatiales), which both supported this study.

**REFERENCES**


