Terrain surfaces and 3D landcover classification from small footprint full-waveform lidar data: Application to badlands

Frédéric Bretar, Adrien Chauve, Jean-Stéphane Bailly, Clément Mallet, Andres Jacome

Manuscript prepared for Hydrol. Earth Syst. Sci.
with version 2.1 of the LaTeX class copernicus.cls.
Date: 3 October 2008

Terrain surfaces and 3D landcover classification from small footprint full-waveform lidar data: Application to badlands

Frédéric BRETAR[1], **Adrien CHAUVE**[1,2], **Jean-Stéphane BAILLY**[2], **Clément MALLET**[1], **and Andres JACOME**[2]

[1]Institut Géographique National - Laboratoire MATIS - 4 Av. Pasteur 94165 Saint-Mandé - FRANCE
[2]Maison de la Télédétection - UMR TETIS AgroParisTech/CEMAGREF/CIRAD - 500 rue J.F Breton, 34095 Montpellier - FRANCE

Abstract. This article presents the use of new remote sensing data acquired from airborne full-waveform lidar systems. They are active sensors which record altimeter profiles. This paper introduces a set of methodologies for processing these data. These techniques are then applied to a particular landscape, the badlands, but the methodologies are designed to be applied to any other landscape. Indeed, the knowledge of an accurate topography and a landcover classification is a prior knowledge for any hydrological and erosion model. Badlands tend to be the most significant areas of erosion in the world with the highest erosion rate values. Monitoring and predicting erosion within badland mountainous catchments is highly strategic due to the arising downstream consequences and the need for natural hazard mitigation engineering. Additionaly, beyond the altimeter information, full-waveform lidar data are processed to extract intensity and width of echoes. They are related to the target reflectance and geometry. Wa will investigate the relevancy of using lidar-derived Digital Terrain Models (DTMs) and to investigate the potentiality of the intensity and width information for 3D landcover classification. Considering the novelty and the complexity of such data, they are presented in details as well as guidelines to process them. DTMs are then validated with field measurements. The morphological validation of DTMs is then performed via the computation of hydrological indexes and photo-interpretation. Finally, a 3D landcover classification is performed using a Support Vector Machine classifier. The introduction of an ortho-rectified optical image in the classification process as well as full-waveform lidar data for hydrological purposes is then discussed.

Correspondence to: Frédéric BRETAR
frederic.bretar@ign.fr

1 Introduction

Remote sensing is an effective set of techniques to collect physical data from the Earth surface used as inputs in erosion/hydrological models or to monitor hydrological fields over large areas (Schultz and Engman, 2000; King et al., 2005). Images obtained in the visible domain with optical sensors can be analyzed for generating 2D landcover and landform maps either automatically by image processing methods (Chowdhury et al., 2007) or by photo-interpretation. The use of the infrared channel helps to detect the vegetation (Lillesand and Kiefer, 1994). In a stereoscopic configuration, images are processed to generate Digital Surface Models (DSMs) (Kasser and Egels, 2002).

More recently, airborne lidar (LIght Detection And Ranging) systems (ALS) provide 3D point clouds of the topography by direct time measurement of a short laser pulse after reflection on the Earth surface. Moreover, such active systems, called **multiple echo lidar**, allow to detect several return signals for a single laser shot. It is particularly relevant in case of vegetation areas since a single lidar survey allows to acquire not only the canopy top (the only visible layer from passive sensors), but also points inside the vegetation layer and on the ground underneath. Depending on the vegetation density, some of them are likely to belong to the terrain. After a classification step in ground/off-ground points (also called filtering process), relevant Digital Terrain Models (DTMs) can be generated. Such DTMs are of high interest for geomorphologists to study erosion processes (McKean and Roering, 2004) or to map particular landforms (James et al., 2007). Moreover, hydrologic models such as TOPOG (O'Loughlin, 1986) or TOPMODEL (Quinn et al., 1991) handle topographic data either as Digital Surface/Terrain Models or as meshes. By classifying terrain under different levels and types of vegetation cover, lidar data, if suitable, could provide new land classification, i.e., terrain, cover maps. This new 3D landcover classification can even

be more related to the hydrological processes that are usually modelled in hydrological production indices as the SCS runoff curve number (USD, 1986), the runoff coefficient in the rational method (Pilgrim, 1987) or the the plant cover factor in Wischmeier and Smith's Empirical Soil Loss Model (USLE) (Wischmeier et al., 1978). Lidar data have also been investigated by Bailly et al. (2008) and Murphy et al. (2008) for drainage networks characterization (Cobby et al., 2001; Antonarakis et al., 2008) and by Mason et al. (2003) as input data for flood prediction problems. For the latter, the authors use lidar data as resampled elevation grids and detect high and low vegetation areas. Vegetation heights are then converted into friction coefficients.

Finally, **multiple echo lidar** data are typically used for the unique possibility of extracting terrain points as well as vegetation heights with high accuracy. When suitable, Hollaus et al. (2005) insist on the possibility to derive the roughness of the ground from lidar point clouds. However, the filtering algorithm used to process lidar data is landscape dependent and the classification result may be altered (Sithole and Vosselman, 2004).

Based on the same technology than multiple echo lidar systems, **full-waveform lidar** systems provide altimeter profiles of the reflected pulse. They represent the laser backscattered energy as a function of time. These profiles are processed to extract 3D points (echoes) but in addition, other interesting features that could be related to landscape characteristics. Depending on the landscape properties (geometry, reflectance) and on the laser diffraction angle (entailing small or large footprint), the recorded waveform becomes of complex shape. An analytical modelling of the profiles provides the 3D position of significant targets as well as the intensity and the width of lidar echoes (Section 3.1). A detailed state-of-the-art of such systems can be found in Mallet and Bretar (2008).

This paper introduces a set of methodologies for processing full-waveform lidar data. These techniques are then applied to a particular landscape, the badlands, but the methodologies are designed to be applied to any other landscape.

Indeed, badlands tend to be among the most significant areas of erosion in the world, mainly in semi-arid areas and in subhumid Mediterranean mountainous areas (Torri and Rodolfi, 2000). For the latter case, more active dynamics of erosion are observed (Regues and Gallart, 2004) with the highest erosion rate values in the world (Walling, 1988). Very high concentrations of sediment during floods, up to 1000 $g.l^{-1}$, were registered (Descroix and Mathys, 2003).

Badlands are actually defined as intensely dissected natural and steeply landscapes where vegetation is sparse (Bryan and Yair, 1982). Badlands are characterized by V-shape gullies that are highly susceptible to weathering and erosion (Antoine et al., 1995). These landscapes result from unconsolidated sediments or poorly consolidated bedrock, as marls, under various climatic conditions governing bedrock disintegration through chemical, thermal or rainfall effects (Nadal-Romero et al., 2007).

The hydrological consequences of erosion processes on this type of landscapes are a major issue for economics, industry and environment: high solid transport, bringing heavily loaded downstream flood, are silting up reservoirs (Cravero and Guichon, 1989) and downstream river aquatic habitats (Edwards, 1969). Therefore, monitoring and predicting erosion within badland mountainous catchments is highly strategic due to the arising downstream consequences and the need for natural hazard mitigation engineering (Mathys et al., 2003). Traditionaly, the monitoring activities in catchments are derived from heavy in situ equipments on outlets or from isolated and ponctual observations within catchments. In complement to these traditional observations, hydrologists are expecting remote sensing to help them to upscale and/or downscale erosion processes and measurements in other catchments, by providing precise and continuous spatial observations of erosion features or erosion driven factors (Puech, 2000). Among other inputs, erosion monitoring and modelling approaches on badlands (Mathys et al., 2003) need maps of landform features, mainly gullies (James et al., 2007) that are driving the way flows and maps of important driven factors of erosion in mountainous badland catchments. These factors are soil and rocks characteristics (Malet et al., 2003), vegetation strata used to derive 3D landcover classes controlling rainfall erosivity, and the terrain topography (Zhang et al., 1996), which allow to derive slope and aspect of marly hillslopes (Mathys et al., 2003).

This paper aims at investigating the potentialities of using full-waveform lidar data as relevant altimeter data, but also as a possible data source for 3D landcover classification focusing on the characterization of badland erosion features and terrain classification. If some papers have been published regarding the interpretation of full-waveform lidar data, most of them are based on large footprint lidar data acquired from satellite plateforms (Zwally et al., 2002). Very few researches have been carried out on the analysis of small footprint full-waveform airborne lidar data (Mallet and Bretar, 2008). Considering their novelty and their complexity, we propose to develop some new and specific guidelines related to their processing (including some physical corrections) and their management. Furthermore, we will show that DTMs generated from lidar data are of high accuracy, even over complex mountainous landscapes, which is consistent with the study of erosion processes. Finally, the extraction of the intensity and the width of each echo is investigated as potential information for landcover classification. Intensity and width are related to the target reflectance as well as to the local geometry (slope, 3D distribution of the target).

This paper begins with a background on full-waveform lidar systems (Section 2.1) as well as a brief presentation of

a management system to handle the data (Section 2.2). We then present the processes to convert raw data into 3D point clouds (Section 3.1). Section 3.2 is dedicated to the development of a filtering algorithm to classify the lidar point cloud into ground/off-ground points as well as on the generation of DTMs. The echo intensity and width extracted from full-waveform lidar data are described in Section 3.3. We focus this section on theoretical developments (detailed in Appendix A), basis of the introduction of intensity corrections. Section 4 presents the badland area whereon investigations have been performed as well as the data: lidar data, orthoimages and field measurements. DTMs produced by our algorithm are then validated by both field measurements (Section 5.1) and by the computation of an hydrological index (Section 5.2) compared with manually (photo-interpretation) extracted crests and thalwegs. We finally present in Section 6 the results of a 3D landcover classification using a first level of terrain vegetation cover classes and based on a supervised classifier: the Support Vector Machines (SVMs). Different features have been tested, including the three visible channels of the orthoimage. The opportunity of using full-waveform lidar data for hydrological purposes is then discussed.

2 Managing full-waveform lidar data

2.1 Background on full-waveform lidar systems

The physical principle of ALS consists in the emission of short laser pulses, with a width of 5-10 ns at Full-Width-at-Half-Maximum (FWHM), from an airborne platform with a high temporal repetition rate of up to 200 kHz in multiple echo mode. They provide a high point density and an accurate altimeter description within each laser diffraction beam. The two way runtime to the Earth surface and back to the sensor is measured. Then, the range from the lidar system to the illuminated surface is recorded (Baltsavias, 1999). A lidar survey is composed of several parallel and overlapping strips (100 m to 1000 m width).

The emitted electromagnetic wave interacts with objects depending on its wavelength. The main influences on the laser light come from artificial or natural objects belonging to the illuminated surface. For ALS systems, near infra-red sensors are used (typical wavelengths from 0.8 to 1.55 μm). The Pulse Repetition Frequency (PRF) depends on the acquisition mode and on the flying altitude. Contrary to multiple echo systems which record only some high energy peaks in real time, full-waveform lidar systems record the entire signal of the backscattered laser pulse. Figure 1 shows raw full-waveform data.

Full-waveform systems sample the received waveform of the backscattered pulse at a frequency of 1 GHz. The footprint size depends on the beam divergence and on the flight altitude. Most commercial airborne systems are **small footprint** (typically 0.3 to 1 m diameter at 1000 m altitude).

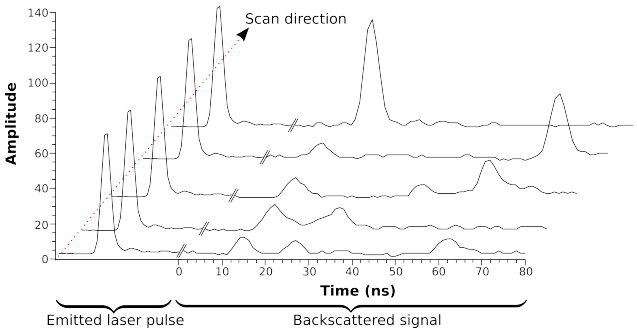


Fig. 1. Raw full-waveform lidar data: five emitted pulses and their respective backscattered signals.

2.2 Handling full-waveform lidar data

Initially, raw full-waveform lidar data are sets of range profiles of various lengths. Raw profiles are acquired and stored in the sensor geometry following both the scan angle of the lidar system and a chronological order along the flight track. After the georeferencing process and the pre-processing step (Section 3.1), raw profiles become vectors of attributes containing, for each 3D point, the x,y,z-coordinates, additional parameters (intensity and FWHM) and a link to the sensor geometry. Managing these data is much more complex than images: the topology (neigbhorhood system, topological queries) is designed to be as efficient as possible when accessing and storing the data. Indeed, the data volume is drastically larger than traditional laserscanning techniques: it takes 140 GB for an acquisition time of 1.6 h with a PRF of 50 kHz. Moreover, a 3D/2D visualization tool is also necessary to handle the attributes, both in the sensor and in the ortho-rectified geometries (cf. figure 2). A specific software has therefore been developed for these purposes.

3 Processing full-waveform lidar data

3.1 From 1D signals 3D to point clouds

Contrary to multiple echo lidar sensors which provide directly 3D point clouds, full-waveform sensors acquire 1D depth profiles along the line of sight for each laser shot. The derivation of 3D points from these signals is composed of two steps:

- The waveform processing step provides the signal maxima location, i.e., the range values, as well as additional parameters describing the echo shape.
- The georeferencing process turns the range value to a $\{x, y, z\}$ triplet within a given geographic datum.

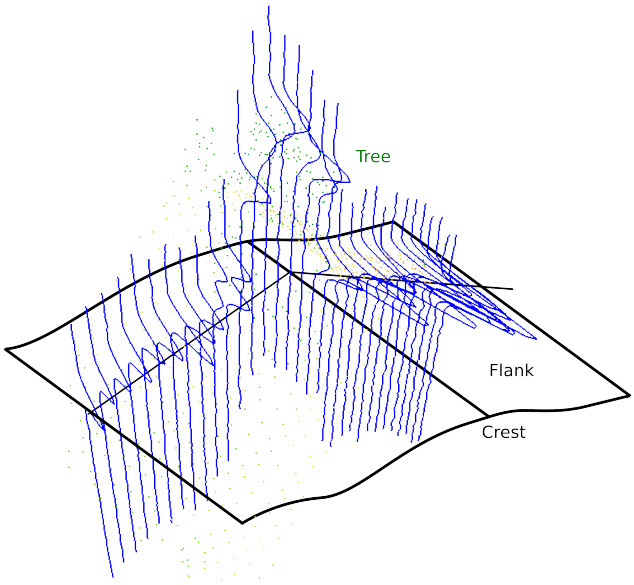


Fig. 2. 3D representation of georeferenced waveforms on a crest. A tree is lying on the crest.

WAVEFORM PROCESSING: It aims at maximizing the detection rate of relevant peaks within the signal in order to foster information extraction. In the literature, a parametric approach is generally chosen to fit the waveform. Parameters of a mathematical model are estimated. The objective is twofold. A parametric decomposition gives the signal maxima, i.e., the range values of the different targets hit by the laser beam. Then, the best fit to the waveform is chosen among a class of functions. This allows to introduce new parameters for each echo and to extract additional information about the target shape and its reflectance.
Our methodology is based on a paper written by Chauve et al. (2007). The authors describe an iterative waveform processing using a Non-Linear Least Squares fitting algorithm. After an initial coarse peak detection, missing peaks are found in the residuals of the difference between the modelled and initial signals. If new peaks are detected, the fit is performed again. This process is repeated until no further improvement is possible. This enhanced peak detection method is useful to model complex waveforms with overlapping echoes and also to extract weak echoes.
The Gaussian function has been shown to be suitable to model echoes within the waveforms (Wagner et al., 2006). Its analytical expression is:

$$f_{\mathrm{G}}(x) = I \exp\left(-\frac{|x-\mu|^2}{2\sigma^2}\right) \qquad (1)$$

where μ is the maximum location, I the peak amplitude, and σ the peak width.
For each recorded waveform, the transmitted pulse is also digitized. By retrieving its maximum location, the time interval between the pulse emission and its impact on a target is known. The range value of the target ensues from the time-of-flight calculation.
In this paper, the echo amplitude will be refered to as *intensity*. However, in the literature, the intensity can also be associated to the total energy of the echo, product of the intensity and σ. The standard deviation σ corresponds to the half width of the peak at about 60% of the full height. In some applications, however, the Full-Width-at-Half-Maximum (FWHM) is often used instead. We have FWHM$= 2\sigma\sqrt{2\,\ln 2}$.

GEOREFERENCING: Similarly to multiple echo lidar sensors, computing the $\{$x, y, z$\}$ coordinates of each echo in a geodetic reference frame from the range value requires additional data. The scan angle is used jointly to the range to calculate the $\{$x, y, z$\}$ position for each point in the scanner coordinate frame. Then, the GPS position of the aircraft, and the sensor attitude values (roll, pitch, heading) for each laser shot are recovered from the full-waveform data file to calculate the $\{$x, y, z$\}$ in a given geodetic datum. Finally, the positions can be transformed in some cartographic projection (French NTF Lambert II Etendu in this paper, see Section 4 for more details).
The transformation formulas cannot be expressed because they differ from a sensor to another. Offset values are different, depending on the configuration of the laser system, GPS and Inertial Measurement Unit (IMU) devices.

After applying the advanced step of waveform modelling, full-waveform lidar data generate denser point clouds than multiple echo data. It is particularly relevant when studying the vegetation structure (Mallet and Bretar, 2008). However, we will see in Section 6 that additional parameters are also of interest for landcover or soil classification.

3.2 From point clouds to DTM

The processing of a lidar point cloud consists in classifying the data, which is generally associated to the resampling of the data on a regular grid. Due to the very fine geometry of a lidar point cloud, many algorithms have been developed to automatically separate gound points from off-ground points (Sithole and Vosselman, 2004). Most of these approaches have good results when the topography is regular, but remain unperfect in case of mixed landscapes and slope conditions: parameters of the algorithms are often difficult to tune and do not fit over a large area. When ground points are mis-classified as off-ground points, the accuracy of the DTM may decrease (it depends on the spatial resolution and on the interpolation method). Inversely, when off-ground points (vegetation or man-made objects) are considered as ground points, the DTM becomes spiky which can be misinterpreted by hydrological models. Vegetated landscapes

with sparse vegetation in a mountainous areas (alpine landscapes) are particularly interesting for the study of natural hydrology and the phenomenoms of erosion (cf. Section 4). Nevertheless, the processing of such landscapes need strong human interactions to correct the classification: typical errors are mis-classification of vegetation points in steep slopes, over-estimation of the DTM in case of dense vegetated areas, under-estimation of the DTM in open slope areas when the local slope is not explicitly introduced in the process.

A methodology that handles these problems has been recently developed (Bretar and Chehata, 2008) and is used in this study to compute the DTMs. It is based on a two step process:

i. The computation of an initial surface using a predictive Kalman filter: it aims at providing a robust surface containing low spatial frequencies of the terrain (main slopes). The algorithm consists in analyzing the altimeter distribution of the point cloud of a local area in the local slope frame. Points of the first altimeter mode (lowest points) belong to the terrain. A DTM value at a specific position depends on the neighboring pixels through their respective uncertainties. The predictive Kalman framework provides not only a robust terrain surface (the slopes are also integrated in the predictive filter), but also an uncertainty σ_{DTM} for each DTM pixel as well as a map of normal vectors $\overrightarrow{n}$.

ii. The refinement of this surface using a Markovian regularization: it aims at integrating micro relieves (lidar points within the uncertainty σ_{DTM}) in a minimization process to refine the terrain description. Formulated in a Bayesian framework, additional prior information (crest, thalweg etc.) can also be integrated in the refinement process.

The lidar point cloud is then classified based on geometric criteria. A lidar point is labelled as GROUND if it is located within a buffer zone defined as the corresponding DTM uncertainty σ_{DTM}. Otherwise, it is considered as OFF-GROUND. In natural landscapes, off-ground points belong mainly to vegetation, and sometimes to human-made features (e.g., electric power lines, shelters). Vegetation areas are described as non-ordered point cloud (high variance) compared to human-made structures. Vegetation points are therefore extracted by fitting a plane on the off-ground points. If the residuals are higher than a defined threshold ($\sim 0.3\,\mathrm{m}$), points are labelled as VEGETATION. Figure 3 summarizes the entire algorithm to calculate a DTM from a lidar point cloud. This classification is not explicitly used in the following, but for generating the validation set related to the supervised classification (Section 6). Figure 4 is a 3D bird view of the orthoimage superimposed on the DTM.

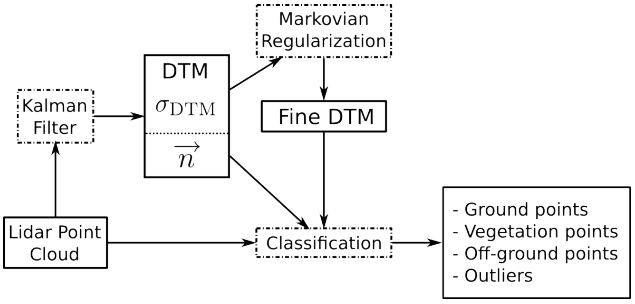


Fig. 3. Flowchart of the geometrical processes of a lidar point cloud.

Fig. 4. 3D bird view of the orthoimage superimposed on the DTM. ©IGN

3.3 Processing intensity and width of lidar echoes

Beyond the 3D point cloud, full-waveform lidar data provide intensity and width of each echo (Section 3.1) that are potential interesting features for landcover classification. The backscattered intensity (or received power) is a function of the laser power, the distance source-target, the incidence angle, the target reflectivity, the absorption by the atmosphere etc. The use of such features in a landscape classification framework necessitates a global coherence between all strips. Correcting the recorded intensity values from some of known contributions makes possible the analysis of "physical" parameters such as the target reflectivity. We propose also to analyze the effect of the incidence angle on the Full-Width-at-Half-Maximum.

3.3.1 The intensity

According to Nicodemus et al. (1977), the scattered radiant flux P_{s} in the zenith/azimuth angles $(\theta_{\mathrm{s}}, \phi_{\mathrm{s}})$ within the cone Ω_{s} is related to the incident flux P_{i} in the direction $(\theta_{\mathrm{i}}, \phi_{\mathrm{i}})$ within Ω_{i} by (Figure 5)

$$P_{\mathrm{s}}(\theta_{\mathrm{s}}, \phi_{\mathrm{s}}, \Omega_{\mathrm{s}}) = \varrho(\Omega_{\mathrm{i}}, \Omega_{\mathrm{s}}) P_{\mathrm{i}}(\theta_{\mathrm{i}}, \phi_{\mathrm{i}}, \Omega_{\mathrm{i}}) \qquad (2)$$

where $\varrho(\Omega_i, \Omega_s)$ is the biconical reflectance.

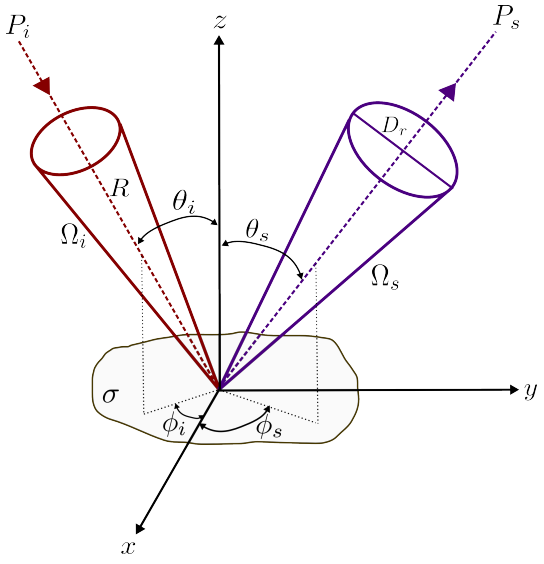


Fig. 5. Coordinate system of the scattered and incident radiant flux.

Introducing the backscattered cross section of the target σ, equation 2 can be rewritten as (Wagner et al., 2008b):

$$P_\mathrm{s} = \frac{D_\mathrm{r}^2}{4\pi R^4 \beta_\mathrm{t}^2} \sigma P_\mathrm{i} \tag{3}$$

where D_r is the diameter of the receptor, R the range from sensor to target, β_t the laser beam width.
With $\sigma = \pi \rho_\mathrm{m} R^2 \beta_\mathrm{t}^2 \cos\theta_\mathrm{s}$ (Höfle and Pfeifer, 2007)

$$P_\mathrm{s} = \frac{D_\mathrm{r}^2 \rho_\mathrm{m}}{4R^2} \cos\theta_\mathrm{s} P_\mathrm{i} \tag{4}$$

where ρ_m is the target reflectance, which depends on the material.

Since the recorded intensity is proportional to the backscatterred flux P_s, correcting intensity values gives access to the target reflectance, and therefore, in case of a Lambertian surface, to the classification of the material. Since the apparent reflecting surface is smaller in case of non-zero incidence angle than in case of zenithal measurements (the cosinus dependency in equation 4), recorded intensity values are corrected from the scalar product of the emitted laser direction and the corresponding terrain local slope extracted from the DTM.

We have also remarked that emitted pulses have significant amplitude variations along the flight track which may alter the spatial homogeneity of returned waveforms. Figure 6(a) represents the ratio between the intensity values of the emitted laser pulse and the average intensity values over the whole strip along the flight track (x-axis). Considering the high PRF of the laser, intensity values are constant along the scan line. The effects of such variations are visible in the returned waveforms as vertical lines (figure 6(b)). We therefore normalized the returned waveforms by the average intensity value of all emitted pulses. The effects of the correction are presented in figure 6(c). One can notice that vertical lines have disappeared.

3.3.2 The Full-Width-at-Half-Maximum

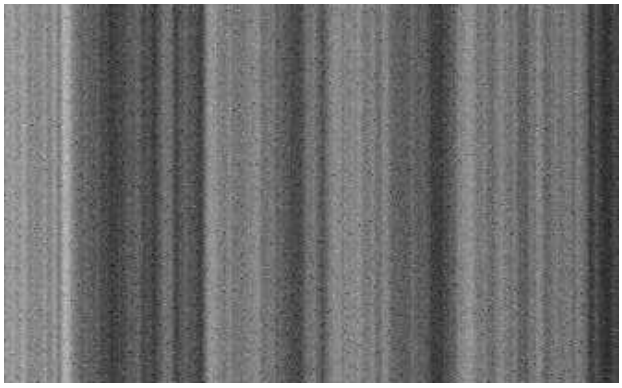
The FWHM has shown some spatial variability in our data set. Considering the badland and alpine landscape, we investigated the influence of the incidence angle on the FWHM only in case of **bare soil areas**. Indeed, the FWHM of under-vegetation ground points may have been modified by the complex optical medium. These investigations have been performed on simulated waveforms reflected by a tilted planar surface. Appendix A gathers the entire developements as well as some discussions. We show that, in our acquisition configuration (low divergence angle, low flight altitude), the FWHM should stay constant with various incidence angles. We cannot extend this conclusion for ground points below the vegetation since the waveform has been modified through the canopy cover. The spatial variability is therefore attributed to a more complex spatial beam response of the surface due to structures and/or reflectance properties.

4 Materials

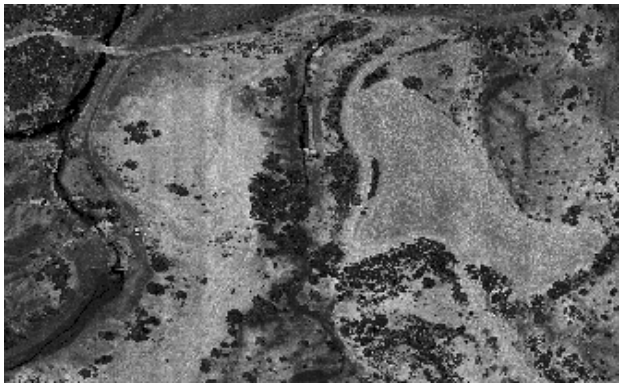
Lidar data have been acquired over the Draix area, France. Draix area is an experimental area on erosion processes in badlands located in the South of the French Alps. It belongs to the Euromediterranean Network of Experimental and Representative Basins (ERB). The Draix area consists in five research experimental catchments, highly equipped and monitored for more than thirty years. Thirteen research units working on erosion and hydrology processes are grouped within the GIS Draix organization (Mathys, 2004). Results for the most two eroded catchments are presented here: they concern the Laval and the Moulin catchments.

4.1 Lidar data

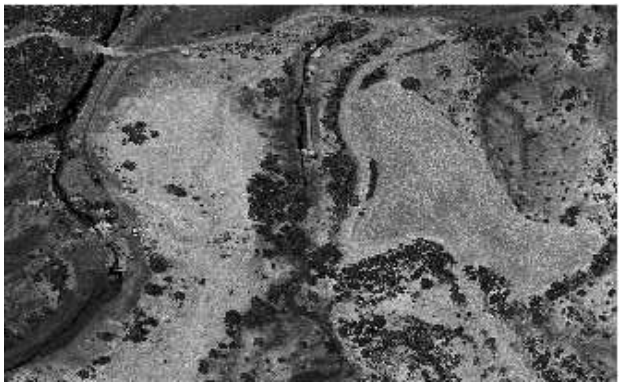
The data acquisition was performed in April 2007 by Sintégra (Meylan, France) using a RIEGL© LMS-Q560 system. This sensor is a small footprint airborne laser scanner and its main technical characteristics are presented in Wagner et al. (2006). The lidar system operated at a PRF of 111 kHz. The flight height was approximatively 600 m leading to a footprint size of about 0.25 m. The point density was about 5 pts/m^2.
The temporal sampling of the system is 1 ns. Each return waveform is made of one or two sequences of 80 samples. For each profile, a record of the emitted laser pulse is also provided (40 samples).



(a) Ratio between the intensity value of the emitted laser pulse and the average intensity values over the whole strip. Values are represented in grey level scale and stretched between 0.72 and 1.35



(b) Raw return intensity of the first echo. Values are represented in grey level scale and stretched between 0 and 150.



(c) Corrected return intensity of the first echo from the laser fluctuations. Values are represented in grey level scale and stretched between 0 and 150.

Fig. 6. Effect of the correction from the laser fluctuations. Images are presented in the sensor geometry.

For this study, three overlapping strips have been used with perpendicular direction. For each of them, a sub-part corresponding to the Moulin and the Laval catchement have been extracted. Strip footprints are presented in figure 8 and denoted S6 (blue), S7 (red), S8 (green).

(a) Orthoimage acquired during the lidar survey (RGB_{RAW}).

(b) Orthoimage extracted from the IGN BDOrtho© (RGB_{IGN}).

Fig. 7. Two orthoimages RGB_{IGN} and RGB_{IGN} over the Draix area.

4.2 Orthoimages

Two orthoimages were available for the study. The first one is extracted from the French IGN data basis BDOrtho©. Acquired in fairly good conditions (almost no shadowed zones) by the IGN digital camera, a physical-based radiometric equalization process has been applied (Paparoditis et al., 2006). The ground resolution is $0.5\,\mathrm{m}$. The triplet of {red, green, blue} channels of the IGN image will be referred in this article to as RGB_{IGN}. The second orthoimage has been calculated from aerial images acquired during the lidar survey by an embedded digital camera. Since the survey has been performed early in the morning, numerous shadowed areas appear. Moreover, no radiometric equalization has been performed entailing a rather poor radiometric quality (see figure 13). The ground resolution is $0.2\,\mathrm{m}$. The triplet of {red, green, blue} channels of this image will be referred in this article to as RGB_{RAW}.

4.3 Field measurements

Quality control points (or ground truth data) were surveyed by a mixed campaign DGPS in Real Time Kinematics (RTK), and a Total Station (ellipsoid: CLARKE 1880 IGN, geoid

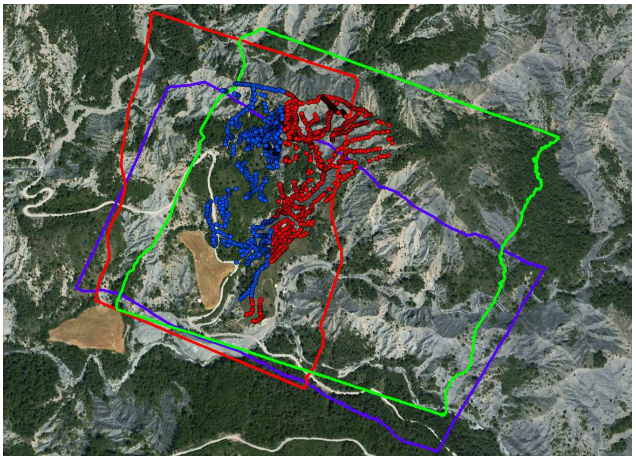


Fig. 8. Lidar strips superimposed on the orthoimage. The blue (resp. red and green) strip is denoted S6 (resp. S7 and S8). Field measurements are also plotted: blue points represent the West subset, red points the East subset.

model RAF98, coordinate reference system NTF Lambert II Etendu). The accuracy was 0.025 m in planimetry, and 0.03 m in altimeter. These points are chosen mainly on thalwegs (bottom of gullies) and crests. Some gullies are so deep that the GPS system (Leica 1200) was not able to work at the same accuracy. Those points were then surveyed with the help of a Total Station (Sokkia SET630R). 3D crossed validation between DGPS and Total Station surveys performed in six points were never greater than one centimeter. A total of 2886 quality control points have been used (cf. figure 8).

5 DTMs analysis

5.1 Qualification of DTMs with field measurements

We present in table 12 a comparison between field measurements and three DTMs generated from each single strip (S6, S7 and S8). One can observe that there is a bias in each DTM w.r.t field measurements as well as a significant standard deviation and RMS.

Looking carefully throught the lidar data and the statistics, a strip adjustment problem was diagnosticated. In order to validate the DTMs, the West part of the field measurements has been used to adjust the DTM while the East part to validate the terrain surface (denoted West↬East in table 11). We also took the dual configuration (East↬West) to test the relevancy of the proposed correction. Here, the adjustement consists in finding the best 3D translation T_{opt} that minimizes a RMS between field measurements and their projection on the DTM. We used a brute force method to explore the entire parameter space. A x, y (resp. z) search step of 0.4 m (resp. 0.1 m) was chosen in relation to the planimetric (resp. altimeter) accuracy of lidar points. The symmetrical validation gives some hints on the real deformations of the DTM, which are generally much more complex than a 3D translation. Table 11 gathers the results of both the optimal correction applied to each DTM and the mean and standard deviation of respectively West↬East and East↬West configurations.

Table 11 shows that the adjustement improves the final accuracy of DTMs both by decreasing the bias and the standard deviation. However, one can notice that the optimal 3D translation varies depending on the West↬East and East↬West configuration. A 0.4 m difference in the y-direction of T_{opt} for DTMs S7 and S6 doubles the standard deviation, while a 0.8 m difference in the y-direction of T_{opt} for DTM S8 has no effect on the final accuracy. These observations tend to show that the deformations between DTMs is not purely a 3D translation, but is of different nature such as polynomial (surface tilt) or non linear (rotation).
As a conclusion, we can say that, after a simple geometric adjustement, the calculated DTMs at 1 m resolution have an absolute altimeter accuracy of some decimeters. A better geometric adjustement should improve this accuracy.

5.2 Qualification of DTMs with hydrological indexes and photo-interpretation

The quality assessment of a DTM for hydrological purposes is not completely satisfying when considering only the altimeter error distribution. Other DTM quality criteria directly connected to the usual hydrological information extracted from DTM may be used: drainage networks, drainage areas, slopes like presented in (Charleux-Demargne, 2001). These criteria are mainly based on the basic landform information related to the first and the second derivative of a DTM. However, these criteria are not easy to use in a qualification process since (1) they are conditioned by both the algorithms and the parameters used to produce the information (e.g., a drainage area threshold in the D8 flow accumulation algorithm), (2) reference data are not easily available (how to survey drainage networks?) and finally (3) the quantification of quality is often not properly defined (how to compare dissimilarities of drainage networks?). Moreover, criteria are usually not generic: it is related to a specific hydrological index.
In order to overcome these problems, a single criteria is proposed for a quantified auto-evaluation of DTMs at a given resolution in erosion areas with an hydological and morphological point of view.
This criteria is the rate of crests and thalwegs observed from an orthoimage that are detected from the convergence index (CI) built on a DTM (Köthe and Lehmeier, 1994). The convergence index corresponds, for each DTM cell, to the mean difference between angle deviations. These angle deviations are calculated in each of the eight adjacent pixels. For an adjacent pixel, the angle deviation is the absolute difference, in degrees, modulo 180, between its aspect and the azimuth to

the central pixel (Zevenbergen and Thorne, 1987). The convergence index is a symetric and continuous index ranging from -90° up to 90°. This index highlights crests when highly positive and thalwegs when highly negative. Figure 9 shows the convergence indexes computed on S6. Main thalwegs and crests appear with respectively highly negative (blue) and positive (red) values.

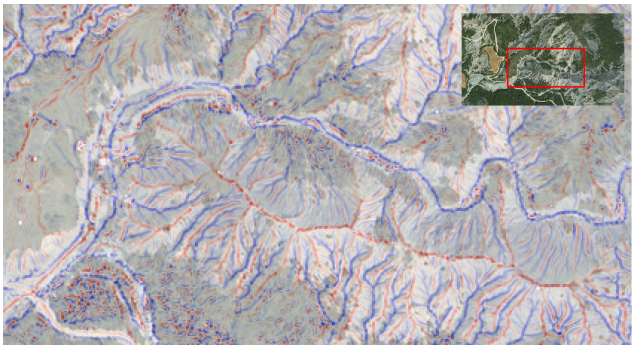


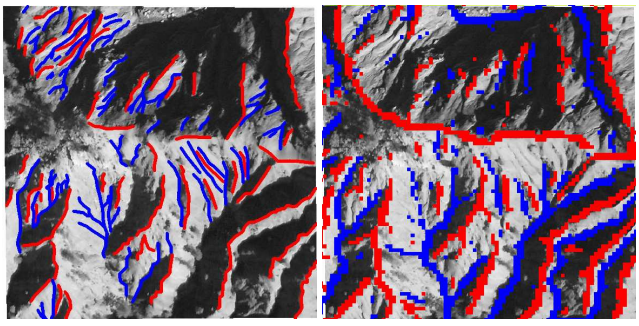
Fig. 9. CI computed on S6 superimposed to the orthoimage.

At a given location, a thalweg (resp. a crest) is considered to be detected in the DTM) if CI values belong to $[-90°, -\eta]$ (resp. to $[\eta, 90°]$, $\eta \in \mathbb{R}$). On a "perfect" DTM without noise, only CI=0 (i.e., η=0) indicates a plane terrain without any crests and thalwegs, whatever the slope is. When dealing with noisy DTM, thresholding the CI with η to retrieve significant crests and thalwegs becomes a challenging task. We therefore simulated a distribution of CI from a set of 1000 virtual noisy DTMs. They were generated with a trend corresponding to a plane of constant slope (e.g., 33° is the mean slope of Draix area). The simulation consists in generating Gaussian random fields (Lantuejoul, 2002) using the LU method (Journel and Huijbregts, 1978) following noise spatial distribution models with parameters: range, nugget and sill (variance) for spatial covariance.

Since the simulated CI distribution is of Gaussian shape, we set η to two times the standard deviation. We accept that five percents of CI values due to hazard on noise can be classified in significant crest and thalweg.

We show some results on a sub-aera of Draix. The simulated CI distribution (performed on 33°slope, Gaussian noise of zero mean and 2.66 standard deviation) provides a threshold value $\eta = 8.46$. We show the results of the thalweg and crest detection on figure 10. Figure 10(a) is a manual delineation of apparent crests and thalwegs. The photo-interpretation process is applied on main structures, but very near linear elements and the ones near sporadic vegetated elements are not considered.
Table 13 presents the quality criteria for S6 and S8.

The values obtained are conditioned by the spatial resolution of the DTM and the relative georeferencing of the data



(a) (b)

Fig. 10. (a) Test area (85 * 85 m) with photo-interpreted thalwegs (blue) and crests (yellow). (b) Detection of significant crests (red) and thalwegs (blue).

(image and DTM). However, when comparing results for S6 and S8, results show that (1) S6 is better representing landforms than S8 and (2) crests are more precisely detected in DTMs than thalwegs.

5.3 Discussion

Regarding the altimeter quality of full-waveform LiDAR DTMs, we obtain a rather precise and accurate relief restitution of a catchment of several square kilometers (about the same as the one obtained with multi-echo LiDAR). The observation of local erosion processes requires a more detailed relief restitution. Other techniques like terrestrial LiDAR or photogrammetry by unmanned aerial vehicles (Jacome et al., 2008) are more accurate and precise, but, are not well adapted to survey large areas. However, considering the altimeter accuracy of DTMs (approximately 0.9 m for 2 standard deviation on the altimeter random error), and that the local ablation speed over Draix area is of 1.5 cm per year (Oostwoud and Ergenzinger, 1998), change detection and monitoring of erosion effects would require a delay between surveys of several decades. Nevertheless, the loss of sediment volume within catchments are not homogeneous and are temporary stored on hill-slope gully networks: $\sim$ 200 tons/km^2 are trapped in the gully network), which corresponds to an approximate of 150 m^3 (Mathys et al., 1996). These volumes are significant enought to shorter time lag for a multidate analysisfull-waveform LiDAR DTM (lower than a decade), even with an accuracy of some decimeters. Only full-waveform LiDAR survey which gives an adequate compromise between precision, accuracy and extent makes possible the monitoring of sediment volume displacement in the gully network at a catchments scale

6 3D landcover classification

6.1 Methodology

Lidar data have been used so far as accurate altimeter data to extract ground points and generate DTMs. The challenges were to automatically process the data in a mountainous landscape with steep slopes and vegetation, the whole with the highest accuracy. We mentionned in the introduction that a landcover map is an important input of hydrological models. We therefore propose in this section to describe the inputs and outputs of a classification framework wherein lidar width and intensity values can be integrated and their benefit evaluated. Indeed, the interpretation of additional lidar parameters has been barely studied and reveals to be of interest for landcover classification. Wagner et al. (2008a) proposed classification rules based on a decision tree for vegetation/non-vegetation areas in a urban landscape using solely the width and the amplitude: a point is considered as VEGETATION if (1) it is not the last pulse of a profile containing multiple returns (2) it is a single return with low amplitude ($\leq$75) and large width ($\geq$1.9 ns). Focusing on the study of the vegetation, Reitberger et al. (2008) have integrated different features to segment individual trees in a graph-cut framework. Among them, the authors show that the feature corresponding to the average intensity on the entire tree plays the most important role in leaf-on conditions, while the ratio between the number of single reflections and the number of multiple reflections is the most important in leaf-off conditions.

Here, we would like to answer the question: do lidar width and intensity values improve a classification pattern in badlands? An efficient supervised classification algorithm called Support Vector Machines (SVM) has been used (Chang and Lin, 2001). This machine learning algorithm is well adapted to deal with high-dimensional feature space since the algorithm complexity does not depend on the data dimension. In recent years, SVM was shown to be relevant technique for remote sensing data analysis (Huang et al., 2002): ability to mix data from different sources, robustness to dimensionality, good generalization ability and a non-linear decision function (contrary to decision trees for instance). In this paper, the 3D lidar point cloud is labelled, thus providing a 3D landcover classification. (Mallet et al., 2008) applied this technique with success for classifying urban areas from full-waveform lidar data.

Four classes have been identified focusing on a first hierarchical level of 3D land cover classification, relevant for badlands landscapes with anthropogenic elements: 1-LAND, 2-ROAD, 3-ROCK and 4-VEGETATION. The three first classes can be ordered on an increasing erosion sensitivity criteria. The first class LAND is taking into account terrain under natural vegetation cover and cultivated areas in grassland. The second one, ROADs, are linear elements with natural (marls), bared but compacted material. These elements are known to impact runoff production within catchments. The third one contains areas with bared black marls in gullies, the main source of sediment production. The latter, vegetation, could be used further to describe the 3D vegetation structure, useful for a more detailed hierarchical level of 3D land cover classification.
SVM algorithm requires its own feature vector for each 3D lidar point to be classified. Only three lidar features have been retained. Indeed, it appears that the larger the number of features, the more difficult to make an interpretation of the results. They are:

- d_{DTM}, the distance between the 3D point and the DTM,
- I, the echo intensity,
- FWHM, the echo width.

Additionally, the RGB_{IGN} and RGB_{RAW} features have been added in the classifier, providing three radiometric attributes (figure 12). Their introduction allows a discrimination between **road** and **land** impossible with the lidar features and improve the classification results. The radiometric quality of the orthoimage has been tested using both orthoimages.
The training set over each of the four classes has been manually defined by selecting a specific mask on the orthoimage (photo-interpretation). We have implemented the SVM algorithm with the LIBSVM software (Hsu and Lin, 2001), selecting the generic Gaussian kernel. For more theoretical explanations, please see (Pontil and Verri, 1997).

6.2 Results and discussion

The data set S6 has been analyzed. Figure 11 shows the histograms of lidar derived features corresponding to the four selected classes. d_{DTM} and the I have bounded values which describe the **vegetation** (resp. $>$ 1 m and between 0 and 20), whereas the width values tend to be uniform between 3 ns and 4.5 ns. **Road** and **land** have similar distributions for lidar derived features, which explains the high confusion values in Table 14. The distributions of **rocks** is flattened for the *distance to the DTM* since many points are choosen in very steep slopes, and are therefore more sensitive to the DTM quality. The intensity of **rocks** is slighty different from the other classes. Figures 12 and 13 shows the histograms of RGB_{IGN}.

The classification is validated with ground truth sets defined both manually on the orthoimage and also using the geometric classification assigned in Section 3.2. We assume that points below the vegetation truth set belong to the class LAND. $\sim 40\,\%$ of the total number of points have been validated. Figure 14 shows the four validation sets for each class. A confusion matrix is then calculated for each configuration. The accuracy of the classification results are quantified by the average accuracy AA, mean of the diagonal values of the

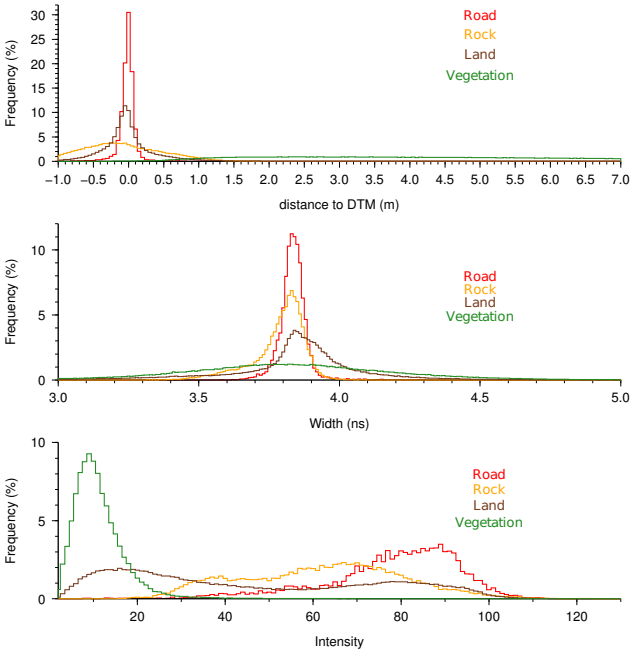


Fig. 11. Histograms of I, FWHM and d_{DTM} for the four classes ROAD, ROCK, LAND and VEGETATION.

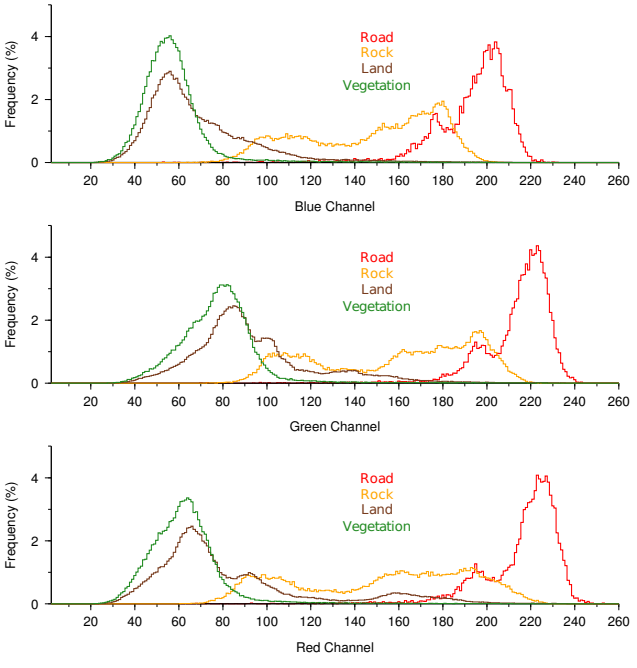


Fig. 12. Histograms of RGB_{IGN} for the four classes ROAD, ROCK, LAND and VEGETATION.

confusion matrix. AA does not depends on the number of points in each validation set.

When using solely lidar derived features $\{d_{DTM}$, I, FWHM$\}$, table 14 indicates that the confusion between classes is not negligible particularly for some of them: ROCK with ROAD reaches $22.4\,\%$, while LAND with ROAD reaches $19.4\,\%$, what was predictable looking through the statistics of the training set (Figure 11). The vegetation has a high percentage of true positive ($94.2\,\%$) and is well detected. With an average accuracy of $79.1\,\%$, it appears that a classification based only on lidar derived features is consistent.

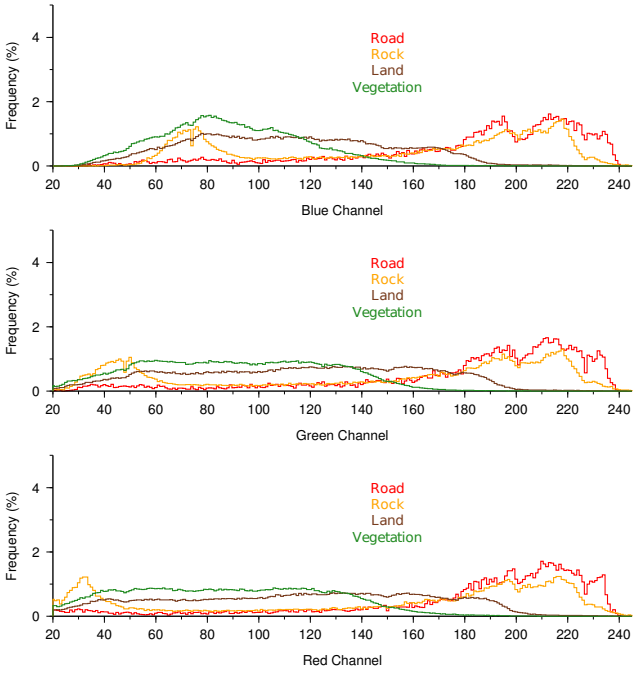


Fig. 13. Histograms of RGB_{RAW} for the four classes ROAD, ROCK, LAND and VEGETATION.

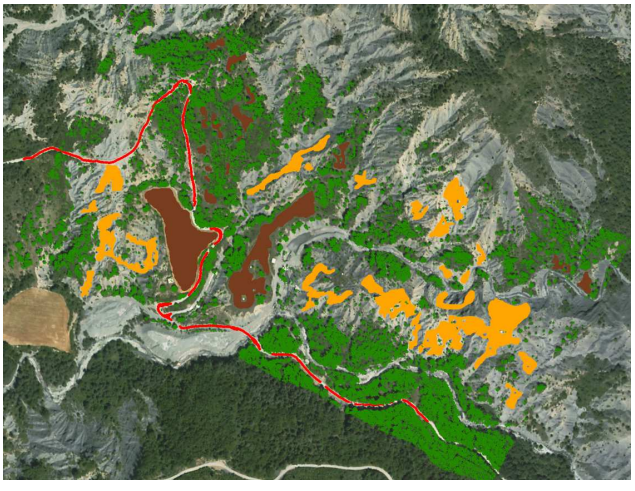


Fig. 14. Ground truth classification of S6 for each class LAND (dark brown), ROAD (red), ROCK (orange) and VEGETATION (dark green).

Before testing the effects of introducing lidar intensity and width, we investigated the impact of the radiometric quality of the orthoimages on the classification results. Tables 15 and 16 are the confusion matrices corresponding to the classification results with respectively $\{d_{DTM}, RGB_{RAW}\}$ and $\{d_{DTM}, RGB_{IGN}\}$. One can observe a significant discrepancy between both radiometric features with an average accuray of $82.1\,\%$ using $\{d_{DTM}, RGB_{RAW}\}$ and $92.2\,\%$ using $\{d_{DTM}, RGB_{IGN}\}$. The true positive values of ROCK (resp. ROAD) increase from $73\,\%$ (resp. $74\,\%$) to $89.3\,\%$ (resp. $93.5\,\%$) when using $\{d_{DTM}, RGB_{IGN}\}$ instead of $\{d_{DTM}, RGB_{RAW}\}$. Moreover, the confusion between several classes decreases significantly: ROAD with ROCK decreases from $10.4\,\%$ to $3.7\,\%$, ROAD with LAND from $13.8\,\%$ to $1.2\,\%$, ROCK with ROAD from $19.8\,\%$ to $8\,\%$, LAND with ROAD from $4.2\,\%$ to $1.5\,\%$. In other words, the use of $\{d_{DTM}, RGB_{IGN}\}$ instead of $\{d_{DTM}, RGB_{RAW}\}$ gives better classification results.

True positive values are higher when using image-based features $\{d_{DTM}, RGB_{IGN}\}$ than $\{d_{DTM}, I, FWHM\}$ and the confusion between classes most of the time decreases: LAND with ROAD decreases from $19.4\,\%$ to $1.5\,\%$, ROAD with ROAD decreases from $22.4\,\%$ to $8\,\%$. Nevertheless, the comparison is more mitigated with $\{d_{DTM}, RGB_{IGN}\}$. Indeed, true positive values of ROAD decrease from $84.2\,\%$ to $74\,\%$ and the confusion between the other classes increases significantly. However, LAND is better classified with less confusion with ROAD ($19.4\,\%$ to $4.2\,\%$). As a results, it appears that even if the average accuracy of a classification using image-based features is better, intensity and width of lidar echoes have interesting discriminative properties.

The results of the introduction of lidar intensity and width in the classification process are shown in tables 17 and 18. There are minor effects on the results when using $\{RGB_{IGN}, I, FWHM, d_{DTM}\}$ instead of $\{RGB_{IGN}, d_{DTM}\}$. The classification is bettered when using $\{RGB_{RAW}, I, FWHM, d_{DTM}\}$ instead of $\{RGB_{RAW}, d_{DTM}\}$, true positive values of ROCK increase from $73\,\%$ to $76.2\,\%$, ROAD increase from $74\,\%$ to $77.1\,\%$, VEGETATION are similar and LAND increase from $85.8\,\%$ to $87.1\,\%$. When comparing $\{RGB_{RAW}, I, FWHM, d_{DTM}\}$ with $\{I, FWHM, d_{DTM}\}$ (table 14), the improvement is particularly consistent for ROCK, LAND and VEGETATION, but true positive values of ROAD decreases from $84.2\,\%$ to $77.1\,\%$ and the confusion with ROCK increases from $7.5\,\%$ to $13.4\,\%$. In fact, the radiometry of roads are sensitive to tree shadows. The combination of the very high resolution of RGB_{RAW} and the time of the survey (early in the morning) feeds the training set with bright and dark (shadow) radiometric values. On the contrary, lidar intensity and width do not depend on the sun configuration. Superimposed on the orthoimage of figure 14, a 3D landcover classification obtained with $\{RGB_{IGN}, I, FWHM, d_{DTM}\}$ is presented in figure 15 and figure 16.

Finally, the quality of the classification depends mainly on the DTM accuracy (represented here as d_{DTM}). Moreover, within the framework of the methodology, it appears that a classification based on $\{I, FWHM, d_{DTM}\}$ is suitable, but gives a worse accuracy than a classification based on $\{d_{DTM}, RGB_{RAW}\}$ or $\{d_{DTM}, RGB_{IGN}\}$. Used on their own, full-waveform lidar data are relevant to discriminate vegetation from non vegetation points, but the confusion between other classes remains not negligible. The intensity and the width do not improve the classification accuracy if the radiometric features have a good separation between classes. Otherwise, the benefit is rather small, but in case of artefacts in a class (like whadow) for which lidar measurements are not sensitive. Inversely, the use of poor radiometric features may alter the classification result of specific landscapes (here ROAD) where intensity and width are well bounded.

7 Conclusions

The different points treated in this paper entail some conclusions. Firstly, the accuracy of the lidar data we worked on (badlands) was proven decimetric. Even if erosion dynamics on these landscapes would require a centimetric accuracy to be studied yearly, DTMs generated from lidar survey are consistent for hydrological sciences at the catchment level. We focused this paper on generating and qualifying DTMs, but also on the automatic computation of a 3D landcover classification. We showed that lidar intensity and width contain enough discriminative information on badlands to be classified in LAND, ROAD, ROCK and VEGETATION with $\sim 80\,\%$ accuracy. The introduction of image-based radiometric features in the classifier improved the accuracy of the classification ($\sim 92\,\%$), provided that they bring relevant discrimination between classes. Otherwise, it would be better to introduce alternative data.

Appendix A

Waveform simulation

This section aims at investigating the received waveform properties of a laser pulse after a reflection onto a tilted plane. The simulation is meant to represent one typical situation that occurs when surveying a vegetation-free area of various slopes. We focus this section on studying the influence of the incidence angle with regard to the FWHM. The simulation is based on the paper written by Kirchhof et al. (2008).

A received waveform $R(x, y, z, t)$ is often written as a temporal convolution product between the emitted laser pulse $s(t)$, the impulse response of the receiver $h(t)$, the spatial beam profile $P(x, y)$ and the illuminated area $S(x, y, z)$.

$$R(x, y, z, t) = s(t) * h(t) * P(x, y) * S(x, y, z) \qquad \text{(A1)}$$

where $*$ is the convolution operator.

Assuming $h(t)$ constant (ideal photodiode, infinite bandwidth and linear frequency characteristic of the receiver) and a uniform spatial distribution $P(x, y)$, equation A1 can be written as

$$R(x, y, z, t) = s(t) * S(x, y, z, t) \tag{A2}$$

The surface response can describe the geometry of the surface as well as reflectance properties. Assuming the surface to be Lambertian, we can write:

$$R(x, y, z, t) = \sum_{\tau=-\infty}^{\tau=\infty} s(\tau - t) S(x, y, z, \tau) \tag{A3}$$

For the simulations, waveforms are sampled at 10 GHzso that small variations of the FWHM should be detected.

- **Derivation of the emitted pulse**

The emitted pulse can be of different shapes. Kirchhof et al. (2008) use a uniform distribution since it fits with their system. Our data showed that the emitted pulse was most of the time of Gaussian shape. We therefore used a Gaussian distribution of a given standard deviation σ_0 to simulate the interactions with the tilted plane.

$$s(t) = \exp\left(-\frac{t^2}{\sigma_0^2}\right) \tag{A4}$$

- **Derivation of the surface response**

We consider the coordinate system described in figure A1.

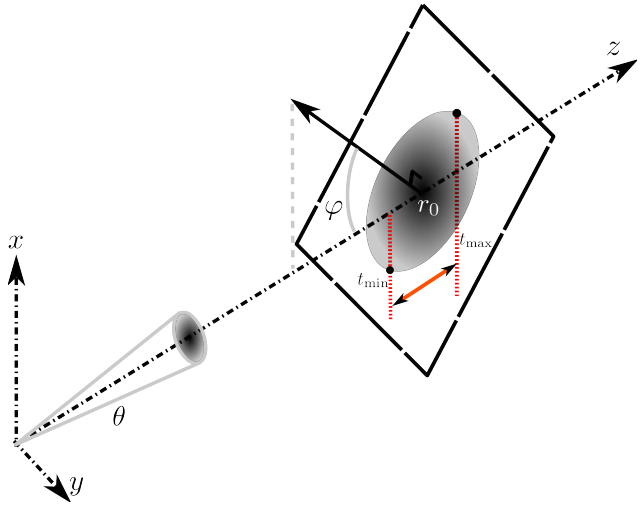


Fig. A1. Coordinate system of the simulation.

The divergence of the laser beam is denoted θ. The normal vector of the plane is oriented in the $(0, x, y)$-plane. The laser pulse and the receiver are located in $(0, 0, 0)$. S is therefore the intersection of the divergence cone and the plane:

$$S(x, y, z, t) = \begin{cases} x^2 + y^2 \leq \left(z \tan(\frac{\theta}{2})\right)^2 \\ x \sin\varphi + z \cos\varphi = r_0 \cos\varphi \end{cases} \tag{A5}$$

$$\Leftrightarrow S(z) = 2\sqrt{\left(z \tan(\frac{\theta}{2})\right)^2 - (r_0 - z)^2 \cot^2\varphi} \tag{A6}$$

$$\text{with } z \in [\frac{r_0}{1 + \tan(\frac{\theta}{2}) \tan\varphi}; \frac{r_0}{1 - \tan(\frac{\theta}{2}) \tan\varphi}]$$

When subsituting $z = \frac{ct}{2}$, (c is the speed of light) we obtain the temporal description of the surface response between t_{min} and t_{max}. From the receiver point of view, a contribution of the surface at a given time is an ellipse section (equation A6) along the y-axis between t and $t + dt$. Considering the very short integration time $\Delta t = t_{\text{max}} - t_{\text{min}}$ for realistic incidence angles φ and the small divergence angle θ, the surface response is symmetric w.r.t $\frac{2r_0}{c}$.

The integration time Δt is particularly of importance in the convolution product: the larger Δt, the flatter the convolved received waveform. Δt depends on the incidence angle φ, on the beam divergence θ and on the range r_0. Figure A2 shows the variations of the integration time w.r.t. the incidence angle φ for various divergence angles. For a given range $r_0 = 500$ m (which is a common flight altitude for a full-waveform lidar survey) and $\theta = 0.3$ mrd (which is the RIEGL©standard), one can observe that the integration time remains short for large incidence angles:

$$\left.\begin{array}{l} \max_{\Delta t<1\text{ns}}(\varphi) < 45° \\ \max_{\Delta t<2\text{ns}}(\varphi) < 65° \end{array}\right\} \theta = 0.3 \text{ mrd}$$

For larger divergence angles, the incidence angles have a strong effect on the integration time. We have

$$\left.\begin{array}{l} \max_{\Delta t<1\text{ns}}(\varphi) < 17° \\ \max_{\Delta t<2\text{ns}}(\varphi) < 31° \\ \max_{\Delta t<7\text{ns}}(\varphi) < 65° \end{array}\right\} \theta = 1 \text{ mrd}$$

$$\left.\begin{array}{l} \max_{\Delta t<1\text{ns}}(\varphi) < 9° \\ \max_{\Delta t<2\text{ns}}(\varphi) < 17° \\ \max_{\Delta t<14\text{ns}}(\varphi) < 65° \end{array}\right\} \theta = 2 \text{ mrd}$$

Figure A3 shows the variations of the integration time w.r.t. the range r_0 for a given incidence angle $\varphi = 30°$ and various divergence angles. We note that the integration time remains again short w.r.t. the range for small divergence angles. For small footprint lidar, seeing that an airborne full-waveform survey is performed at a flight altitude less than 1000 m, the effect is negligible on the convolution. It is no

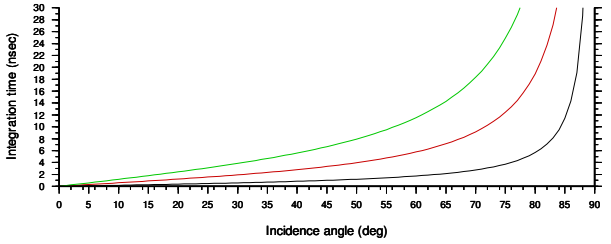


Fig. A2. Variations of the integration time $\Delta t = t_{\max} - t_{\min}$ with regard to $\varphi \in [0; 90° - \frac{\theta}{2}]$. $r_0 = 500$ m, $\theta = 0.3$ mrd (black), $\theta = 1$ mrd (red) and $\theta = 2$ mrd (green).

longer true when dealing with large footprint satellite data.

$$\left.\begin{array}{l}\max_{\Delta t<1\text{ns}}(r_0) < 868\text{ m} \\ \max_{\Delta t<2\text{ns}}(r_0) < 1733\text{ m}\end{array}\right\} \theta = 0.3\text{ mrd}$$

$$\left.\begin{array}{l}\max_{\Delta t<1\text{ns}}(r_0) < 260\text{ m} \\ \max_{\Delta t<2\text{ns}}(r_0) < 520\text{ m} \\ \max_{\Delta t<4\text{ns}}(r_0) < 1040\text{ m}\end{array}\right\} \theta = 1\text{ mrd}$$

$$\left.\begin{array}{l}\max_{\Delta t<2\text{ns}}(r_0) < 260\text{ m} \\ \max_{\Delta t<4\text{ns}}(r_0) < 520\text{ m}\end{array}\right\} \theta = 2\text{ mrd}$$

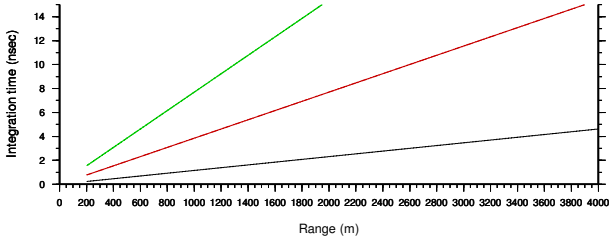


Fig. A3. Variations of the integration time $\Delta t = t_{\max} - t_{\min}$ with regard to $r_0 \in [200\ m; 4000\ m]$. $\theta = 0.3$ mrd (black), $\theta = 1$ mrd (red), $\theta = 2$ mrd (green) and $\varphi = 30°$.

• Variations of the FWHM w.r.t. the incidence/ divergence angle

We noticed previously that the integration time modifies the surface response. It is particularly true for large divergence/incidence angles as well as for large ranges. We would like to investigate the results of the convolution between the emitted laser pulse ($\sigma_0 = 2.1$ ns $\Leftrightarrow$ FWHM $= 5$ ns) and the surface response for various incidence/divergence angle and various ranges. Table 19 summarizes our experiments. **For a weak divergence angle** $\theta = 0.3$ **mrd**, the variations of the FWHM w.r.t. the incidence angle and to the range are negligible. **For a small incidence angle** $\varphi = 10°$, the effect of the divergence angle and of the range are negligible. **For an incidence angle of** $\varphi = 30°$, the variations of the FWHM increases with the range and the divergence angle, but remain negligible for operational flight conditions of a full-waveform airborne survey and a GigaHz sampler. **For a large incidence angle** $\varphi = 60°$, the variations become significant both with increasing the divergence angle and the range.

Acknowledgements. The authors would like to deeply thank the GIS Draix for providing the full-waveform lidar data and for helping in ground truth surveys. They are grateful to INSU for its support to GIS Draix through the ORE program.

References

Urban hydrology for small watersheds, Technical Release 55, United States Department of Agriculture, Natural Resources Conservation Service, Conservation Engineering Division, 2nd edn., 1986.

Antoine, P., Giraud, D., Meunier, M., and Ash, T. V.: Geological and geotechnical properties of the Terres Noires in southeastern France: weathering, erosion, solid transport and instability, Engineering Geology, 40, 223–234, 1995.

Antonarakis, A. S., Richards, K. S., Brasington, J., Bithell, M., and Muller, E.: Retrieval of vegetative fluid resistance terms for rigid stems using airborne lidar, Journal of Geophyscal Research, 113, 2008.

Bailly, J., Lagacherie, P., Millier, C., Puech, C., and Kosuth, P.: Agrarian landscapes linear features detection from LiDAR elevation profiles: application to artificial drainage network detection, International Journal of Remote Sensing, 29 (11-12), 3489–3508, 2008.

Baltsavias, E. P.: Airborne Laser Scanning: Basic relations and formulas, ISPRS Journal of Photogrammetry and Remote Sensing, 54 (2-3), 199–214, 1999.

Bretar, F. and Chehata, N.: Terrain Modelling from lidar range data in natural landscapes: a predictive and Bayesian framework, Tech. rep., Institut Géographique National, available at http://hal.archives-ouvertes.fr/hal-00325275/fr/, 2008.

Bryan, R. and Yair, A.: Perspectives on studies of badland geomorphology, in: Badland Geomorphology and Piping, edited by Bryan R, Y. A., GeoBooks (Geo Abstracts Ltd), pp. 1–12, Norwich, UK, 1982.

Chang, C.-C. and Lin, C.-J.: LIBSVM: a library for support vector machines, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2001.

Charleux-Demargne, J.: Qualité des modèles numériques de terrain pour l'hydrologie. Application à la caractérisation du régime de crues des bassins versants, Ph.D. thesis, Université de Marne-La-Vallée, France, 2001.

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M., and Puech, W.: Processing full-waveform lidar data: modeling raw signals, in: IAPRS, vol. 36 (Part 3/W52), Espoo, Finland, 2007.

Chowdhury, P. R., Deshmukh, B., and Goswami, A.: Machine Extraction of Landforms from Multispectral Images Using Texture and Neural Methods, in: Proc. of the International Conference on Computing: Theory and Applications, pp. 721–725, Washington, DC, USA, 2007.

Cobby, D. M., Mason, D. C., and Davenport, I. J.: Image processing of airborne scanning laser altimetry data for improved river

flood modelling, ISPRS Journal of Photogrammetry and Remote Sensing, 56 (2), 121–138, 2001.

Cravero, J. and Guichon, P.: Exploitation des retenues et transport des sdiments, La Houille Blanche, 3-4, 292–296, 1989.

Descroix, L. and Mathys, N.: Processes, spatio-temporal factors and measurements of current erosion in the French Southern Alps: a review, Earth Surface Processes and Landforms, 28 (9), 993–1011, 2003.

Edwards, D.: Some effects of siltation upon aquatic macrophyte vegetation in rivers, Hydrobiologia, 34 (1), 29–38, 1969.

Höfle, B. and Pfeifer, N.: Correction of laser scanning intensity data: Data and model-driven approaches, ISPRS Journal of Photogrammetry and Remote Sensing, 62 (6), 415–433, 2007.

Hollaus, M., Wagner, W., and Kraus, K.: Airborne laser scanning and usefulness for hydrological models, Advances in Geosciences, 5, 57–63, 2005.

Hsu, C.-W. and Lin, C.-J.: LIBSVM: a library for Support Vector Machine, software available at http://www.csie.ntu.edu.tw/∼cjlin/libsvm, 2001.

Huang, C., Davis, L., and Townshend, J.: An assessment of support vector machines for land cover classification, International Journal of Remote Sensing, 22 (4), 725–749, 2002.

Jacome, A., Puech, C., Raclot, D., Bailly, J., and ROUX, B.: Extraction d'un modèle numrique de terrain par photographies par drone, Revue des Nouvelles Technologies de l'Information, to appear, 2008.

James, L., Watson, D., and Hansen, W.: Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA, CATENA, 71 (1), 2007.

Journel, A. and Huijbregts, C.: Mining Geostatistics, Academic Press, London, 1978.

Kasser, M. and Egels, Y.: Digital Photogrammetry, Taylor & Francis, 2002.

King, C., Baghdadi, N., Lecomte, V., and Cerdan, O.: The application of Remote Sensing data to monitoring and modelling of soil erosion, CATENA, 62 (2-3), 79–93, 2005.

Kirchhof, M., Jutzi, B., and Stilla, U.: Iterative processing of laser scanning data by full waveform analysis, ISPRS Journal of Photogrammetry and Remote Sensing, 63 (1), 99–114, 2008.

Köthe, R. and Lehmeier, F.: SARA System - System zur Automatischen Relief-Analyse, 1994.

Lantuejoul, C.: Geostatistical Simulation: Models and Algorithms, Springer Verlag, Berlin, 2002.

Lillesand, T. and Kiefer, R.: Remote Sensing and Image interpretation, John Wiley & Sons, 1994.

Malet, J., Auzet, A., Maquaire, O., Ambroise, A., Descroix, L., Esteves, M., Vandervaere, J., and Truchet, E.: Investigating the influence of soil surface characteristics on infiltration on marly hillslopes, Earth Surface Processes and Landforms, 28 (5), 547–560, 2003.

Mallet, C. and Bretar, F.: Full-Waveform Topographic Lidar: State-of-the-Art, ISPRS Journal of Photogrammetry and Remote Sensing, to appear, 2008.

Mallet, C., Bretar, F., and Soergel, U.: Analysis of full-waveform lidar data for classification of urban areas, Photogrammetrie Fernerkundung GeoInformation (PFG), 5/2008, 337–349, interne, 2008.

Mason, D. C., Cobby, D. M., Horritt, M. S., and Bates, P. D.: Floodplain friction parameterization in two-dimensional river flood models using vegetation heights derived from airborne scanning laser altimetry, Hydrological Processes, 17 (9), 1711–1732, 2003.

Mathys, N.: Information available http://www.grenoble.cemagref.fr/etna/oreDraix/oreDraix.htm, 2004.

Mathys, N., Brochot, S., and Meunier, M.: L'érosion des Terres Noires dans les Alpes du Sud : contribution l'estimation des valeurs annuelles moyennes (bassins versants expérimentaux de Draix, Alpes-de-Haute-Provence, France), Revue de géographie alpine, pp. 17–27, 1996.

Mathys, N., Brochot, S., Meunier, M., and Richard, D.: Erosion quantification in the small marly experimental catchments of Draix (Alpes de Haute Provence, France). Calibration of the ETC rainfall-runoff-erosion model, CATENA, 50 (2-4), 527–548, 2003.

McKean, J. and Roering, J.: Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry, Geomorphology, 57 (3-4), 331–351, 2004.

Murphy, P., Meng, J. O. F., and Arp, P.: Stream network modelling using lidar and photogrammetric digital elevation models: a comparison and field verification, Hydrological Processes, 22 (12), 1747–1754, 2008.

Nadal-Romero, E., Regués, D., Marti-Bono, C., and Serrano-Muela, P.: Badland dynamics in the Central Pyrenees: temporal and spatial patterns of weathering processes, Earth Surface Processes and Landforms, 32 (6), 888–904, 2007.

Nicodemus, F., Richmond, J., Hsia, J., Ginsberg, I., and Limperis, T.: Geometrical Considerations and Nomenclature for Reflectance, Tech. rep., U.S. Department of Commerce, 1977.

O'Loughlin, E.: Prediction of surface saturation zones in natural catchments by topographic analysis, Water Resources Research, 22 (5), 794–804, 1986.

Oostwoud, W. D. and Ergenzinger, P.: Erosion and sediment transport on steep marly hillslopes, Draix, Haute-Provence, France: an experimental field study, CATENA, 33, 179–200(22), 1998.

Paparoditis, N., Souchon, J.-P., Martinoty, G., and Pierrot-Deseilligny, M.: High-end aerial digital cameras and their impact on the automation and quality of the production workflow, ISPRS Journal of Photogrammetry and Remote Sensing, 60, 400–412, 2006.

Pilgrim, D. H., ed.: Australian rainfall and runoff, Institution of Engineers, Canberra, Australia, 1987.

Pontil, M. and Verri, A.: Properties of support vector machines, Tech. Rep. AIM-1612, MIT, Cambridge, USA, 1997.

Puech, C.: Utilisation de la télédétection et des modèles numériques de terrain pour la connaissance du fonctionnement des hydrosystèmes, Habilitation à diriger des recherches, Grenoble University, 2000.

Quinn, P., Beven, K., Chevallier, P., and Planchon, O.: The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models, Hydrological Processes, 5 (1), 59–79, 1991.

Regues, D. and Gallart, F.: Seasonal patterns of runoff and erosion responses to simulated rainfall in a badland area in Mediterranean mountain conditions (Vallcebre, Southeastern Pyrenees), Earth Surface Processes and Landforms, 29 (6), 755–767, 2004.

Reitberger, J., Krzystek, P., and Stilla, U.: Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees, International Journal of Remote Sensing, 29 (5), 1407–

1431, 2008.

Schultz, G. and Engman, E., eds.: Remote Sensing in Hydrology and Water Management, Springer-Verlag, Berlin, 2000.

Sithole, G. and Vosselman, G.: Experimental Comparison of Filter Algorithms for Bare-Earth Extraction from Airborne Laser Scanning Point Clouds, ISPRS Journal of Photogrammetry and Remote Sensing, 59 (1-2), 85–101, 2004.

Torri, D. and Rodolfi, G.: Badlands in changing environments: an introduction, CATENA, 40 (2), 119–125, 2000.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., and Studnicka, N.: Gaussian Decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner, ISPRS Journal of Photogrammetry and Remote Sensing, 60 (2), 100–112, 2006.

Wagner, W., Hollaus, M., Briese, C., and Ducic, V.: 3D vegetation mapping using small-footprint full-waveformairborne laser scanners, International Journal of Remote Sensing, 29 (5), 1433–1452, 2008a.

Wagner, W., Hyyppa, J., Ullrich, A., Lehner, H., Briese, C., and Kaasalainen, S.: Radiometric calibration of full-waveform small-footprint airborne laser scanners, in: IAPRS, vol. 37 (Part 1), Beijing, China, 2008b.

Walling, D.: Soil erosion research methods., chap. Measuring sediment yield from river basins, pp. 39–73, Soil and water conservation society, Iowa, USA, 1988.

Wischmeier, W., , and Smith, D.: Predicting rainfall erosion losses: a guide to conservation planning - Agriculture Handbook, 537, US Dept Agric., Washington, DC, 1978.

Zevenbergen, L. and Thorne, C.: Quantitative Analysis of Land Surface Topography, Earth Surface Processes Landforms, 12, 1987.

Zhang, L., O'Neill, A. L., and Lacey, S.: Modelling approaches to the prediction of soil erosion in catchments, Environmental Software, 11 (1-3), 123–133, 1996.

Zwally, H. J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., Bufton, J., Dezio, J., Hancock, D., Harding, D., Herring, T., Minster, B., Quinn, K., Palm, S., Spinhirne, J., and Thomas, R.: ICESat's laser measurements of polar ice, atmosphere, ocean, and land, Journal of Geodynamics, 34 (3-4), 405–445, 2002.

Fig. 14. Orthoimage of the Draix area.

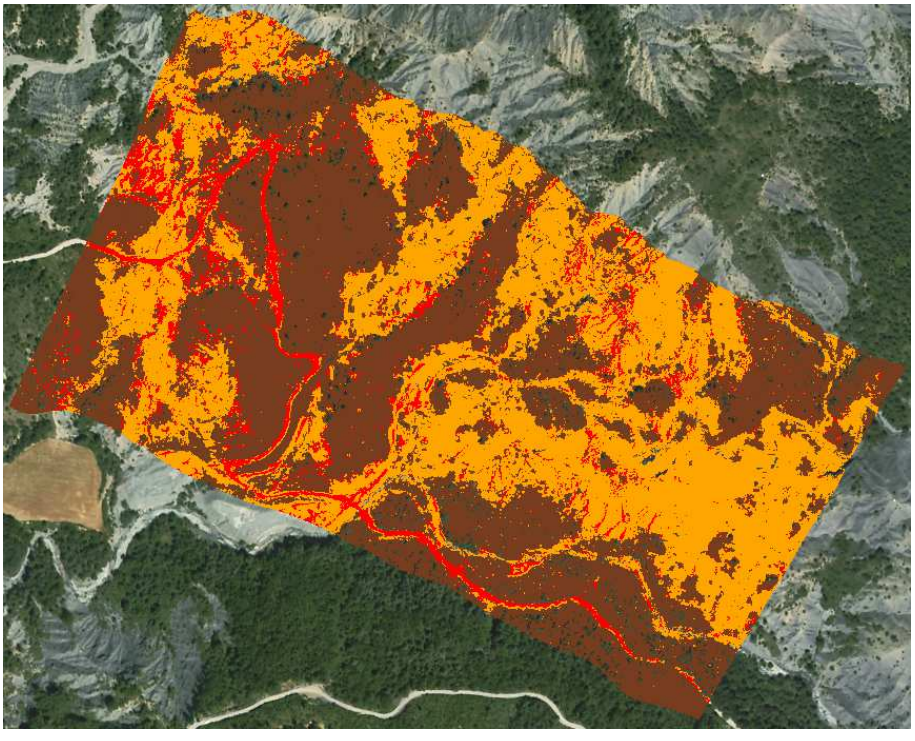


Fig. 15. Classification results: LAND (dark brown), ROAD (red), ROCK (orange).

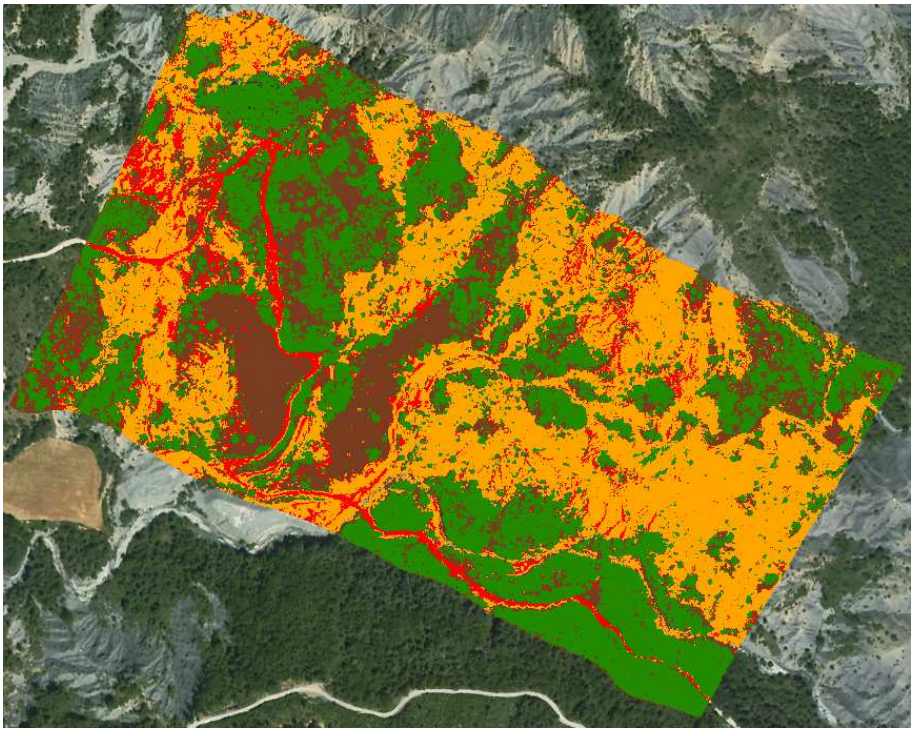


Fig. 16. Classification results: LAND (dark brown), ROAD (red), ROCK (orange) and VEGETATION (dark green).

Table 11. Field comparison of the DTM **after** adjustement.

		T_{opt} (m)	Mean±Stddev (m)
S6	West↬East	[-1.6 0.4 0.2]	**-0.13 ± 0.46**
	East↬West	[-1.6 0.8 0.2]	-0.11 ± 0.88
S7	West↬East	[0.4 1.6 0]	**0.06 ± 0.33**
	East↬West	[0.4 1.2 0]	-0.04 ± 0.62
S8	West↬East	[-0.8 -0.8 0.2]	**-0.06 ± 0.41**
	East↬West	[-0.8 -1.6 0]	-0.04 ± 0.42

Table 12. Field comparison of the DTM **before** adjustement.

DTM	# Pts	Mean $\pm$ Stdd (m)	RMS (m)
S6	1749	0.18 ± 1.05	1.07
S7	2797	-0.20 ± 0.79	0.82
S8	2886	0.41 ± 0.65	0.77

Table 13. Morphological quality criteria results

DTM	S6	S8
Detected crests (%)	72.6	47.1
Detected thalwegs (%)	53.5	44.7
Overall (%)	62.8	45.8

Table 14. Confusion matrix corresponding to the classification with {I, FWHM, d_{DTM}}.

# points		ROCK	ROAD	VEGET^{o}	LAND
71216	ROCK	**69.6**	22.4	0.4	7.3
13244	ROAD	7.5	**84.2**	0.1	6.6
402995	VEGET^{o}	0.9	0	**94.2**	4.7
279321	LAND	9.1	19.4	2.7	**68.6**
	AA	79.1 %			

Table 15. Confusion matrix corresponding to the classification with $\{d_{DTM}, RGB_{RAW}\}$.

# points		ROCK	ROAD	VEGET^o	LAND
71216	ROCK	**73**	19.8	0.9	5.9
13244	ROAD	10.4	**74**	0.2	13.8
402995	VEGET^o	0.9	0.4	**95.8**	2.8
279321	LAND	7	4.2	2.9	**85.8**
	AA	82.1 %			

Table 16. Confusion matrix corresponding to the classification with $\{d_{DTM}, RGB_{IGN}\}$.

# points		ROCK	ROAD	VEGET^o	LAND
71216	ROCK	**89.3**	8	0.2	2.2
13244	ROAD	3.7	**93.5**	0	1.2
402995	VEGET^o	0.8	0.2	**95.7**	3.2
279321	LAND	4	1.5	4.3	**90**
	AA	$92.2\,\%$			

Table 17. Confusion matrix corresponding to the classification with $\{_{\text{DTM}}$, I, FWHM, $\text{RGB}_{\text{IGN}}\}$.

# points		ROCK	ROAD	VEGET^{o}	LAND
71216	ROCK	**88.2**	10.4	0.1	1
13244	ROAD	3.7	**93.8**	0	1
402995	VEGET^{o}	0.1	0.2	**96.5**	3.1
279321	LAND	3.5	1.6	4	**90.7**
	AA	$92.3\,\%$			

Table 18. Confusion matrix corresponding to the classification with $\{d_{DTM}, I, FWHM, RGB_{RAW}\}$.

# points		ROCK	ROAD	VEGET^{o}	LAND
71216	ROCK	**76.2**	19.5	0	4
13244	ROAD	13.4	**77.1**	0.1	7.8
402995	VEGET^{o}	0.3	0.1	**95.3**	4.2
279321	LAND	5	5	2.7	**87.1**
	AA	83.9 %			

Table 19. FWHM and variations of the return pulse (**bold**) in ns w.r.t. a Gaussian emitted pulse of FWHM=5 ns - Range=500 m, 1000 m, 2000 m.

r_0	$\theta \downarrow \ \varphi \rightarrow$	10°		30°		60°	
500 m	0.3 mrd	5.0	**0**	5.0	**0**	5.1	**0.1**
	1 mrd	5.0	**0**	5.1	**0.1**	6.2	**1.2**
	2 mrd	5.0	**0**	5.5	**1.5**	9.7	**4.7**
1000 m	0.3 mrd	5.0	**0**	5.0	**0**	5.41	**0.41**
	1 mrd	5.0	**0**	5.5	**0.5**	9.7	**4.7**
	2 mrd	5.2	**0.2**	7.1	**2.1**	19	**14**
2000 m	0.3 mrd	5.0	**0**	5.2	**0.2**	6.7	**1.7**
	1 mrd	5.2	**0.2**	7.1	**2.1**	19.0	**14**
	2 mrd	5.8	**0.8**	12.7	**7.7**	39.0	**34**