POLITECNICO DI TORINO

Master degree course in Environmental and land engineering

Master Degree Thesis

Calibration of a stochastic model for riparian vegetation dynamics from LiDAR acquisitions

The case study of Cinca River

Supervisors:
Prof. Carlo Camporeale
Prof. Paolo Vezza
Prof. Monica Riva

Candidate:
Melissa Latella
s240202

Academic Year 2017-2018

Abstract

Carbon cycle has recently been revised in view of the understanding of the contribution of inland waters to the definition of the total C budget. Particularly, the role of rivers has been revaluated, recognizing their relevance as active components of this cycle, and not merely as passive pipes that convey carbon from lands to oceans, and thus driving numerous academic institutions to point their researches to the so-called *River Carbon Cycle*. As the contribution of Politecnico di Torino in such direction is the investigation of the carbon sequestration in riparian corridors, this dissertation was inserted in this frame by the definition of its two main objectives: the evaluation of the use of LiDAR to estimate riparian biomass and the calibration of a stochastic model, which describes riparian vegetation growth according to topographic and hydrological constraints, on the basis of real LiDAR data measurements.

LiDAR, *Light Detection And Ranging systems*, is a remote sensing technique that is based on the principle of measuring distances by laser light and it has a wide range of applications, as forestry. Although numerous LiDAR-based models for the estimation of above-ground biomass in forest stands have been regressed, no studies concern riparian corridors, where tree population is generally younger and sparser. Thus, the reliability of these forest models in riparian environments was tested. Firstly, an insight about the state of the art of LiDAR technology and its airborne implementation for forestry was provided. Secondly, a selection of literature models was applied to the study area, which was a segment of Cinca River (Spain), in order to choose the most reliable one, after the processing of LiDAR data with FUSION/LDV, which is a free software package released by the US Department of Agriculture (USDA). Finally, the influence of grid discretization on results and the consistency of the chosen model with consolidated allometric formulas were evaluated. The achieved results were the definition of a procedure that combines the use of GIS and FUSION/LDV for determining vegetation statistics and the choice of the model proposed by Means et al. (2000) for the conversion of these statistics in above-ground biomass. This model demonstrated to be implementable with a relatively coarse grid discretization (i.e. cellsize 10x10 m), thus requiring short computation time, and to return biomass values that are consistent with the ones provided by allometric formulas.

The distribution of phreatophyte riparian vegetation can be described by a stochastic model that was provided by Camporeale and Ridolfi in 2006. According to this,

vegetation dynamics are influenced by the randomness of hydrological fluctuations and topography of the riparian transect. Thus, input data for the model are hydrological and geometric information as well as parameters linked to biological features of vegetation. Despite literature values were already available for most of these parameters, the one that represents the ratio between the rates of, respectively, vegetation decay and growth, still needed to be properly set. To this purpose, a method to process the available data and calibrate the model was defined. According to this method, hydrological data are processed with HEC-RAS and MATLAB to obtain the probability distribution function and integral scale of water levels, while the geometry is set by the realization of a DTM with FUSION/LDV. The probability distribution function of vegetation and its two first moments are obtained by the implementation of the previously defined procedure for processing LiDAR data. Finally a MATLAB script based on least squares allows to to minimize the deviation among computed values of dimensionless biomassand real LiDAR-derived data. The definition of this procedure was done referring to the hydrological data and LiDAR acquisitions that were available for the Cinca River's study area.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The carbon cycle is one of the fundamental biogeochemical processes that enable life on Earth. It comprises fluxes regulated by geological and biological processes, that means carbon exchanges among the various environmental matrices, the long-term sequestration in carbon sinks and its eventual release. Although carbon cycle has been studied since the end of XVIII century [26], it has been recently updated in view of the revaluation of the active contribution of inland waters to the definition of the total C budget [13], [19].

As a consequence, institutions have started to address their researches to a deeper understanding of the rivers dynamics in this frame, particularly focusing on the role of riparian vegetation.

The conventional carbon cycle

Carbon is one of the major component of organic compounds and part of minerals, such as carbonates, and it is mutually exchanged among the various environmental matrices that constitute the Earth system. This continuous series of parallel processes is called *Carbon Cycle* and, in its conventional description, it consists of carbon circulation among three reservoirs [13], [19], [49]:

- *Atmosphere:* it is the smallest in terms of storage as it relevance is mainly due to the linkage it provides to the other two matrices. Import of carbon in atmosphere derives from outgassing from water and biological respiration.

- *Lands:* they store a large amount of carbon both in soil and vegetation. Import of carbon derives either from the uptake of vegetation during photosynthesis or from the dry and wet deposition of particles from atmosphere.

- *Oceans:* they constitute the largest reservoirs, storing carbon in both its inorganic and organic speciations. Dissolved inorganic carbon (DIC), namely the ensemble of dissolved CO_2, HCO_3^- and CO_3^{2-}, derives from the weathering of rocks; while

dissolved organic carbon (DOC) and particulate organic carbon (POC)[1] consist of both living organisms and fragments of dead ones. Part of the input carbon is directly uptaken from atmosphere, while another part is conveyed by surface streams and groundwater[2].

Definition of *River Carbon Cycle*

Inland waters constitute solely the 1% of Earth's surface and, in the conventional carbon cycle, are considered as pipes that passively convey carbon from lands to oceans and thus not contributing to the C budget. Nevertheless, in the last decade, studies have shown that they play an active role in carbon cycle, as burial and outgassing could strongly influence the amount of carbon conveyed to the ocean [13].

Focusing to the so-called *River Carbon Cycle*, Cole et al. (2007) [19] demonstrated that streams are place of active C transformation and that their inclusion in C budget could reduce imbalances that currently occur when computing C fluxes among reservoirs. Also, they provided roughly estimates that show how much the import of carbon from land to rivers exceeds the amount of carbon delivered by rivers to oceans, thus highlighting the relevance of the actions of streams in the carbon cycle. According to them, lands deliver approximately 1.9 $Pg \cdot y^{-1}$ of C to rivers, while solely 0.9 $Pg \cdot y^{-1}$ are conveyed to oceans. The remaining 1.0 $Pg \cdot y^{-1}$ are partially buried in sediments along the streams (≈ 0.23 $Pg \cdot y^{-1}$) or emitted to atmosphere by outgassing, that means by the expiration of water biota (≈ 0.75 $Pg \cdot y^{-1}$). As the import from terrestrial ecosystem is equal or exceeds the gross primary production of riparian environments, rivers play both the role of net sources of CO_2 to the atmosphere and net sinks of C in sediments.

Focusing on the role of riparian vegetation

As anthropic activities are strongly influencing carbon fluxes, thus altering the pre-existing equilibrium and increasing the release of carbon dioxide into the atmosphere [49], various techniques have been developed to deal with CO_2 excess, promoting innovative sinks for carbon capture and storage. The simplest one is the nature-inspired forestation, or rather the forced establishment of stands of trees in unvegetated areas, in order to enhance the conversion from carbon dioxide to biomass [15]. Nevertheless, in the frame of forestation, the relevance of riparian environments in C budget is usually neglected, focusing solely on the capability of forests to sequester carbon dioxide.

---

[1] DOC and POC are both convey to oceans by streams. DOC constitutes the major transfer of carbon from terrestrial ecosystems and it flows with water. On the contrary, POC usually tends to settle and its movement consists of a series of discrete events led by hydrodynamics lift and drag and, therefore, influenced by the hydrological regime. However, separation between DOC and POC is not sharp, as their properties vary with continuity and the two phases continuously interact along the river flow [12].

[2] River Carbon Cycle refers to the processes occurring along streams and comprising C transfer and transformation.

Indeed, carbon sequestration in river corridors could exceed the storage in mature forests, as the continuous process of growth and uprooting, exerted by streams on riparian vegetation, enhances both the storage of carbon in the dead biomass, which is buried or transported downstream by the river itself, and the capture of additional carbon by the growth of new colonizing species. Hence, the contribution of riparian vegetation for carbon sequestration may be essential, especially in those areas where the anthropic expansion has reduced the spaces that could have been available for the ordinary forestation, and, therefore, needs to be more deeply investigated.

Chapter 2

Objectives

As the role of streams in the carbon cycle has been revaluated, numerous academic institutions are now pointing their researches to deepen the various aspects of the River Carbon Cycle. Among them, Politecnico di Torino is focusing on evaluating the amount of carbon that can be potentially sequestered by vegetation in riparian corridors, comparing it with the amount stored in the floodplain. In such frame, this dissertation had two main objectives: the evaluation of the use of LiDAR measurement to estimate riparian vegetation biomass and the calibration of a stochastic model, which describes riparian vegetation growth according to topographic and hydrological constraints, on the basis of real data deriving from raw LiDAR data acquisition and processing.

LiDAR for estimating riparian biomass

Despite LiDAR technologies are widely used in forestry and a number of both site-specific and large-scale models to convert airborne LiDAR data in biomass have been released, neither specific models addressed to relate LiDAR measurements with riparian vegetation have been defined nor the reliability of existing ones has been assessed. Thus, one of the two objectives of this dissertation was to understand the state of the art in the use of airborne LiDAR for estimating above-ground biomass in riparian corridors and to define a methodology to this purpose. Such methodology was successively applied to the study area in order to obtain the vegetation data necessary for the model calibration.

Stochastic model calibration

In 2006, Camporeale and Ridolfi [16] defined a stochastic models that relates the distribution of riparian vegetation with the topography of the river transect, the hydrological regime and the intrinsic features of local vegetation. Although most of the parameters referring to vegetation had already been determined for the most common tree species in temperate climate, the one related to vegetation decay and its ratio with the rate of growth, still needed to be properly set on the basis of real measurements. The second objective of this dissertation was, therefore, to define a procedure for the stochastic model calibration and implementing it in the study area.

Chapter 3

Study case

3.1 Context

The study area was a segment of Cinca River, 13 km long, in the region of Aragón, Spain. A complete description of Cinca river is reported in *Estudio de la calidad ecológica integral de los tramos fluviales más importantes del Río Cinca*, [61], [62].

Cinca River has its source in the Pineta Valley, in Pireneos Mountains, and it flows from NE towards SW, ending in the confluence with Segre River. It is one of the major tributaries of Ebro River, being 191 km long and having a catchment area, which is shown in Figure 3.1, of 10,000 m^2. Despite the relevant level of anthropization, especially in its middle and lower segments where there is a high number of hydroelectric plants, small reservoirs, mainly dedicated to large irrigation infrastructures and industries, and urban settlements, the contribution of Cinca River to Ebro discharge shows a marked seasonal regime [62]. Cinca River has numerous affluents; among these, the most important *right-bank* tributaries are Vellos, Ara, Vero and Alcanadre Rivers and *left-bank* are Barrosa, Cinqueta, Nata, Ésera, Isábena, Sosa and Clamor Amarga (also called Tamarite) Rivers.

With respect to hydromorphology and ecology, headwaters are in good conditions, while the lower segments are affected by anthropic interventions that reduce the water level and alter the water quality. Nevertheless, although water quality decreases moving downstream, due to the anthropic pressure, the ecological status of the river is defined at least as *moderate* in all its segments, that means riparian habitats are not strongly affected by artificial water regulation and vegetation dynamics occurs along the river. Riparian vegetation comprises shrubs and young trees, even though its spread in the floodplain has been reduced and controlled to create room for crop fields and artificial embankments [62].

3.2 Study area

The choice of a segment of Cinca River as study area was done according to morphological features, or rather looking for sections where the river was not channelized, but wandering, with bars and islands. Also, the pressure of crop fields and infrastructures was

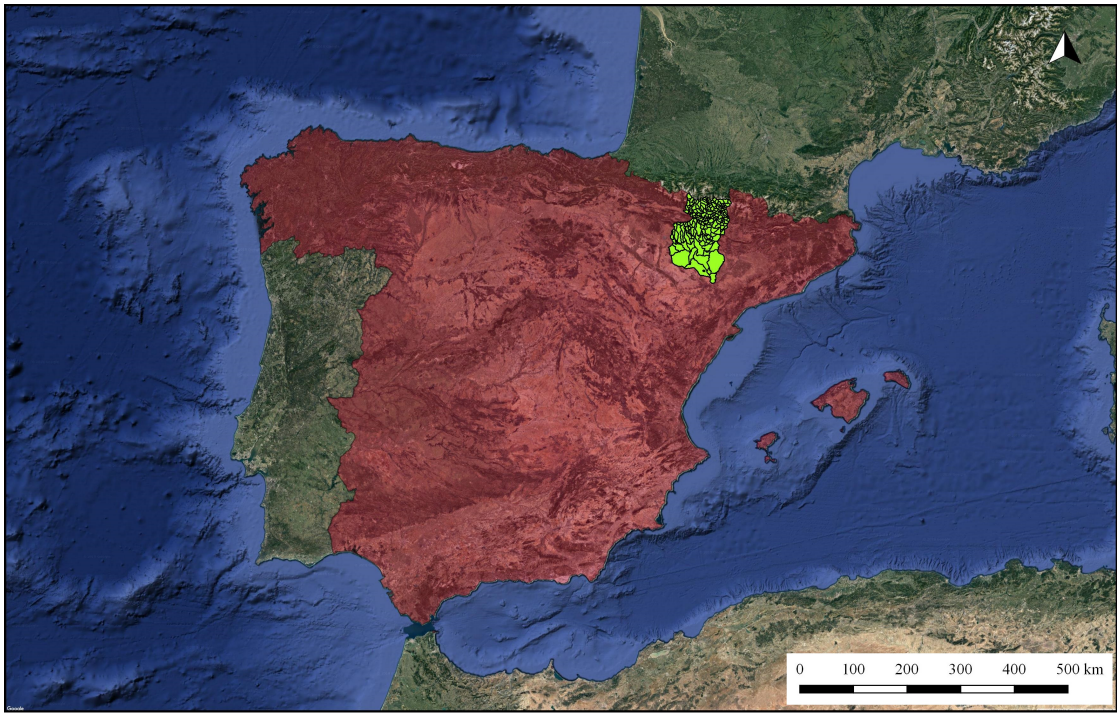

Figure 3.1. Geographic location of Cinca River catchment. Red colour indicates Spain, while green colour indicates Cinca catchment.

considered in order to find an area where riparian vegetation had enough room to grow. Finally, the availability of temporal series of daily streamflows was taken into account to describe local hydrological characteristics of the river and calibrate the stochastic vegetation model

The selected segment is delimited by two tributaries of Cinca River, starting at the confluence of Alcanadre River, close to the town of Ballobar, and downstream the confluence with Clamor Amarga River, which is close to the town of Zaidín. Its length is approximately 13 km and the corresponding catchment area is 9,409 km^2.

Hydromorphological features

Cinca River, whose bed is constituted by Quaternary sediments, flows in an alluvial floodplain made of conglomerates, gravel and clay. This floodplain is more than 2 km wide and it is surrounded by low hills of soft materials, as sandstones and marls [62]. Orthophotos show that the selected segment can be classified as wandering, having some bars and islands.

Riparian vegetation

As crop fields occupy part of the floodplain, riparian vegetation can grow only along a narrow band between the fields and the water or above the islands. Most of the vegetation is constituted by shrubs, whose dominant species are bitter willow *(Salix Elaeagnos)*, black willow *(Salix Atrocinerea)* and white willow *(Salix Alba)*. Also, there are some mastics *(Pistacia Lentiscus)*, black poplars *(Populus Nigra)* and white poplars *(Populus alba)*. Grass-like plants, as common reeds *(Arundo Donax and Phragmites Australis)*, and bulrush *(Typha sp.)*, grow where inundations are more frequent [61].

3.3 Available hydro-morphological data

Overall information about the study area

The report *Estudio de la calidad ecológica integral de los tramos fluviales más importantes del Río Cinca*, carried out by *United Research Services Corporation (URS)* and commissioned by *Comisaría de Aguas de la Confederación Hidrográfica del Ebro (CHE)*, provides an insight about the hydromorphological and biological status of the Cinca River and its major tributaries (i.e. Alcanadre, Ésera, Ara e Isábena Rivers). Particularly, volume II, *Río Cinca: Informe de síintesis* [62], consists of an overview of the entire river, expliciting hydromorphological, physico-chemical and ecological parameters, while volume III, *Río Cinca: Atlas* [61], deepens the description focusing to one of the 14 segments, which constitute the river, per chapter. The study area is part of the 13^{th} segment, however no bathymetric cross sections were found.

Hydrological data

Hydrometric data were derived from the network of gauging stations that monitor water discharges and levels in the Ebro catchment. They were provided by *Confederación Hidrográfica del Ebro* [3]. As there were no stations along the selected segment, the data used in this study were taken from hydrometric station 9017 that is located in the town of Fraga, downstream the town of Zaidín, and that refers to a catchment area including both the catchment of the selected segment and the catchment of the Clamor Amarga tributary. In Figure 3.2 the catchment of the station 9017 (on the left) is compared to the catchment of the study segment (on the right), while in Figure 3.3 the selected segment is highlighted by colour red and the location of the two gauging stations by black dots.

Gauging stations 9017 and 9225 provided the temporal series of average, minimum and maximum daily discharge and water level, respectively expressed in m^3s^{-1} and in m, despite only the average daily discharges were used for calibrating the stochastic model. Temporal availability of data for both the gauging stations are summarized in Table 3.1.

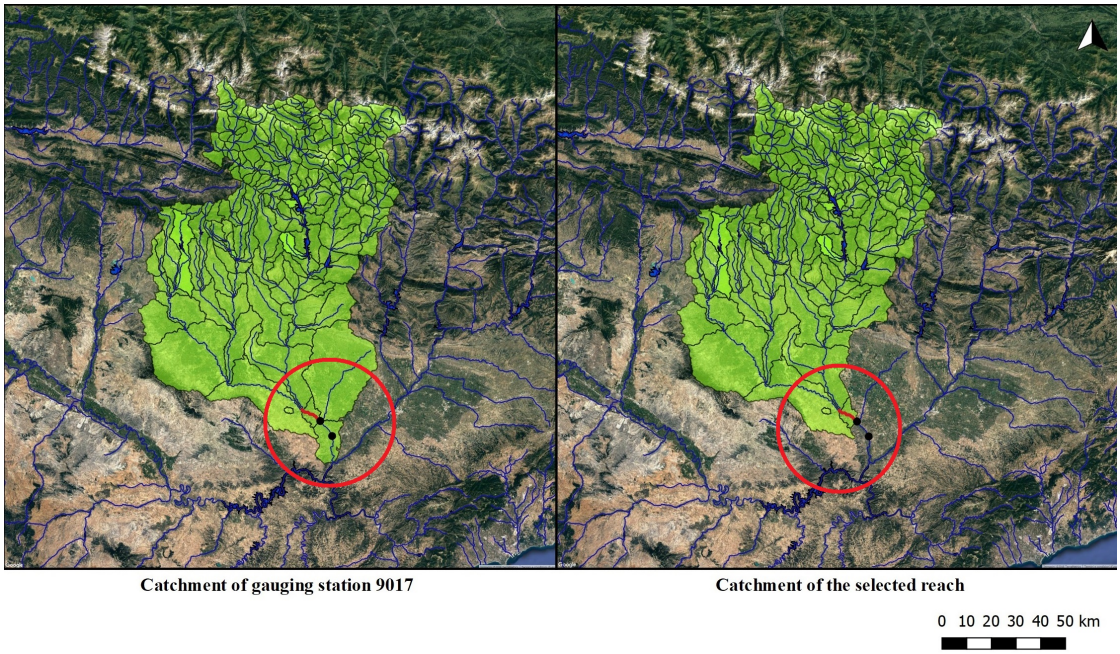

Figure 3.2. Comparison between catchment areas.

Table 3.1. Length of temporal series for the selected gauging stations.

Station	Date of measurements
9017	from 03/10/1947 to 31/12/2017
9225	from 22/04/2008 to 31/12/2017

Two main issues accompanied the availability of hydrometric data. Firstly, hydrometric useful data comprised only flow rates[1], while the calibration of the stochastic model required as input the probability density function and the integral scale of water level in a plethora of points along the segment; to solve this problem, a conversion from these to water levels was done, as explained at Section 9. Secondly, the available data did not refer to the sole selected segment, as station 9017 is located at the ending section of Cinca River and station 9225 is related to Clamor Amarga affluent; thus, it was necessary to understand if the contribution of Clamor Amarga was negligible so that data from station 9017 could be referred directly to the selected segment.

[1]Even though water levels measured at station 9017 and 9225 were available, they could not be used. In fact, water level is a site-specific parameter depending on many factor, such as river width, slope and roughness at the measuring section. Thus, values that are measured close to hydrometric stations are not representative for the entire segment, unless the feature of the channel are the same.

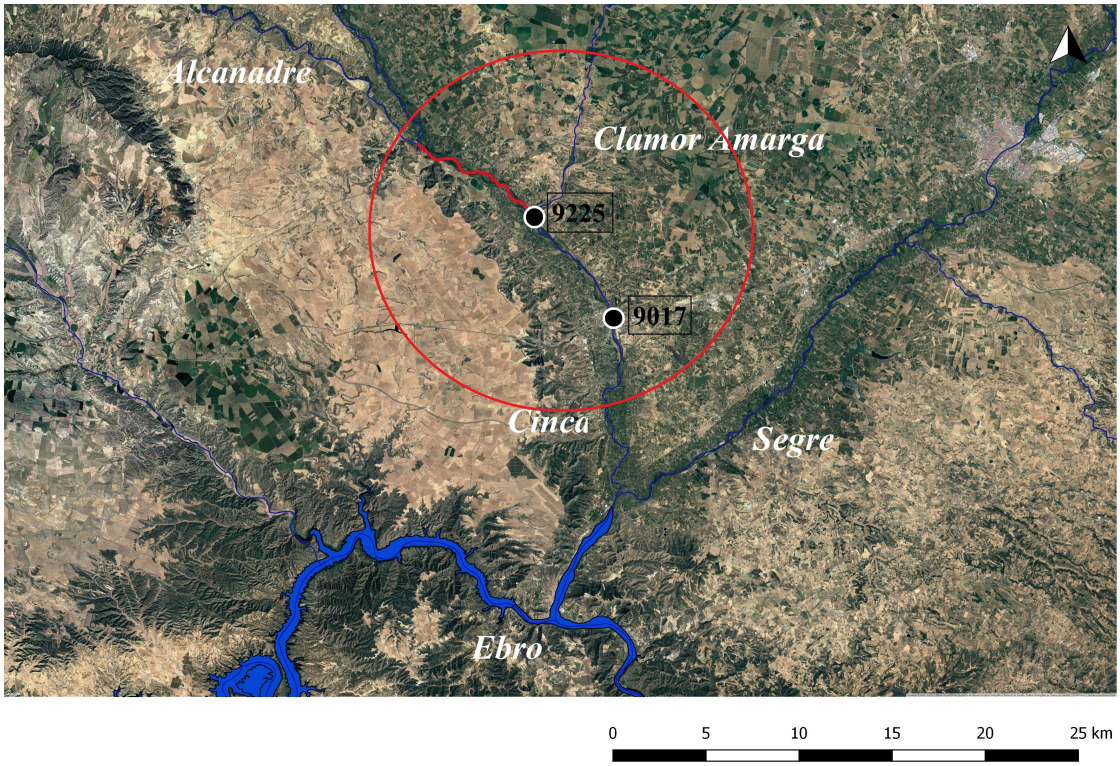

Figure 3.3. Position of the selected segment (red line) and the hydrometric stations (black dots).

To this purpose, data from station 9017 were compared to data from station 9225, which is located at the end of Clamor Amarga tributary, just before its confluence with Cinca River, in order to highlight an eventual influence of the discharge of Clamor Amarga on the Cinca regime. As shown in Figure 3.4 and Figure 3.5, the contribution of Clamor Amarga River was negligible as it represented, on average, only 6.5% of the total discharge of Cinca River with $RMSE$ equal to 2.7 $m^3 \cdot s^{-1}$. Figure 3.4 and Figure 3.5 refer to measurements from 22nd April 2008 to 31st December 2017, as this represents the available recording interval for station 9225.

LiDAR data

LiDAR data for biomass estimation were downloaded by *Centro de Descargas*, an online platform of *Centro Nacional de Información Geográfica (CNIG)*, where all the data generated by *Instituto Geográfico Nacional (IGN)* are available for public users. LiDAR acquisitions above Spanish territory are made according to the *Plan Nacional de Ortofotografía Aérea (PNOA)*, whose aim is creating a 3D point clouds for the entire nation. The involved equipment is a small-footprint and discrete-return airborne laser scanner,

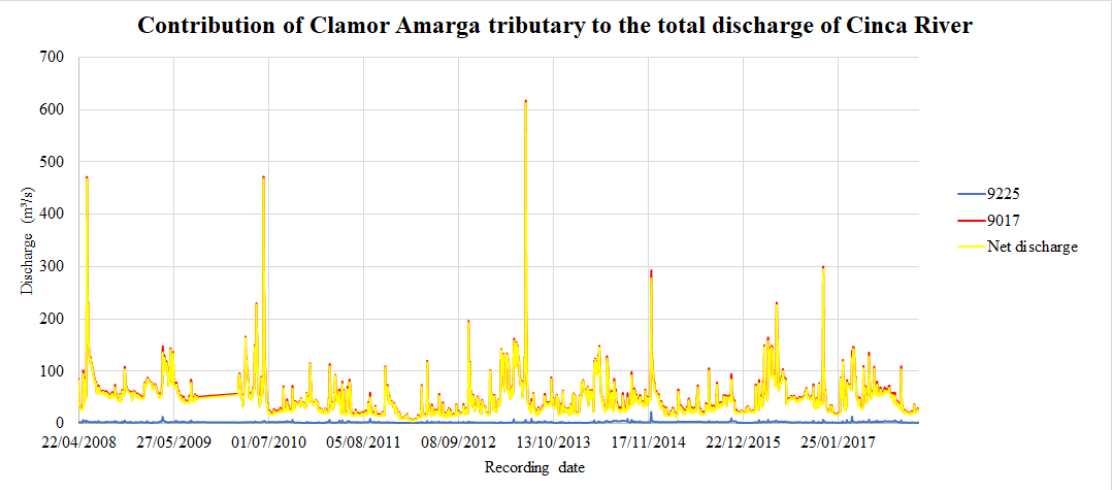

Figure 3.4. Analysis of the possibility to neglect the contribution of Clamor Amarga tributary to the total discharge of Cinca River.

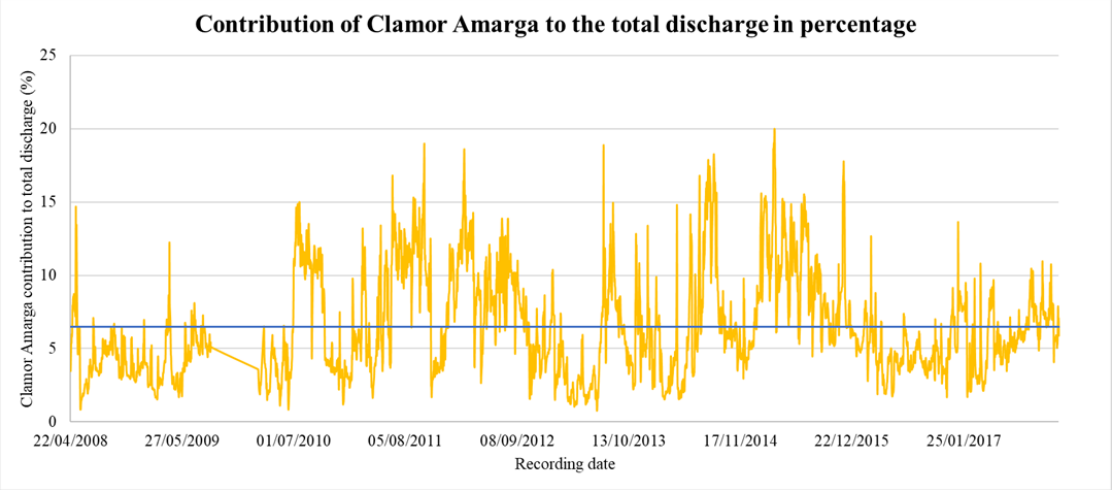

Figure 3.5. Quantification of Clamor Amarga discharge in percentage. The blue line indicates the average percentage, equal to 6.5.

the geodetic reference system is ETRS89 and the cartographic projection is UTM. The point density is 0.5 points·m^{-2} as first coverage and 1.0 points·m^{-2} as second coverage, while the elevation RMSE is less than 0.20 m. Due to the high precision of PNOA acquisitions, the 3D point clouds allow users to realize accurate terrain and surface models. More in detail, the study area was scanned on 20th October 2016 with a point density of 0.5 points·m^{-2}. LiDAR datasets were downloaded as *.laz* files, each of which covers a square portion of territory with side 2 km long [7]. Thus, nine datasets, whose names are

reported in Table 3.2, had to be merged to completely cover the study area, as shown in Figure 3.6.

The intent of the PNOA was to have at least 0.5 points·m^{-2} [1] and the used datasets have a nominal return density of 0.68 points·m^{-2}, as found out by processing data with FUSION/LDV software (see Section 5.1) and by observing the resulting first returns density images in Figure 3.7. In Figure 3.8, the intensity of returning pulses is shown, ranging from black (i.e. no return, pulses completely absorbed by the surface) to white (i.e. pulses completely reflected by the target surface).

In this dissertation, LiDAR outcomes were used to evaluate models for biomass estimation and to provide vegetation data for the stochastic model calibration.

Table 3.2. List of used datasets, downloaded from Centro de Descargas [1].

2016
PNOA_2016_ARA_266-4612_ORT-000-COL
PNOA_2016_ARA_268-4612_ORT-000-COL
PNOA_2016_ARA_268-4610_ORT-000-COL
PNOA_2016_ARA_270-4610_ORT-000-COL
PNOA_2016_ARA_272-4610_ORT-000-COL
PNOA_2016_ARA_272-4608_ORT-000-COL
PNOA_2016_ARA_274-4608_ORT-000-COL
PNOA_2016_ARA_272-4606_ORT-000-COL
PNOA_2016_ARA_274-4606_ORT-000-COL

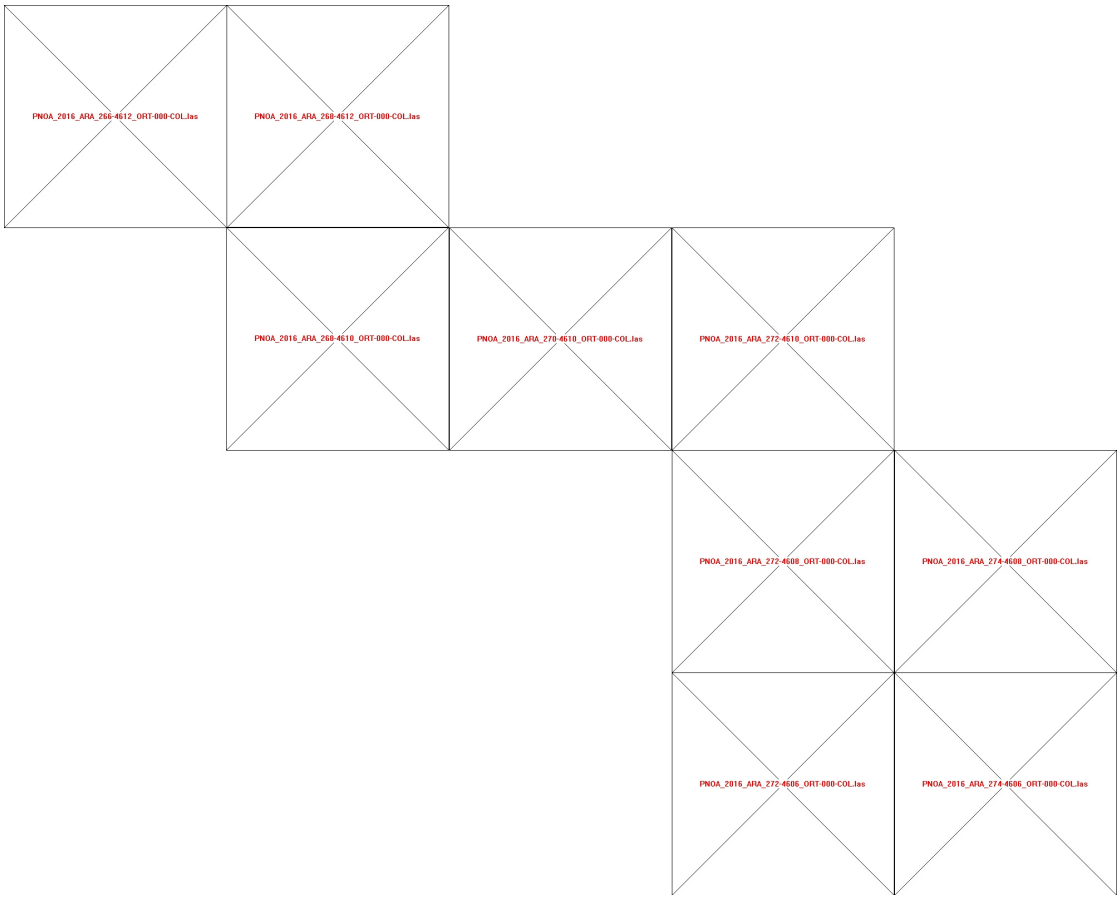

Figure 3.6. LiDAR data coverage of the study area. Red writings are the name of the data file downloaded by CNIG online platform 9[1].

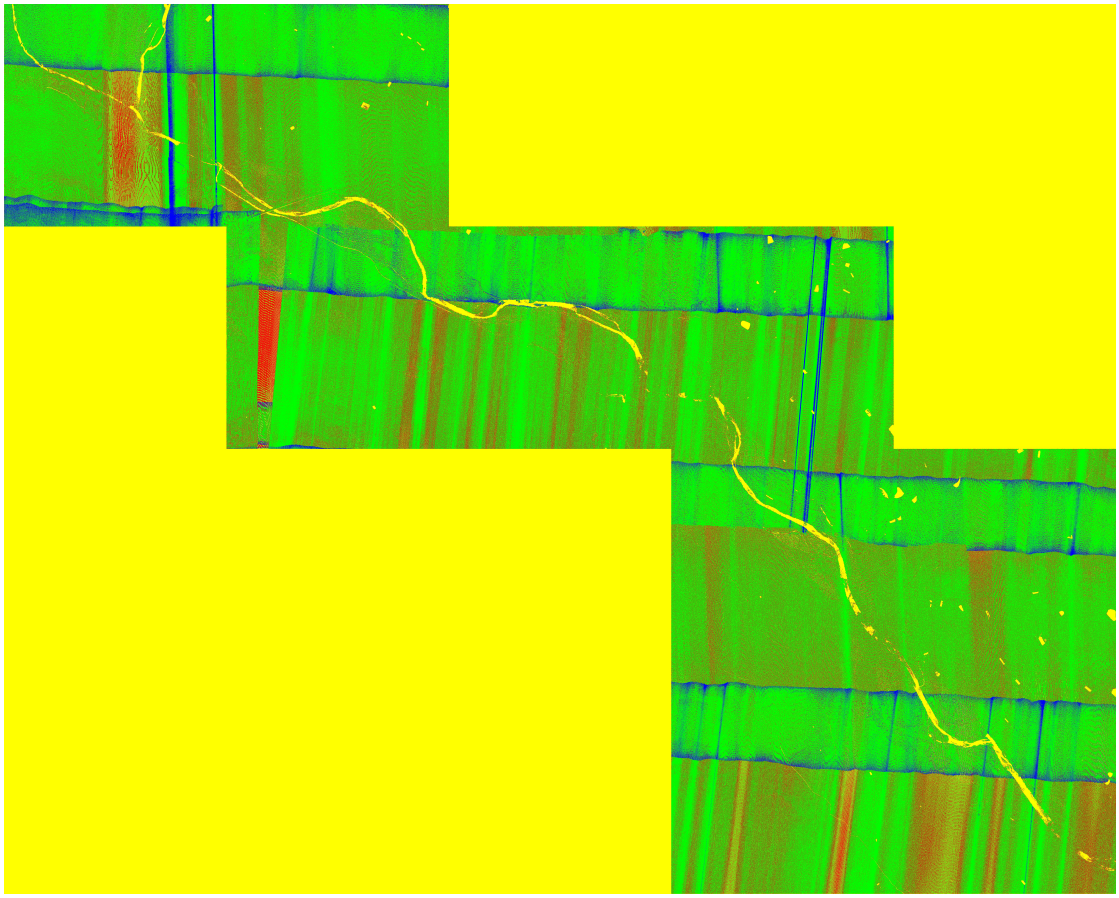

Figure 3.7. First returns density. Yellow indicates no points, red density less than 1 point·m^{-2}, green density ranging from 1 to 3 point·m^{-2} and blue density higher than 3 point·m^{-2}.

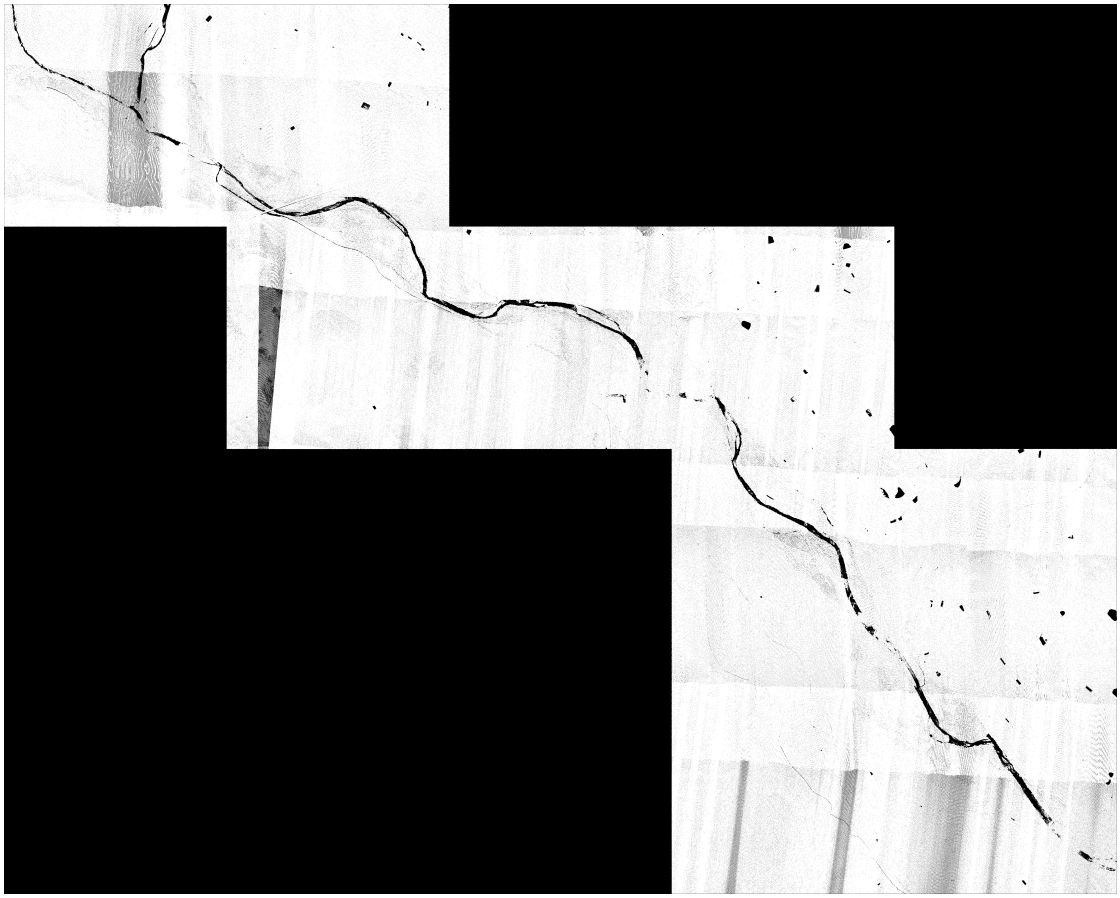

Figure 3.8. Intensity of return in grey scale. Black indicates complete absorption of pulses while white indicates complete reflection.

Part I

LiDAR for estimating riparian biomass

Chapter 4

Introduction

Light Detection And Ranging systems (LiDAR) is a remote sensing technique that is based on the principle of measuring distances by laser light and it has a wide field of application, ranging from meteorology to traffic management.

To reproduce 3D digital elevation model, LiDAR technology is usually mounted on an aircraft so that a wide area can be investigated [45]; LiDAR provides a 3D point cloud that represents the location where surfaces have been hit by the laser beam and that can be processed to reproduce the topographic surface of the ground and the elevation model of vegetation and buildings. Although LiDAR was born in 1960s, commercial airborne laser-scanner technologies started to develop solely in the second half of 1990s [24]. Since that moment, the number of fields of application of such technology has increased so fast that the advantages, but also the obstacles, in the use of LiDAR in these contexts has been solely partially studied, leaving a wide room for further researches.

This dissertation aimed to evaluate airborne LiDAR potentialities for the estimation of above-ground biomass in riparian corridors. One of the main objective was to compare and evaluate the models that have already been proposed and used in ordinary forestry to transform LiDAR raw data in a measure of above-ground biomass. The goal was to assess their validity also for estimating riparian vegetation and define a standard methodology for this kind of investigations in river corridors. LiDAR data related to the study area (see Section 3.2) were processed with FUSION/LDV, a free software package released by *United States Department of Agriculture (USDA)*.

4.1 LiDAR functioning principles and equipment

The functioning principle of LiDAR systems is based on the measurement of return times of emitted signals. As the speed of light is known, return times are transformed in distance data that, in turn, are used to reproduce elevation models of Earth's surface.

LiDAR is an active remote sensing system since it does not simply receive pulses from external sources, but it emits its own signal. This signal is an electromagnetic pulse that ranges from ultraviolet to infrared light (i.e. from 0.1 μm to 2.0 μm in

wavelength[1]). Once a pulse is emitted, it travels towards the Earth's surface, hits the target objects and it is reflected back to LiDAR sensors that register the elapsed time of this return. As the delay is related to distance through the speed of light, the farther the target surface is, the longer the associated return time is. However, not all the surfaces respond to laser pulses at the same way. Solid surfaces, as buildings and ground, completely reflect the pulse, whereas water totally absorbs it, avoiding the pulse to come back to the aircraft and provide a time information. Vegetation acts differently as the top of the canopy reflects part of the pulse, providing a return time that is defined *first return*, while the other part passes through it and is reflected by the ground. Usually, in vegetated areas, 1 to 5 returns correspond to one emitted pulses, allowing a vertical characterization of the vegetation structure [40]. Thus, raw data provided by LiDAR systems is a three-dimensional point cloud representing the Earth's surface with both its natural and anthropic elements.

Many are the ways to classify LiDAR systems. One differentiation can be done according to the features of recorded signals, distinguishing in *discrete return* systems, if for each emitted signal they record solely the major peaks in the returning electromagnetic wave and associate a point in space to each peak, and *full waveform recording*, if LiDAR systems digitize the entire reflected energy giving a complete vertical profile [24], [68]. They can further be divided in *profiling*, if they record solely along the imaginary vertical line that runs from the sensor to the ground, and *scanning*, if they captured a series of adjacent profiles along an imaginary horizontal swathe on the ground that runs on either side of the sensor [68]. Another classification can rely on the size of laser beam diameter at the ground level, discriminating *small-footprint* systems, if diameter ranges from 0.15 to 1 m and *large-footprint* systems, in case the diameter is larger than 10 m [68]. Also, systems can be categorized according to the platform they are mounted on:

- *Ground-based:* from a historical point of view, the first application of this type was meteorology, as LiDAR can profile aerosols, water vapor, temperature and other atmospheric parameters [39]. Nowadays, portable versions of this technology, also called *terrestrial laser scanning*, are used to measure the damage of anthropic structures and ground movements as landslides and debris flows, but also to scan archaeological artifacts; the equipment of these systems usually consists of a small size and light weight laser scanning device and a laptop to register data [35].

- *Spaceborne:* these kind of LiDAR systems are usually large-footprint and full waveform recording and operate in various contexts. Examples for this category are *Mars Orbiter Laser Altimeter (MOLA), Lunar Laser Ranging Instrument (LLRI), Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO)* and *Geoscience Laser Altimeter System (GLAS)*. *MOLA* and LLRI are used in astronomy; the former is part of the Mars Global Surveyor spacecraft launched in

[1]Wavelength is linked to the target of LiDAR surveys. Meteorology usually requires infrared waves (1.5 - 2 μm) or ultraviolet (around 0.25 μm), terrestrial mapping near-infrared (around 1.0 μm) and bathymetry blu-green light (0.5 - 0.6 μm).

1996 to map Mars surface [6], while the latter was launched in 2008 and designed to provide data about lunar topography with a resolution of 10 m [5]. *CALIPSO* was launched in 2006 to obtain high precision profile of clouds and aerosols [2]. *GLAS* was launched in 2003 to study ice-sheets altimetry and evolution, measure vegetation and collect atmospheric data; it emits 40 laser pulses per second, capturing spots having a footprint of around 70 m and spaced 170 m apart along Earth's surface [4], [68].

- *Airborne:* as the name suggests, the equipment of airborne LiDAR systems is mounted on a moving platform, namely a single- or twin-engine plane or a helicopter, and comprises different instruments [38], [68]:
  - *Laser scanning unit:* it both emits and receives electromagnetic pulses. Although several configurations of this technology have been released, generally commercial systems for geomatics and forestry belong to the *small-footprint* typology and they work emitting discrete signals, typically around 200,000 pulses per second. Usually small-footprint systems are *discrete return* whereas large-footprint systems are *full waveform recording* as they return a complete vertical profile of vegetation. Depending on the sensor model, the scanner uses a rotating prism or a oscillating mirror to address the pulses across a line of the land below the aircraft, whilst the motion of the aircraft itself allows for the collection of parallel lines. The simultaneous combination of such a longitudinal and transverse motion allows for the achievement of a complete coverage of the investigated area.
  - *Global Positioning System (GPS):* part of this instrumentation is located on the ground as reference points, while the core is positioned on the aircraft. It provides the position and orientation of the laser scanning unit, during the emission and reception of pulses, in an absolute coordinate systems. Thanks to this information, also the distance calculated from recorded return time can be transformed in a XYZ point in the reference absolute coordinate systems.
  - *Inertial Measurement Unit (IMU):* it monitors and measures roll, pitch and yaw of the aircraft in order to correct the coordinates provided by GPS.
  - *Computer:* it controls the functioning of the whole equipment and stores collected data. Its clock records the exact time of emission and reception of pulses.

  As it is explained in the following section, fields of application of airborne LiDAR are numerous, due to their versatility and capability in providing low-cost, fast and easy post-processing of data.

With respect to the possibility to detect vegetation and measure its biomass, all the three category of LiDAR are commonly used. On the one hand, terrestrial laser scanner can capture the features of a single tree or a small stand [35]. On the other hand, spaceborne full waveform-recording systems provide information about large areas, despite the

accuracy of measured vegetation height declines with increasing slope of investigated terrain; accuracy is further affected by the intrinsic features of this instrumentation[2]. Also the adoption of full waveform-recording systems in forestry has been studied during the last decades, even though this technology is usually more employed in research than in practice, where *Airborne Laser Scanning (ALS)* systems are largely preferred. In fact, ALS can provide reliable values for inventory parameters[3] regardless the density of pulse returns[4], especially if regression models to relate vegetation metrics given by LiDAR and biomass are calibrated with on-field measures.

4.2 Examples of Airborne LiDAR application

During the 1960s, Fiocco and Smullin built the first LiDAR system [23], implementing the theoretical studies of Schawlow and Townes about maser [53] and the ruby laser invented by Maiman[5] [37]. Meteorology was the first field of application of LiDAR as well as topography [39]; for example, in 1971 a laser altimeter was used to scan the moon surface during Apollo 15 mission [10]. Since its first development, LiDAR technology has been mounted on flying platform; however, the spatial accuracy of measures done by these airborne systems considerably improved after the creation of GPS, in the 1970s.

Similarly to other remote sensing techniques, at the beginning of their development LiDAR systems were private and dedicated to research. It was solely starting from 1995 that LiDAR was on commerce for the public, allowing for a speeding up of its advancement [24]. First commercial applications mostly concerned agriculture, geomatics and forestry, involving both terrestrial and airborne systems. On the one hand, terrestrial laser scanner were, and currently are, widely used in farming and intensive agriculture, as a LiDAR-based detection can be used to characterize weeds that have grown in crop fields, once the typical height of the crop vegetation and weeds have been defined, allowing for large scale monitoring of field productivity [11]. Also, the association of LiDAR data about crop height with measures of green laser return intensity indicates the nitrogen content in some food crop, such as wheat plantations, for which the N-concentration

[2]Indeed, the penetration of highly dense canopy and the detection of ground require high pulse energy and, therefore, a low pulse rate, leading to low resolution and scarce density of sampling in the study area [68].

[3]Forest inventory parameters: indicators of the main features of trees in a stand, namely basal area, diameter at breast height (DBH), maximum, minimum and mean height, timber volume, number of trees per hectare, total biomass, canopy bulk density, canopy fuel load [66].

[4]Actually, some studies have pointed out that low densities of returns lead to an underestimation of vegetation height [27], [42], but in many cases this effect is negligible [44], [64].

[5]Maser and laser are the acronyms respectively of *Microwave Amplification by Stimulated Emission of Radiation* and *Light Amplification by Stimulated Emission of Radiation*. The basic functioning is the same for both technologies but they operate in different fields of electromagnetic spectrum as the maser works with microwaves (i.e. λ = 1 - 100 mm), while the laser with visible light (i.e. $\lambda = 4{\cdot}10^{-6}$ - $7{\cdot}10^{-6}$ mm). LiDAR systems are based on laser technology.

in the early development stages is crucial for achieving the potential maximum productivity; hence, LiDAR becomes a diagnostic tool to be used for managing plantations [22]. On the other hand, the advance of airborne systems (ALS) mainly concerned forestry and geomatics, fields where LiDAR potentialities have been investigated, as example, for landscape monitoring and wood production planning. Currently, the main uses of airborne LiDAR can be sum up as:

- *Creating high resolution DTMs:* accuracy of DTMs that have been realized starting from LiDAR data is proven not only in bare areas, where the error is usually less than 0.15 m, but also in vegetated ones; in fact, although the error slightly increases with canopy density, it always remains negligible, allowing for a sub-meter accuracy [20], [32], [51].

- *Mapping coastline and bathymetry:* accuracy of LiDAR measurements is high not only in topography, but also in bathymetry[6], allowing users to create high precision maps of shorelines or flooding zones and monitor evolution of coastlines, thus supporting coastal research and environmental management; if coupled with hyperspectral imagery, airborne LiDAR also provides maps of wetlands, beaches, coral reefs, and submerged aquatic vegetation [32].

- *Mapping vegetation structures:* due to limited temporal and economic resources and, sometimes, to safety concern and a scarce physical accessibility to the investigated site, field surveys allow for data collection solely in circumscribed areas [42]. On the contrary, remote sensing represents a fast and cheap solution to cover wide areas [45]. The representation of processed LiDAR data in a GIS environment provides a useful tool for analyzing, developing projects and monitoring, once the relationship between LiDAR measures and on-field inventory parameters have been calibrated [42]. One illustrative example is the case of New Zealand that developed a carbon inventory system relying on ALS acquisitions [25].

- *Estimating wood volume for timber and fuel production:* models that relate LiDAR data and wood volume or vegetation biomass may facilitate the quantification of potential timber production and wildfire hazard, becoming a useful tool for forest planning and management [40], [57].

- *Assessing the consequences of defoliation and insect infestations:* while usually defoliation is assessed by time-consuming on-field sampling, scholars are now proposing to use LiDAR technique as an innovative mapping and monitoring tool for estimating the damage caused by insect infestation. This assumption relies on the hypothesis that an increased proportion of returns corresponds to a more intense

[6] In this application emitted electromagnetic waves must be a blue or green laser in order to maximize the capability of the pulse to penetrate water; LiDAR system for topographic purposes usually operates with near-infrared signals that, conversely, are completely absorbed by water and cannot investigate underwater ground [32].

defoliation [64]. Also, defoliation can be estimated by comparing the values of leaf area index (LAI)[7] at different times, having defined LAI as dependent on the degree of pulse penetration through the canopy layer [55].

In this dissertation, solely the potentialities of LiDAR in mapping vegetation structures and estimating volume, or biomass, are analyzed. In the following section, a brief review of selected articles and studies concerning these purposes is reported.

4.3 A review of existing approaches to relate biomass to vegetation size

Biomass estimation through allometric equations has been widely studied and a large numbers of models have been tested. The basic principle is to correlate plant morphological parameters to volume, or biomass, by calibrating regression models whose data usually derive from on-field measurements. The first models were calibrated for small stands, providing site-specific equations that neither have a good fit on more extended areas, where the features of vegetation lose their homogeneity, nor can distinguish among tree components[8] [30], [58], [59]. Therefore, starting from the 1990s, new allometric models for large-scale estimation have been developed. One of the most representative examples is the series of equations set by Jenkins et al. in 2003 [30] and subsequently adopted by USDA Forest Service to develop the carbon budget of United States. Jenkins et al. aimed to create generalized equations that well fitted the vegetation features regardless its location across the country and that were consistent in terms of form, definition and required input data. To this purpose, they merged information from existing literature in order to set a database of the various models that had been calibrated on U.S. tree samples until that time. Collected models were organized according to sample taxonomy and geographic location and subsequently processed to return generalized equations for each species group. These equations allow to predict above-ground biomass, and the relative subdivision in tree components, and they require as input data solely DBH[9], while the other parameters are given once the tree species are identified. The authors highlighted that the choice of DBH instead of mean height as independent variable was due to the scarce accuracy in the measurements of this latter in case of closed-canopy stands. However, also this extrapolation of large-scale biomass data from allometry demonstrated to be affected by a high number of uncertainty. As Jenkins et al. reported, large variability in biomass values was due to noise in measurements and the

[7] Leaf area index (LAI): dimensionless parameter that characterizes the tree canopy. For broadleaf species, it is equal to the ratio between the one-sided green leaf area and the ground surface area (m^2/m^2) [55].

[8] Tree components: foliage, coarse roots, stem bark and stem wood [30].

[9] DBH: *Diameter at Breast Height*, is the diameter of the trunk of a standing tree, conventionally measured at 1.37 m from the ground.

small size of sample plots, while it seemed not to be related to regional vegetation patterns. Moreover, inconsistency in definitions and measuring methods among the various models could have affected the process of generalization.

Jenkins et al. generalized equation is:

$$AB_{\text{dry}} = exp(\beta_0 + \beta_1 ln(D)), \tag{4.1}$$

where AB_{dry} is above-ground dry biomass expressed in kg, D is DBH expressed in cm and β_0, β_1 are parameters that depend on tree species. In the case of Cinca River, β_0 and β_1 can be set as, respectively, -2.2094 and 2.3867[10]. The authors also provided an equation to calculate the contribution of each tree component to the total biomass:

$$ratio = exp\left(\beta_{0,\text{ ratio}} + \frac{\beta_{1,\text{ ratio}}}{D}\right) \tag{4.2}$$

Set of β_{i} values for *hardwood* are listed in Table 4.1.

Table 4.1. Value of parameters for determining the ratio of each tree component and total biomass [30].

Biomass component	$\beta_{\mathbf{0,\ ratio}}$	$\beta_{\mathbf{1,\ ratio}}$
Foliage	-4.0813	5.8816
Coarse root	-1.6911	0.8160
Stem bark	-2.0129	-1.6805
Stem wood	-0.3065	-5.5240

In the attempt to overcome the uncertainties affecting allometric large-scale estimations, the use of remote sensing techniques has started to be analyzed. The advantages related to remote sensing are mainly the required short time for data acquisition, the low-cost for investigations on wide areas and the independence from ground accessibility to the site. Despite agencies promoted satellite imagery [10], [2], [4], [5], [6], also airborne laser scanning has been largely developed, releasing a plethora of regression models for volume and biomass estimation. However, the objective of this dissertation was modelling riparian vegetation in temperate regions and, therefore, a large part of literature was neglected, as most of the studies were conducted in different climates, for example tropical regions, where the growth of vegetation follows different temporal and spatial scales. In the following paragraphs a short review of the proposed models is reported.

He et al. (2013) [29]

The authors developed equations for relating physical features of vegetation with biomass, since they stated that LiDAR could provide the most accurate measurements of ground

[10] Values given by the authors for *Hardwood - aspen/alder/cottonwood/willow*.

and canopy elevation, even though in conditions of extreme slope or high canopy closure. By attempting various combination of statistics inventory parameters with biomass, the authors found good correlations between mean first return elevation, or the 10^{th} and 90^{th} percentiles of first returns, and above-ground and stem biomass, while the first return above 2 m, or the 75^{th} and 90^{th} percentiles of first returns elevation, were correlated to crown biomass. In general, crown biomass showed a weaker correlation with LiDAR data. Also, they found that canopy cover could be related to biomass values. Although this paper focused solely on *Picea Crassifolia* species and its related sample area was located at a very high altitude, the model was accounted among the ones to be tested for riparian biomass estimation, because the operating conditions were very similar to those of the study site. In fact, the instrumentation consisted in a small footprint airborne laser scanning with low density (about 1 point·m^{-2}) and discrete returns. Moreover, the procedure to convert raw data into statics inventory parameters was similar to the one that was defined, in this dissertation, for riparian corridors (see Section 5.2). The proposed equation for above-ground biomass estimation is:

$$AB = -9.013 + 10.812\overline{h} + 25.105CC \tag{4.3}$$

AB is above-ground biomass expressed in Mg·ha^{-1}, $\overline{h}$ is the mean canopy height and CC is the canopy cover that ranges from 0 (i.e. no cover) to 1 (i.e. completed cover).

Lefsky et al. (2002) [34]

The authors' aim was to return a simple model for estimating above-ground biomass in temperate and boreal regions with airborne LiDAR acquisitions, as it would reduce the effort and expense of fieldwork for following investigations about biomass and carbon storage. Firstly, they defined one regression model for each of the analyzed biomes (i.e. boreal coniferous, temperate coniferous and temperate deciduous) by relating field measurements with LiDAR data and, finally, they combined the three models in one generalized relationships between above-ground biomass and LiDAR-derived mean canopy height:

$$AB = 0.378\overline{h}^2 \tag{4.4}$$

AB is above-ground biomass expressed in Mg·ha^{-1}, while $\overline{h}$ is the mean canopy height expressed in meters.

However, the authors processed large-footprint SLICER waveform data to define this equation and not specifics about its application to small-footprint and discrete-return data were given.

Li et al. (2008) [36]

By using data acquired by a high-density discrete return LiDAR equipment, the authors analyzed the possibility to develop LiDAR-based models for estimating above-ground biomass with a small set of LiDAR metrics that have an easy intelligible biological meaning, in contrast with the models offered by literature that were often made up of a

large number of variables, difficult to interpret in a physical perspective. Furthermore, they compared models that had been obtained by using different methods for selecting the variables, such as stepwise regression, Bayesian modelling average and principle component analysis.

As the authors studied three different forest types, they released four different sets of equations (i.e. one for each forest type and one generalized) containing relationships that had been defined by using each of the three selection methods. In this dissertation, only the generalized equations were tested, since all the forest types differed from typical riparian vegetation in temperate regions. Stepwise regression and Bayesian methods returned the same model:

$$LN(AB) = 5.49 + 0.42\overline{h} + 5.18h_{\text{cv}} - 00.66h_{50} + 0.66h_{75} - 0.3h_{90} + 2.98CC \tag{4.5}$$

While the principle component analysis gave two different equations:

$$LN(AB) = 11.23 - 0.42h_1 + 0.7h_2 - 0.64h_3 \tag{4.6}$$

$$LN(AB) = 5.64 + 0.11\overline{h} + 5.66h_{\text{cv}} + 3.14CC \tag{4.7}$$

AB is above-ground biomass expressed in kg·ha^{-1}, CC is the canopy cover, $\overline{h}$ is the mean canopy height, h_{cv} is the coefficient of variation of canopy height and the subscript numbers indicate the percentile of canopy height.

Means et al. (1999) [41]

The authors investigated the use of large-footprint scanning airborne LiDAR in mature forests, providing a model for above-ground biomass estimations that is valid if the considered stand has a biomass of at least 1000 Mg·ha^{-1}:

$$AB = 55 + 0.385\overline{h}^2 \tag{4.8}$$

AB is the above-ground biomass expressed in Mg·ha^{-1} and $\overline{h}$ is the mean canopy height expressed in m.

Means et al. (2000) [40]

According to the authors, LiDAR showed a potential for commercial applications in forest industry, allowing for the estimation of the main physical features of stands (i.e. height, basal area and volume) with reduced or null field work. This statement was confirmed by scatter-plots and regression analyses that returned relationships between small footprint LiDAR data and these features.

Even though large footprint LiDAR could provide reliable estimates, the authors focused on small footprint instruments because they were the ones commercially available at that time. The proposed height and stem wood volume equations are:

$$\overline{h} = 7.69 + 1.90h_{90} - 1.23h_{\text{max}} \tag{4.9}$$

$$ln(V) = 2.532 + 0.05651h_{80} + 2.355CC_{20} - 0.1581h_0 \tag{4.10}$$

considering the entire tree population and:

$$\overline{h} = 1.75 + 0.87h_{90} - 1.59h_0 \tag{4.11}$$

$$V = -83.4 + 16.29h_{80} + 9327CC_{90} \tag{4.12}$$

neglecting the old-growth stands. V is the volume expressed in $m^3 \cdot ha^{-1}$, h is the vegetation height expressed in meters and CC is the canopy cover. Subscript numbers indicate the percentiles and the maximum height.

Despite the outcome of these equations is the stem wood volume of a stand and the authors did not mention the intent of estimating the total biomass, fresh wood biomass can be calculated starting from them, if the fresh wood density is known. Obviously, this biomass could underestimate the real value as it does not comprise the contribution of foliage to the total amount. Values of fresh wood density ρ_{fresh} for various tree species can be directly found in literature or they can be calculated starting from literature values of dry density ρ_{dry} (i.e. wood-specific gravity) and typical moisture content θ [63]:

$$\rho_{\text{fresh}} = \rho_{\text{dry}} \cdot (1 + \theta) \tag{4.13}$$

Observations on selected models

Although numerous models have been created to estimate above-ground biomass in forest stands, no studies have been addressed to riparian corridors, where tree population is generally younger and sparser. Thus, one of the objectives of this dissertation was to test the reliability of these forest models in the case of their implementation for riparian vegetation. As no local field data of the Cinca River were available, the evaluation was done by comparing the obtained results one another and with literature data.

It is important to highlight that solely above-ground biomass was considered in this dissertation. Although the below-ground fraction can account for 20-26% of the whole biomass, methods for its estimation are poorly developed, being neither frequently implemented on-field nor standardized [50]. The use of LiDAR technology for this kind of estimations has been scarcely studied and limited literature currently refers to it, despite some attempts made with discrete return LiDAR showed a good correlation between LiDAR data and below-ground biomass, due to its strong relation to the size of trees, which it is easily measured with LiDAR [44].

A final consideration, regarding the LiDAR data that are commonly used for biomass estimation, must be done. Indeed LiDAR systems provide information about both elevation of hit surfaces with respect to the ground and intensity of the returning pulses, but the presented models concern solely regressions between elevation metrics and biomass. Return intensity, which is related to the ratio between the amount of energy detected by the receiver for a given reflection point and the amount of total energy emitted for the laser pulse, is rarely used in LiDAR analyses [24], [27], [38] and the few available studies involving return intensity, mainly focus on its potential for discerning surface

materials[11]. Song et al., for example, [56] analyzed the classification of urban elements, as asphalt, roofs, trees and grass by using LiDAR intensity data, though specifying that these usually are affected by high noise due to the variation of intensity response from target surfaces if the angle of reflection changes[12]. Other studies have been made to create a classification procedure to separate different land covers by using a combination of intensity and elevation data [14], [17], despite the reliability of intensity-based classification of vegetation and individual trees is debated [28].

In the paucity of literature about intensity-based biomass estimates, an illustrative example was provided by García et al. [27]. In the context of a mixed species Mediterranean forest, the authors showed how LiDAR could provide a measure of biomass by combining elevation and intensity statistics. According to them, intensity-related variables could better explain the variance of biomass and increase the accuracy of its estimation, especially when the study area is occupied by heterogeneous tree species. However, although the defined model was provided by the authors and it was developed for a Spanish forests, it was neglected in the development of this dissertation because of three main reasons:

- *Biological inconsistency:* the proposed model was regressed considering pine forests (*Pinus nigra Arn., Pinus sylvestris L. and Pinus Pinaster Ait.*) and oaks (*Quercus faginea Lam., Quercus ilex L. and Quercus pyrenaica Willd.*) that do not belong to the same genus of typical trees in riparian corridors.

- *Impossibility of a proper intensity calibration:* the reliability of this type of estimates depends on the calibration of intensity values that, in turn, usually relies on the knowledge of scan angle and/or altitude of the flying platform for each emitted pulse. In the context of this dissertation, solely the average height of the scanning flight was provided by CNIG thus preventing a proper calibration.

- *Inadequacy to the thesis requirements:* the proposed equation was considered too complex for the purposes of this dissertation. Indeed, the final goal of this review was not only to select a model for estimating biomass in riparian corridors, but also to choose a simple and linear procedure to define the probability distribution function of this biomass for a given altitude. García et al. model was inadequate to this purpose as the equation they proposed requires too many steps and leads solely to an average value for biomass, neglecting its probability distribution.

[11] Return intensity is related to the geometry of the acquisition (e.g. scan angle, distance source-target), but also to the density of foliage in tree canopies [27].

[12] The solution proposed by the authors was to normalize intensity data by the angle of reflection even though other scholars, as Hasegawa [28], stated that sometimes this correction is not applicable and that not all the surface can be classified according to intensity.

Chapter 5

Method

In order to process LiDAR data provided by *Centro Nacional de Información Geográfica*, both GIS and specific software for managing LiDAR data were used. QGIS, which is a free and open-source geographic information system [9], was chosen as GIS environment because of its easy-to-learn interface and the large variety of operations and plugins it offers, whilst FUSION/LDV software package was selected as instrument for handling LiDAR datasets.

Before defining a procedure to test the biomass estimation models with an embedded implementation of these software, a study about FUSION/LDV functioning and potentialities was required. At Section 5.1 a selection of FUSION/LDV operations is described in order to facilitate the comprehension of the procedure illustrated at Section 5.2.

5.1 LiDAR data processing with FUSION/LDV

FUSION/LDV is a system of software for the processing and visualization of LiDAR data. This software package is free and it has been developed by the Forest Service of US Department of Agriculture as a research tool for working with LiDAR data, whose size usually requires great computational efforts for commercial GIS software [38]. The package comprises:

- *Task-specific command-line programs:* running one or more of this programs allows for filtering, clipping and processing LiDAR data. Tailoring them is possible by rewriting the command-line, so that the final outcome can be customized according to the users' requirements.

- *FUSION software:* it provides the primary interface to display the outcome of the command-line programs.

- *LDV - LiDAR Data Viewer*: it provides an environment for three-dimensional visualization and analysis of data subsets.

Despite the large set of operations that the software package offers, just a sequence of few command lines was sufficient for areal biomass estimation. The following paragraph offers an insight into these basic operations, noting that specifications for each program are stated in the FUSION/LDV handbook [38].

Lda2Las

Raw LiDAR data usually are stored in *.las* format. However large datasets are compressed in *.laz* format and require a conversion prior to the import in the FUSION/LDV environment. Thus, the *Lda2Las* command-line program must be run:

```
Lda2Las [switches] InputFile OutputFile
```

InputFile is the *.laz* file, while *OutputFile* is the generated *.las* file[1]. *Switches* are optional commands to customize programs.

FilterData

Unless the data provider guarantees the quality of the datasets, it is recommended to filter data before processing them. The relative command-line program is *FilterData*:

```
FilterData [switches] FilterType FilterParms WindowSize
OutputFile DataFile
```

FilterType is the filtering algorithm [2], *WindowsSize* is the size of the window used to computed the standard deviation of elevation or the maximum and minimum values, *OutputFile* is the filtered *.las* file. *DataFile* is the *.txt* file that lists the original data file.

PolyClipData

Usually LiDAR data are stored in datasets that cover a square area (e.g. the data used in the case study were subdivided in square areas with a side 2 km long). If the aim is to analyzed solely a part of this square, it is recommend to clip the dataset with *PolyClipData*:

```
PolyClipData [switches] PolyFile OutputFile DataFile
```

InputFile is the original *.las* file, while *OutputFile* is the clipped *.las* file. *PolyFile* is the *.shp* file that defines the area of interest.

---

[1] the *.las* file constitutes the *raw data* in FUSION interface.

[2] FUSION supports four algorithms: *outlier* that removes returns above or below the mean elevation ± the product between the standard deviation of elevation and a user-defined constant; *outlier2* that is an upgrade of the former; *minimum* that removes all returns except the ones with minumum elevation; *maximum* that removes all returns except the ones with maximum elevation.

Catalog

FUSION does not display *.las* data. Thus, an associated *.bpm* grey scale intensity image[3] must be created with *Catalog*:

```
Catalog [switches] DataFile CatalogFile
```

DataFile is the *.txt* file containing the path of the *.las* file and *CatalogFile* is the name of the outcome. Useful switches in this case are:

- */image* creates an image files that shows the area covered by LiDAR data.
- */index* creates indices for LiDAR files.
- */coverage* creates an image that shows nominal coverage for all the datasets listed in the *.txt* file.
- */firstdensity:area,min,max* creates an image that shows the density of first returns in the area represented by each pixel. *area* is the pixel area, *min* and *max* define the interval of acceptable point density per unit area. If the first return densities of a cell is comprised between the minimum and maximum set values, the visualization colour is green, if it is lower red and if it is higher blue.
- */intensity:2,0,255* creates a grey scale intensity image using the average intensity for all first returns within each pixel. *area* is the pixel area, *min* and *max* are, respectively, the minimum and maximum intensity value.

GroundFilter

Once the data have been clipped and filtered, and the grey scale intensity image has been created, the ground points can be extracted from the 3D cloud with *GroundFilter*:

```
GroundFilter [switches] OutputFile CellSize DataFile
```

OutputFile is the *.lda* file containing the ground points, *Cellsize* is the size of the cell that is used to select the ground points and create the intermediate surface. It may differ from the size of cell in the final DTM. *DataFile* is the *.txt* file that lists the *.las* input file. Usually the following switches are used:

- */gparam:#*, */wparam:#*, */aparam:#*, */bparam:#* are the parameters used in the weight equation (**??**). Default values are respectively: -2.0, 2.5, 1.0, 4.0.
- */tolerance:#* is the tolerance value for the final filtering of the ground points, meaning that only the points beyond # units from the final intermediate surface model are excluded from the output file. The default value corresponds to the weight.

[3] the *.bpm* file must be associated to *image* command in FUSION interface.

- */iteration:#* is the number of iterations for the filtering process. The default value is 5.

- */lda* forces the output file to be written in *.lda* format even when the input is a *.las* file.

The filtering algorithm for creating the intermediate surface comprises a system of three equations and it has been adapted from Kraus & Pfeifer [33] by FUSION/LDV developers. It consists of an iterative process: the outcome of the first stage, that is a surface that lies between the true ground and the canopy surface, is computed considering equal weights for all the points of the 3D cloud and becomes the input for the subsequent computation. Starting from the second step, weights are defined by Kraus & Pfeifer system of equations:

$$p_{\mathrm{i}} = 1 \qquad for\ v_{\mathrm{i}} \leq g \tag{5.1}$$

$$p_{\mathrm{i}} = \frac{1}{1 + a(v_{\mathrm{i}} - g)^b} \quad \text{for } g < v_{\mathrm{i}} \leq (g + w) \tag{5.2}$$

$$p_{\mathrm{i}} = 0 \qquad for\ (g + w) < v_{\mathrm{i}} \tag{5.3}$$

If v_i is the vertical distance between the i point and the computed surface, ground points are expected to be under this surface, while canopy point above it. The parameters a and b determine the steepness of the weight function, g indicates which points are related to the maximum weight (i.e. weight 1 is assigned to points located below the surface by more than g along the vertical direction) and w defines the upper limit for point to have an influence on the ground surface modelling; all the points above the surface by more than w have a weight equal to zero. After the fixed number of iteration is run, bare-earth points are used to create the final intermediate surface. These points are those that satisfy the conditions set by Equation 5.1 and Equation 5.2, unless the */tolerance:#* switch is used, since in that case all the points within the specified tolerance distance from the final surface are considered bare-earth points and used for the surface creation.

GridSurfaceCreate

GridSurfaceCreate command-line program reproduces the 3D model of ground surface in *.dtm* format[4]:

```
GridSurfaceCreate [switches] SurfaceFile CellSize XYunits Zunits
CoordSys Zone HorizDatum VertDatum DataFile
```

[4]In FUSION environment, the *.dtm* file of ground surface must be associated to the label *Bare earth*.

SurfaceFile is the output *.dtm* file, while *CellSize* is the size of the grid cell of the DTM. *XYunits* and *Zunits* must be specified as *M* for meters or *F* for feet. *CoordSys* is the coordinate system for the surface, defined as 0 if unknown, 1 for UTM, 2 for state plane, and *Zone* is the corresponding zone for the surface in the chosen coordinate system, defined 0 if unknown. *Horizdatum* is the horizontal datum for the surface, defined as 0 if unknown, 1 for NAD27, 2 for NAD83 and *VertDatum* is the vertical datum for the surface, defined as 0 for unknown, 1 for NGVD29, 2 for NAVD88, 3 for GRS80. *DataFile* is the *.txt* file containing the path of the input *.las* file.

CanopyModel

DTM is not the solely elevation model that LiDAR data processing can produce. The *CanopyModel* command-line programs allows users to create 3D surface models that reproduce the shape of canopy[5]. By modifying the command line, two different types of surface models can be realized:

- *Canopy Surface Model (CSM)*: 3D elevation models whose z-coordinates refer to elevation from sea level.
- *Canopy Height Model (CHM)*: 3D elevation models having the same shape of the relative CSMs, but whose z-coordinates refer to elevation from the ground surface.

For obtaining the CSM, the relative command line is:

```
CanopyModel [switches] SurfaceFile CellSize XYunits Zunits
CoordSys Zone HorizDatum VertDatum DataFile
```

The meaning of the various elements of the command line is the same of *GridSurface-Create*. If the desired output is the CHM, the switch */ground:GroundSurfaceFile.dtm* must be added.

Dtm2Ascii

In order to visualize the surface models created with *GridSurfaceCreate* and *Canopy-Model* in GIS enviroment, *Dtm2Ascii* must be used:

```
Dtm2Ascii [switches] InputFile OutputFile
```

InputFile is the *.dtm* file to convert, while *OutputFile* is the *.asc* file to import in GIS environment. Usually the */raster* switch is used; it interprets the DTM points as the attribute for a cell, adjusting the origin of the ASCII grid file so that the lower left data point is the center of the lower left grid cell.

[5] In FUSION environment, the *.dtm* file of canopy surface must be associated to the label *Canopy*.

GridMetrics

To extract statistic information about elevation and intensity of the returns, *GridMetrics* program may be used. It splits the covered area according to a grid and computes descriptive statistic parameters[6] for the data contained in each cell. The output is a ASCII text file in the *.csv* format whose rows represent one cell and columns are the computed parameters. The sintax of the command line is:

```
GridMetrics [switches] GroundFile HeightBreak CellSize
OutputFile DataFile
```

GroundFile is the *.dtm* file representing the computed DTM, *HeightBreak* is the elevation value considered for cover calculation [7], *Cellsize* is the size of the cell that is used to compute the statistics, *OutputFile* is the *.csv* file with the results and *DataFile* is the *.txt* file that lists the *.las* input file. Useful switches may be:

- */outlier:min,max* defines the lower and upper limits of the elevation interval outside which points are considered as outliers.
- */strata[x1,x2,...,xn]* counts the returns in different height strata. *x1,x2,..,xn* represent the limit of these strata.

Csv2Grid

Finally, to convert *.csv* data in a ASCII format to be visualized in GIS, *CSV2Grid* may be used:

```
Csv2Grid [switches] InputFile Column OutputFile
```

InputFile is the *.csv* file, *Column* is the parameter to be visualized (e.g. mean elevation) and *OutputFile* is the outcoming raster.

5.2 Testing and evaluation of biomass estimation models

As literature offers models to estimate forest biomass, FUSION/LDV software package and GIS were used to test them in the study area and assess their reliability for riparian vegetation. Firstly, the whole study area was considered and models were used

[6] *GridMetrics* returns the position of the cell in the grid, the elevation (or intensity) maximum, minimum, mean, mode, standard deviation, variance, coefficient of variation, interquantile range, skewness, kurtosis, the L-moments and the 1^{st}, 5^{th}, 10^{th}, 20^{th}, 25^{th}, 30^{th}, 40^{th}, 50^{th}, 60^{th}, 70^{th}, 75^{th}, 80^{th}, 90^{th}, 95^{th}, 99^{th} percentiles. Moreover the elevation statistics comprise the number of the various returns (i.e. first, second, etc.) above the heightbreak.

[7] Cover for a specific area is calculated as the ratio between the number of returns above the heightbreak and the total returns. Heightbreak is an arbitrary value, but it is commonly set equal to 1.37 that corresponds to the height, in cm, at which the DBH is measured.

to estimate biomass along all the segment and in the relative floodplain. Models were compared each other in order to select the ones that returned consistent results and the influence of chosen cell size was assessed. Secondly, the consistency of the selected models was tested comparing LiDAR-estimated biomass with the value obtained by allometric formulas, in those areas where vegetation could have grown undisturbed. The selection of this area was done according to the comparison of various orthophotos over time, in order to understand where vegetation growth had not been affected by flow fluctuations.

Creation of DTM

The *.laz* compressed datasets, related to the study area, were converted in a comprehensive *.las* file by running *Lda2Las* and the corresponding grey scale image was obtained with *Catalog*. Importing the grey scale images in GIS environment, the exact shape of the study area was selected and saved in a *.shp* file to be used to clip the dataset with *PolyClipData*[8]. Since the moderate level of anthropization, the selection of the floodplain was done also according to the artificial boundaries, such as high traffic roads, facilities and crop fields.

Secondly, *GroundFilter* was run; *g* and *w* parameters were set respectively equal to -1 and 2.0, while the tolerance was fixed as 0.1. Various numbers of iterations were tried, but no relevant differences were noticed in the outcomes. A final number of 20 was decided to find a balance between precision and short computational time. *CellSize* was defined equal to 2, as FUSION/LDV developers suggests to set the same size of LiDAR acquisition [38].

Finally, the DTM was created with *GridSurfaceCreate*. Before running it, the command line had been modified by defining:

- *CellSize* = 2, equal to the units of LiDAR data in the study area [1].
- *XYunits*, *XYunits* = *M* meaning meters.
- *CoordSys* = 1, as the coordinate system is UTM.
- *Zone* = 31 N, which is the UTM zone number of the study area.
- *HorizDatum*, *VertDatum* = 0, as both the horizontal datum and vertical datum were unknown.

In this case, the *FilterData* operation was neglected, since the source of the data explicits that they had been already preprocessed [1].

[8]Clipping the shape of the study area was necessary as each dataset covers a square area 2 km x 2 km, including not only the riverbed and floodplain, but also the limiting area that are not object of this dissertation.

Biomass estimation

The command line *GridMetrics* was run various times to obtain elevation statistics by discretizing the study area in square cells with side 2, 10 and 20 m long. Cell sizes were defined according to the procedures adopted by the authors of the selected models, that are summarized in Table 5.1.

Table 5.1. Selected models for biomass estimation. Overlined characters indicate averaged values, while subscript numbers indicate percentiles.

Author	**Model**	**Eq.Ref.**
He et al. (2013)	AB= -9.013+10.812$\overline{h}$ + 25.105CC	4.3
Lefsky et al.(2002)	AB = 0.378$\overline{h}^2$	4.4
Li et al.(2008)	ln(AB)=5.49+0.42$\overline{h}$+5.18h_{cv}-00.66h$_{50}$+0.66h$_{75}$-0.3h$_{90}$+2.98CC	4.5
	ln(AB)=11.23-0.42h$_1$+0.7h$_2$-0.64h$_3$	4.6
	ln(AB)=5.64+0.11$\overline{h}$+5.66h$_{cv}$+3.14CC	4.7
Means et al. (1999)	AB = 55+0.385$\overline{h}^2$	4.8
Means et al. (2000)	ln(V) = 2.532+0.05651h$_{80}$+2.355CC$_{20}$-0.1581h$_0$	4.10
	V = -83.4+16.29h$_{80}$+9327CC$_{90}$	4.12

Equation 4.3 was implemented on 20x20 m cells [29], while Equation 4.10 and Equation 4.12 on 10x10 m cells [40]. As the other models worked on the whole area [34], [36], [41], the various discretizations were compared. Also Equation 4.1 (i.e. Jenkins et al.) was implemented as it represented a useful instrument to have an order of magnitude of above-ground biomass, despite the authors declared that this generalized equation might be affected by uncertainties and it had to be used carefully in site-specific applications [30]. This model calculated dry above-ground biomass AB_{dry}, and therefore solely the dry wood amount was accounted and converted in fresh biomass by considering the multiplying factor $(1+\theta)$, where θ is the moisture content of wood. The ratio of dry wood and total dry biomass was computed with Equation 4.2.

Following a similar procedure, also biomass in two sub-areas was evaluated. The first sub-area consisted of one[9] of the nine datasets downloaded from CNIG [1] for the year 2016 and it was used to analyze the behaviour of estimation model when varying cell size; the reduction of study area for this analysis was done with the purpose of shortening computational time. The second sub-area was a small plot, shown in Figure 6.6, defined by observing orthophotos taken over time and choosing one spot where vegetation had grown without being affected by river flow variability. In this spot, vegetation is expected to have reached its maximum growth[10] and, therefore, the value of biomass provided by the chosen model should be comparable with the one returned by the allometric formula of Jenkins et al. that, as explained in the following sections, does not account for flood

[9]Dataset name: PNOA_2016_ARA_270-4610_ORT-000-COL.

[10]If vegetation has reached its maximum size, its biomass value represents the carrying capacity for that species in those specific conditions (e.g. water and nutrients availability, mechanic stress).

uprooting or excessive exposure.

Error metrics

For each implemented model, biomass values were computed for the three chosen cell sizes (i.e. 2x2, 10x10 and 20x20 m). In those cases the authors had specified cell size (He et al. 2013 and Means et al. 2000) solely the results relative to that discretization were reported; in the other cases, results were presented as the average value among the various discretizations ± the standard deviation.

The analysis of the influence of cell size on results was carried out in two different ways. If a specific cell size had been indicated, the mismatch among the reference value (i.e. the biomass computed for the specific cell size) and the values obtained by varying discretization was measured with three error metrics and percentage change. Otherwise, general considerations on differences among results were done. Percentage change was computed as:

$$V = \frac{AB_{\text{final}}}{AB_{\text{initial}}} \cdot 100 - 100, \tag{5.4}$$

where negative values of V indicate lower estimates, while positive higher estimates.

The chosen error metrics were:

- *Mean Bias Error (MBE):* it provides a measure of the magnitude of the errors in the set of values; due to its linearity, it weights at the same way all the individual differences, even though the negative and positive errors could compensate each other returning and underestimate the error.

$$MBE = \frac{1}{N} \cdot \sum_{i=1}^{N} AB_{\text{i}} - AB_{\text{ref}} \tag{5.5}$$

  N indicates the total amount of i-th values AB_{i} to compare with the reference one AB_{ref}. Negative MAE values indicate that on overall the computed AB are smaller than the reference value.

- *Mean Absolute Error (MAE):* similarly to MBE it weights at the same way all the individual differences, but, conversely to MBE, the presence of absolute value prevents considerations about the direction of the error and avoids the measure of the errors to be underestimated.

$$MAE = \frac{1}{N} \cdot \sum_{i=1}^{N} |AB_{\text{i}} - AB_{\text{ref}}| \tag{5.6}$$

  MAE values ranges from 0 to ∞; lower MAEs correspond to smaller errors and higher accuracy.

- *Root Mean Square Error (RMSE):* as errors are squared before the average calculation, RMSE associates higher weights to large errors. Square differences avoid considerations about the direction of the error.

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (AB_{\mathrm{i}} - AB_{\mathrm{ref}})^2} \tag{5.7}$$

  As MAE, $RMSE$ values ranges from 0 to ∞ and accuracy increases when $RMSE$ decreases.

Although $RMSE$ is one of the most reported metrics [67], used especially when large errors must be avoided, in this dissertation also the other two metrics were calculated as their comparison could lead to interesting considerations. MBE was the only metrics, among the chosen three, that indicated the direction of the error. MAE was compared to $RMSE$ in order to understand the variability of occurred errors: the more $RMSE$ exceeded MAE, the less errors were distributed homogeneously among the measures.

Chapter 6

Results and discussion

6.1 Implementation of selected models

The DTM of the study area that was created with FUSION/LDV, starting from available LiDAR data, is shown in Figure 6.1. As the floodplain ranges from 0 to approximately 5.4 meters in relation to the water surface at the time of the acquisition, the DTM was subdivided in 54 altitude bands in order to facilitate the visualization in GIS, setting tbe water surface as the lowest band. The distribution of vegetation in the study area is shown in Figure 6.2, where the mean vegetation height, computed for the finest cell grids (i.e. 2x2 m), is displayed. The selected models were implemented and evaluated both individually, according to literature values and allometric formulas, and in comparison each other.

Jenkins et al. 2003

Equation 4.1 required as input data tree-specific parameters β_0, β_1, whose values for the study area were found in literature [30], and DBH. As processing of LiDAR data provided the average vegetation height, Equation 8.7, was used to find the relative DBH and compute biomass[1]. Also the moisture content θ was set for this equivalent vegetation, with a value equal to 1.055. The resulting above-ground dry biomass was, on average, 43.4 $\pm$ 0.5 Mg·ha^{-1}. Considering the sole wood contribution (e.g stem and branches) and converting it in fresh biomass, the resulting value was 86.7 $\pm$ 1.1 Mg·ha^{-1}

He et al. 2013

Equation 4.3 was implemented on 20x20 m cells and the resulting above-ground biomass was 41.4 Mg·ha^{-1}. This value was perfectly aligned with the interval proposed by the

[1]The tree-specific parameters of Equation 8.7 were set considering average values between poplars and willows. As explained at Section 8.1, an equivalent vegetation with intermediate characteristics between these two species was considered for the study area.

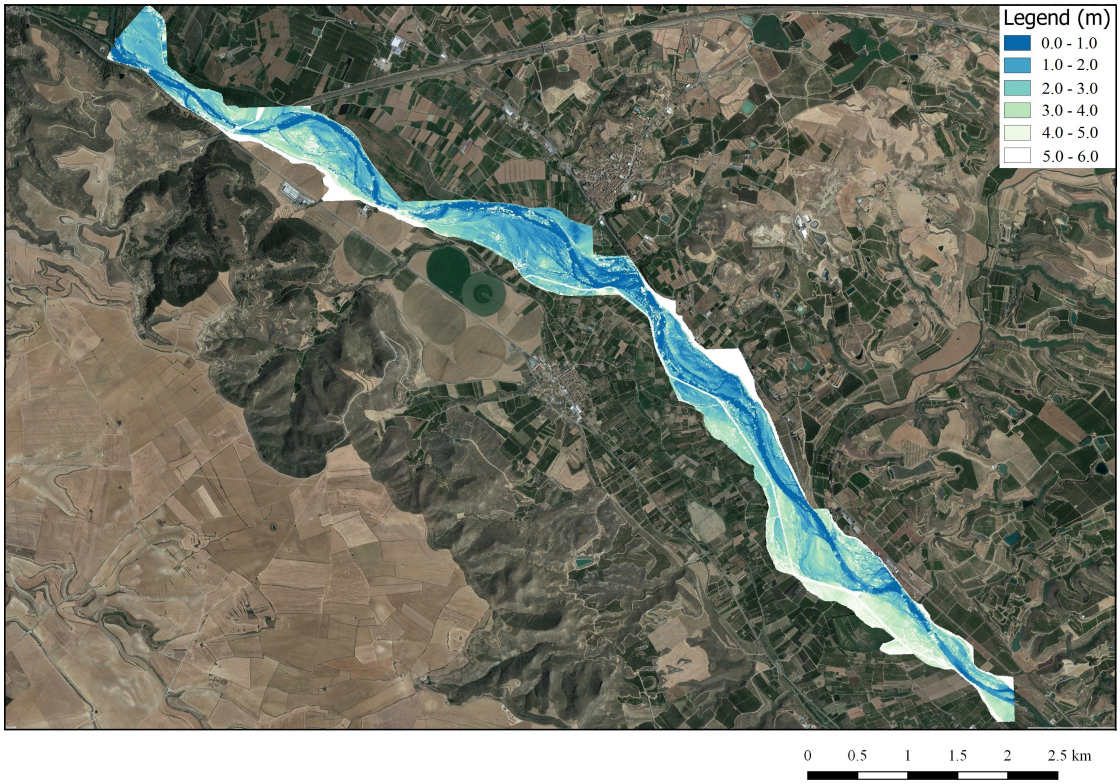

Figure 6.1. DTM of the study area. Legend values indicate the topographic height in relation to the mean water surface.

authors [29], according to whom above-ground biomass ranges from 0 to 230 Mg·ha^{-1}. The computed value could be valid for the shrubs that occupied the study area.

Lefsky et al. 2002

As the authors computed above-ground biomass considering the average height of vegetation in the entire study area [34], Equation 4.4 was implemented experimenting various configurations; for each configuration the average value among the three different cell sizes was considered as each discretization led to a different values for $\overline{h}$. Firstly, a biomass value of 7.9 $\pm$ 1.3 Mg·ha^{-1} was obtained defining a value of $\overline{h}$ for each cell. Secondly, $\overline{h}$ was set as the average value of all the study area and the resulting biomass was 3.6 $\pm$ 1.0 Mg·ha^{-1}.

The results of the two configurations were less than the minimum proposed by the authors for deciduous forests in temperate climates, 11.4 Mg·ha^{-1} [34]. Nevertheless, the authors did not specify the age of the young forest they considered and differences in results may be explained by high renovation rate of riparian vegetation due to continuous uprooting caused by floods. In addition to this, it must be considered that the models

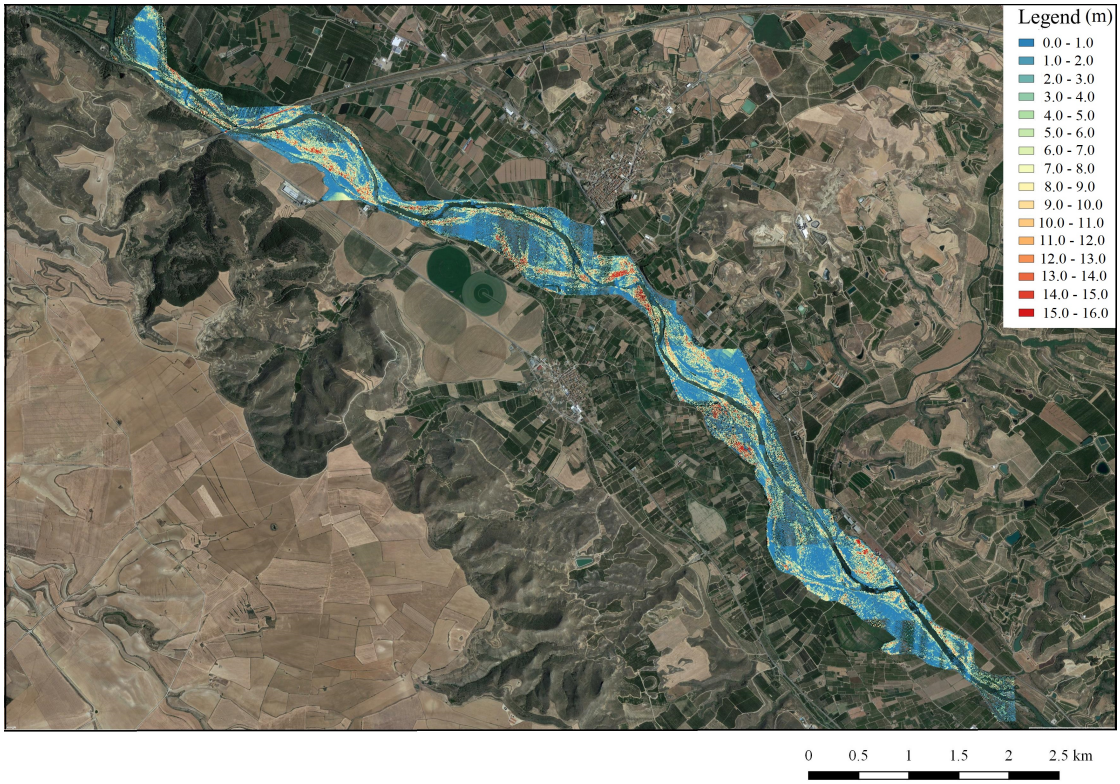

Figure 6.2. Vegetation coverage of the study area. Legend values indicate the vegetation height.

were created for *full waveform recording* LiDAR systems and no specifications about its implementation for *discrete return* were given by the authors.

Li et al. 2008

All the three models, Equation 4.5, Equation 4.6 and Equation 4.7, gave unrealistic results. One cause of this discrepancies could have been the principles on which the models relied; the authors, in fact, did not consider the canopy cover, or rather the ratio between first returns above heightbreak and total first returns, as a independent variable, but filtered the data to obtain only the return above 2 m [36].

As an attempt, *GridMetrics* was run other six times, three filtering values above 1.3 m (2x2, 10x10 and 20x20 m cell grids), which is the reference breast height, and three filtering above 2 m (2x2, 10x10 and 20x20m cell grids). However, also in those cases the results were neither consistent each other nor realistic.

Means et al. 1999

Similarly to Lefsky et al., this model estimates biomass starting from the average vegetation height of the whole study area. Therefore, also Equation 4.8 was tested for different configurations, both changing the cell size and considering the average height for individual cells and the entire area.

Despite this model referred to large-footprint scanning airborne LiDAR, it returned consistent values for all the combinations. The resulting biomass was 63.1 ± 1.3 Mg·ha^{-1} considering $\overline{h}$ for each cell and 58.7 ± 1.0 Mg·ha^{-1} considering a $\overline{h}$ for all the area.

Means et al. 2000

Before implementing the proposed models, an equivalent fresh wood density for the mixed riparian vegetation was defined. As the study area is mostly occupied by willows (ρ_{fresh} = 0.82 Mg·m^{-3}) and poplars (ρ_{fresh} = 0.66 Mg·m^{-3}) , the chosen density was an average value, equal to 0.74 Mg·m^{-3}. Models were first applied to 10x10 m cells, as specified by the authors [40], and subsequently tested for 2x2 and 20x20 cell size in order to evaluate the influence of grid size on results.

Equation 4.10 returned 58.6 Mg·ha^{-1} for 10x10 m cells, while Equation 4.12 returned 2736.9 Mg·m^3, a value that could be referred to very mature forests and it is not realistic in the case of riparian corridors.

6.2 Comparison among models

As the models provided by Li et al. 2003 (Equations 4.5, 4.6 and 4.7) and the second model of Means et al. 2000 (Equation 4.12) did not seem reliable for riparian biomass estimation through small footprint discrete LiDAR, they were discarded from the shortlist.

The remaining models were compared in order to understand their mutual consistency. The value provided by Jenkins et al. allometric formula (i.e. approximately 87 Mg·ha^{-1}) was considered as a reference for the order of magnitude of biomass in the study area. He et al. 2013 (Equation 4.3) underestimated biomass, probably because it was regressed for a specific tree species that do not belong to temperate riparian vegetation; Lefsky et al. 2003 (Equation 4.4) strongly underestimated biomass presumably because the models were regressed for full waveform-recording LiDAR data; Means et al. 1999 (Equation 4.8) and Means et al. 2000 (Equation 4.10) led to consistent results, especially when Equation 4.8 was computed with an average value of height for all the area, regardless the average values returned for each discrete cell. It must be highlighted that Equation 4.8 returns a biomass equal to 55 Mg·ha^{-1} when average height is equal to zero (i.e. it considers that a certain amount of biomass is always present, even when there are no trees) and this could lead to mismatches with Equation 4.10. Assuming, for example, no crown cover (CC=0), Equation 4.10 gives values lower than 55 Mg·m^3 in all those cases that h_{80}<15.4 m, which are vegetation heights likely to occur in riparian

corridors. Hence, Equation 4.8 is not reliable in those cases, where vegetation in proximity of the water is usually young and small size. However, increasing CC, the height threshold decreases and the range of applicability of Equation 4.8 enlarges. In Table 6.1, the explained results are reported.

Biomass values provided by Equation 4.10 were assessed as consistent also with the one computed through Jenkins et al. generalized equation. In fact, even though the former seemed an underestimation, it was considered that:

- Equation 4.10 returned volume values that were transformed in biomass through fresh wood density; hence, the foliage biomass, which accounts for around 3% of total biomass for the study case [30], was neglected.
- Equation 4.1 could provide only indicative values for small-scale areas and, above all, that it does not considered the vegetation removal action of rivers.

Table 6.1. Above-ground biomass computed values (Mg·ha^{-1}) for the study area.

Cell size	**He et al. 2013**	**Lefsky et al. 2003** cells average height	**Lefsky et al. 2003** areal average height	**Means et al. 1999** cells average height	**Means et al. 1999** areal average height	**Means et al. 2000**
	Eq. 4.3	**Eq. 4.4**	**Eq. 4.4**	**Eq. 4.8**	**Eq. 4.8**	**Eq. 4.10**
2	33.9	8.9	3.4	64.1	58.5	57.6
10	30.9	6.4	2.8	61.5	57.8	58.6
20	41.4	8.5	4.7	63.6	59.7	68.3

6.3 Influence of grid discretization

The analysis of the influence that grid discretization could have on results was carried out solely for the two most relevant models, or rather Equation 4.8, considering an areal average for $\overline{h}$, and Equation 4.10. Anyway, the choice of cell size for grid discretization seemed not to greatly affect the results, as shown in Figure 6.3. In the case of Equation 4.10, the results for 10x10 cell size were considered as reference values, since 10x10 m is the grid size used by the authors [40]; errors were generally overestimation for coarser discretization, while slightly and negligible underestimation for smaller cells. With respect to Equation 4.8, not a similar observation may have been done as the authors did not computed biomass by discretizing the area in cells, but they considered the whole territory, and not a cell size can be fixed as reference value in comparison to which evaluating eventual mismatches among results; moreover, variations of results for changing cell size are less than 2 Mg·ha^{-1} and could, therefore, be neglected.

When considering the whole study area, the choice of grid size slightly influenced results for biomass estimation. However, it was necessary to verify if discretization represented a discriminating factor when the estimation had to be done for very narrow and complex areas, as in the case of splitting the study segment according to topographic bands (see Section 8). In that case, in fact, it could happen that areas were less wide than

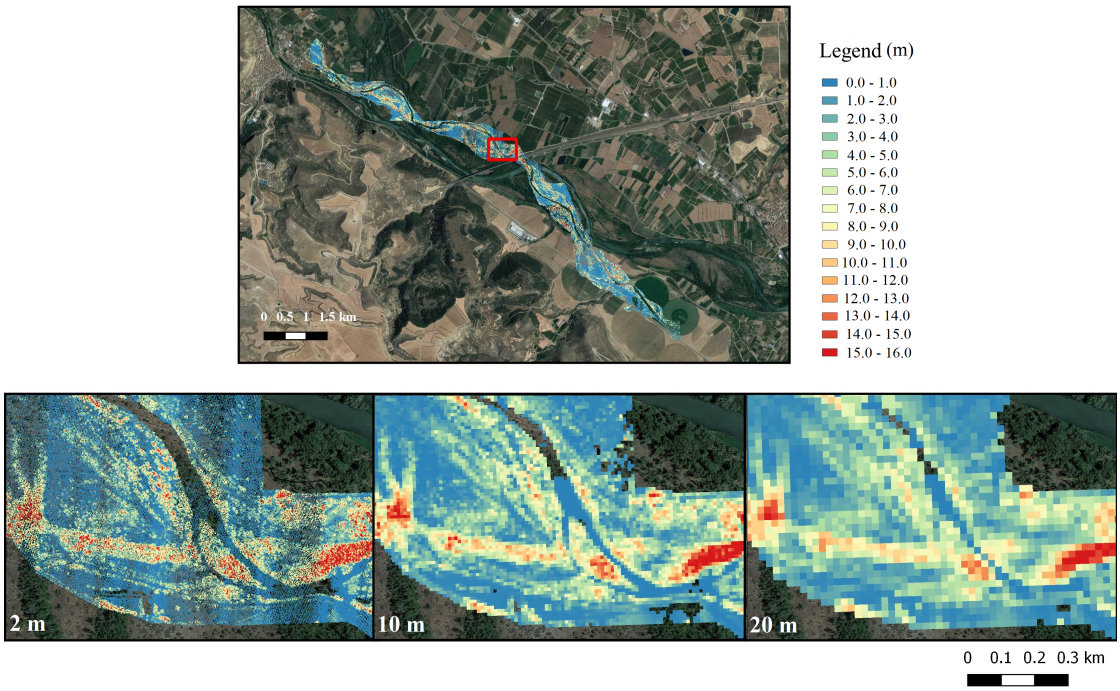

Figure 6.3. Comparison among different cell size for the same selected spot.

the cell size and, according to GIS visualization, it seemed that larger cells tended to overcome the boundaries of the selected band, computing biomass also for the adjacent territories, as visible in Figure 6.4. Thus, a subplot of the study area was taken as reference site to analyze the effect of cell size on biomass estimation and verify that the bias shown in Figure 6.4 was solely related to visualization and not to numeric results. As reported in Figure 6.5, all the cell sizes led to similar biomass trends, except for Equation 4.10 with 2 m cells that tended to overestimate biomass of around 40% in comparison to the reference value (i.e. Equation 4.10 with 10 m cells). Furthermore, *GridMetrics* computed for 2 m cells did not recorded vegetation returns for certain altitude bands. 20 m cells slightly underestimated biomass, with values that are usually 5% lower than the reference ones, while, as stated at Section 6.2, Equation 4.8 led to consistent results with the reference values (underestimation of 2%), unless the biomass is less than 55 $Mg \cdot ha^{-1}$, as happened in the study area above altitudes of 5 m in relation to the river bed.

An overview about error metrics for the various discretizations, assuming Equation 4.10 as reference, is reported in Table 6.2. Metrics confirmed what the percentage change had already highlighted. Equation 4.10 with 2 m cells tended to overestimate biomass (MBE is positive), with large errors; the same equation but with 20x20 m cells, had a good match with reference values, only slightly underestimating biomass (MBE is

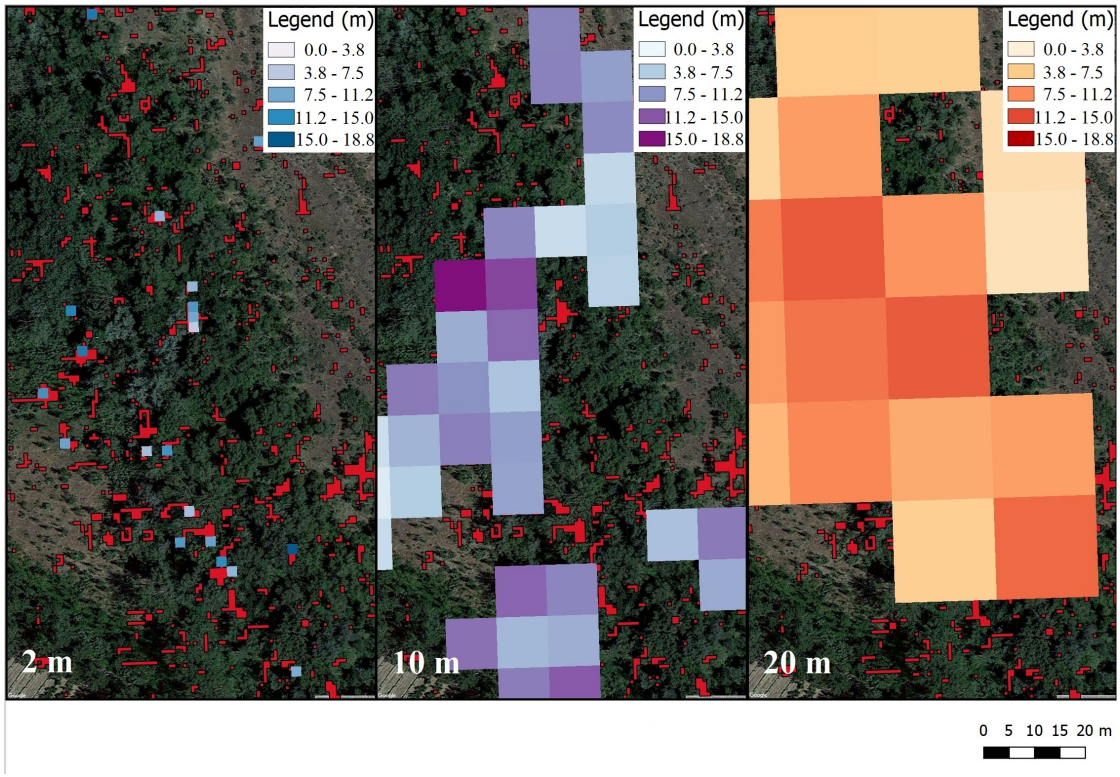

Figure 6.4. Comparison among statics about vegetation height, for a specific altitude band (in red), computed with different cell discretizations.

negative), despite errors were not homogeneously distributed ($RMSE$ is almost the double of MAE). The similarity between Equation 4.8 and the reference case was high and led to the best error metrics when the sole altitude bands below 5 m in relation to the river bed were considered, while it was unacceptable when the entire altitude spectrum was taken in account.

Table 6.2. Error metrics for varying cell size. The second row represents the case in which cell size is 2x2 m for Equation 4.8, but results above 5 m in altitude are excluded because biomass is less than 55 Mg·ha^{-1}.

Cell size	MBE	MAE	RMSE
2 (Eq. 4.8)	-7.2	11.0	19.2
2 (Eq. 4.8)	-2.7	4.9	5.8
2 (Eq. 4.10)	15.1	28.8	94.5
20 (Eq. 4.10)	-3.1	4.8	7.3

In summary, the chosen model for defining the distribution of biomass in the study

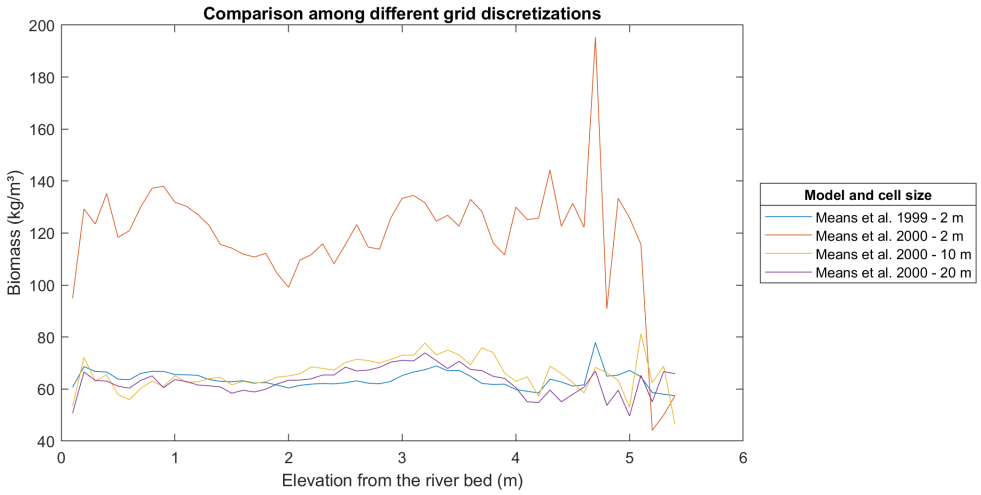

Figure 6.5. Comparison among biomass trends for different cell sizes.

area was Means et al. 2000 (i.e. Equation 4.10, associated to the discretization of the area in 10 m cells). This represented a good compromise between reliability of results and computational time. Although Equation 4.8 for 2x2 m cells returned consistent results with Equation 4.10, it was discarded because of the lack of results for specific altitude bands, the incapacity of recording biomass value lower than 55 Mg·ha^{-1} and the long computational time it would have entailed.

6.4 Comparison of computed carrying capacity

Implementing the chosen model (i.e. Means et al. 2000, Equation 4.10, with 10x10 m cells) on a small plot where vegetation is expected to have grown undisturbed, thus achieving a biomass value close to its carrying capacity, the reliability of the model was assessed. Indeed, the biomass the model returned had to be comparable to the value provided by the consolidated allometric model of Jenkins et al. (Equation 4.1), which does not take into account the variability of conditions of riparian environments.

The selected spot is shown in Figure 6.6. Jenkins et al. formula was implemented by considering the average vegetation height deriving by LiDAR acquisition and returned a biomass value of 138.6 Mg·ha^{-1} that represents the maximum biomass of the specific tree population in the study site when reaching that height and that could be considered as a local carrying capacity. As Equation 4.10 returned a biomass value of 112.4 Mg·ha^{-1} and, therefore, was consistent with Jenkins et al., the model was assessed as sufficiently reliable and used to achieve the second objective of this dissertation.

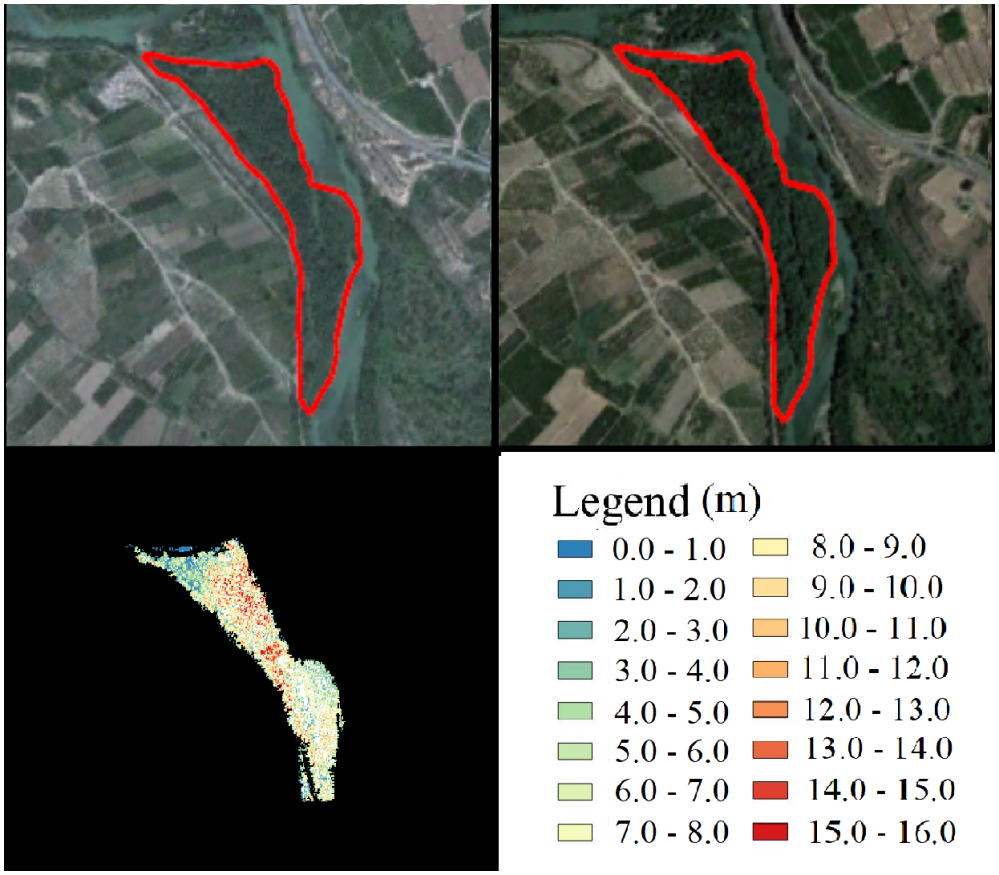

Figure 6.6. Selected area for comparing the biomass computed with the selected model and allometric formulas.

Part II

Stochastic model calibration

Chapter 7

Modelling riparian vegetation dynamics

In 2006, Camporeale and Ridolfi [16], formulated a stochastic process describing the distribution of phreatophyte riparian vegetation by modelling the impact that the randomness of hydrological fluctuations has on its growth. One of the two objectives of this dissertation was to define a procedure for calibrating this model, starting from hydrological data and the outcomes of LiDAR acquisitions that were available for the study area.

Calibration focused on one parameter that still needed to be properly set, k, which represents the ratio between the rate of vegetation decay during inundations and the rate of vegetation growth during exposure, and it required a set of parameters and input data. The former were mainly found in literature, while the latter consisted of geometric, hydrometric and biological site-specific information and were collected in different ways:

- *Hydrometric data:* time series of river discharges, which had been provided by *Confederación Hidrográfica del Ebro* [3], were processed in order to obtain statistics about water levels.
- *Biological data:* the required information concerning riparian vegetation biomass were obtained by processing raw LiDAR data previously downloaded from CNIG online platform [1].
- *Geometry*: the topography of the study area was defined by the outcomes of LiDAR data processing, while the configuration of the phreatic surface was assumed as horizontal.

In summary, the software used for the calibration were FUSION/LDV, QGIS, HEC-RAS and MATLAB. As explained in Part III, vegetation data were processed with FUSION/LDV, while geometric data also required the use of QGIS. HEC-RAS allowed to define the hydrological rating curves. All the computations were made by creating scripts in MATLAB.

7.1 Processes and factors influencing riparian vegetation evolution

Considering the riparian environment as constituted by three components, namely the stream, the topographic site and the riparian vegetation, river dynamics can be explained as an ensemble of processes driven by the interactions occurring among the elements of this triad. As a matter of fact, each of these components continuously exerts disturbances to the others so that they mutually influence their behaviour. In response to these disturbances, the riparian environment evolves in time and space, forced to shift from the existing status towards a new equilibrium. Disturbances mainly consist of river flow fluctuations, geomorphological processes, as erosion and sedimentation, and the action of the colonizing vegetation and they cannot be deterministically predicted because of the randomness of their occurrence and their effects [65].

Assuming a vegetation-centric perspective, riparian vegetation is strongly controlled by geomorphological processes and river flow fluctuations in each phase of its life cycle [16], [52], [65], though its mutually feedback actions on riparian morphology and the local hydrodynamic field. Regarding the initial phase of vegetation life, geomorphological processes contribute to the creation of an adequate site for the establishment of pioneer vegetation, since the deposition of fresh alluvial substrate leads to optimum conditions for seed germination, but they could also prevent vegetation to born because of burial and scouring. Flow regime, or rather the combination of stream velocities and water stages, creates room for new vegetation, too, as high flows can remove pre-existing vegetation (i.e. uprooting due to flooding). Moreover, flow regime influences species selection and zonation by conditioning the dispersion of reproductive material by hydrochory and the moisture and nutrients conditions of soil substrates. After seed germination, vegetation grows accordingly to the moisture conditions and availability of nutrients and oxygen, but it can be inhibited by mechanical stress or change in conditions. Thus, during this phase, flow variability can affect the size the vegetation can reach both regarding its above-ground components and the roots. Low water stages induce drought stress affecting vegetation growth; particularly, if the level of phreatic surface declines below the minimum level that allows vegetation survival, an overall biomass reduction in the riparian corridor occurs. Also, high flows cause stress to vegetation. High water levels induce anoxia, reducing the amount of oxygen available at the root depth, whereas high velocities cause vegetation damage and uprooting and are usually associated to high sediment load fluxes that can bury existing vegetation. In each location, the site-specific conditions (i.e, moisture, depth of water table, nutrients, oxygen) define the carrying capacity of a vegetation species, or rather the maximum size a tree can achieve in those conditions.

Hydrogeomorphological disturbances are not the solely processes by which the recruitment, zonation, growth and decline of riparian vegetation are influenced, as also secondary factors play a relevant role in changing the existing status of riparian vegetation and modify the conditions that favour its development [16], [21], [52], [54], [65]:

- *Hydrological factor:* hydrodynamics field usually is not homogeneous and variability in roughness and local change in flow direction can alter its influence on sediment transport and vegetation growth.

- *Biological factors:* regardless the conditions in the riparian environment, vegetation growth can be affected by interspecific competition among different species. Invasion of allochthonous species, due to both natural convey of alien seeds by the stream and anthropic import, causes a reduction of indigenous plants simultaneously to an increase of alien ones, leading to variations of the overall amount of biomass.

- *Wood debris factor:* the accumulation of wood debris promotes the establishment of new vegetation. Wood debris constitute an obstacle to the ordinary flow, promoting scouring upstream and sedimentation downstream, thus creating a new island that can be colonized. Colonization is also favoured by the abundance of nutrients provided by dead wood.

- *Anthropic factor:* human activities strongly alter the flow regime, both in quantity and quality, and have, therefore, an influence on vegetation growth and survival. Furthermore, regardless hydrological conditions, biomass can be reduced by tree cutting as well as the species selection and zonation can be controlled by human pressure (e.g. import of allochthonous species and establishment of settlements and crops fields along rivers).

- *Animals factor:* mechanical damage or uprooting of weeds and trees in their early phase of growth can be induced by the action of herbivorous, even though the unpredictability and low probability of this occurrence usually prevent it to be mentioned, while modelling vegetation dynamics.

- *Other external factors:* finally, vegetation can be influenced both directly by random events such as wildfires, which completely remove the existing vegetation while creating optimum soil conditions for the colonization of fire-adapted species[1], and indirectly by climate change, which is gradually varying hydrological regime that, in turn, modifies its effects on riparian vegetation.

7.2 Description of the stochastic model

With the purpose of keeping the model analytically tractable [16], the authors decided to include solely the key processes conditioning the evolution of riparian vegetation by setting a series of simplifying hypotheses. Firstly, interspecific interactions in vegetation species, due to synergy or competitiveness, were overlooked, despite the features of the

[1]Fire-adapted species: species that can recover from wildfire faster than others and that are, therefore, most favoured in colonizing areas with a high wildfire hazard [54].

various species were taken in account to set the vegetation-related parameters. Secondly, in order to maintain constant the geometric asset of the problem, a steady river morphology was assumed, neglecting sedimentation, erosion and the influence that vegetation may have on these processes. Thirdly, the time delay between the vertical movements of the river water surface and the related change of groundwater level in the adjacent unconfined aquifer were disregarded considering that this delay is usually negligible in comparison to the timescale of interactions between vegetation and groundwater. Also, other crucial aspects were neglected, such as the eventual action of the external factors listed in the previous section (e.g. heterogeneity in flow hydraulics, accumulation of wood debris, influence of water quality on vegetation growth and decay, feedback of the the various species to climate change and the presence of herbivorous animals). By excluding these factors, the model could focus on the basic actions that regulate the dynamics of growth and decay of vegetation in riparian corridors, returning the steady state probability density function of vegetation biomass and its main statistics.

To sum up, according to this model, the stochasticity of vegetation dynamics is due to flow variability. This randomness is described by the variable h, representing the water levels, its probability density function and the integral scale. Also the other variables are random variables, as they all depend of water level (e.g. depth of phreatic surface δ, probability of inundation P_{I} and exposure P_{E}, carrying capacity β and growth rate α).

Temporal scale

The timescale, to which the model refers, ranges from a day to several years, as it accounts for both the short-term effect of inundations and the slower processes of vegetation growth. However, the results of the models must be evaluated in the long-term, as the simplifications could affect the short-term interpretations.

Geometric setting

This model refers to the spatial scale of a reach and the geometry configuration is set as a generic riparian transect with vertical datum coincident to the minimum water level, as shown in Figure 7.1. This configuration can be completely described by four dimensionless variables:

$$x = \frac{x^* - \overline{w}^*}{\overline{w}^*} \tag{7.1}$$

$$h = \frac{h^* - \overline{h}^*}{\overline{h}^*} \tag{7.2}$$

$$\zeta = \frac{\zeta^* - \overline{h}^*}{\overline{h}^*} \tag{7.3}$$

$$\eta = \frac{\eta^* - \overline{h}^*}{\overline{h}^*} \tag{7.4}$$

x is the transversal coordinate, whose origin is located at the intersection between the river bed and the minimum water level, w is the river width, h, ζ and η indicate the elevation of, respectively, the river water surface, the phreatic surface level and the topographic elevation of riverbed, in reference to the vertical datum. Overlined characters refer to the average value, while $*$ indicates dimensional parameters.

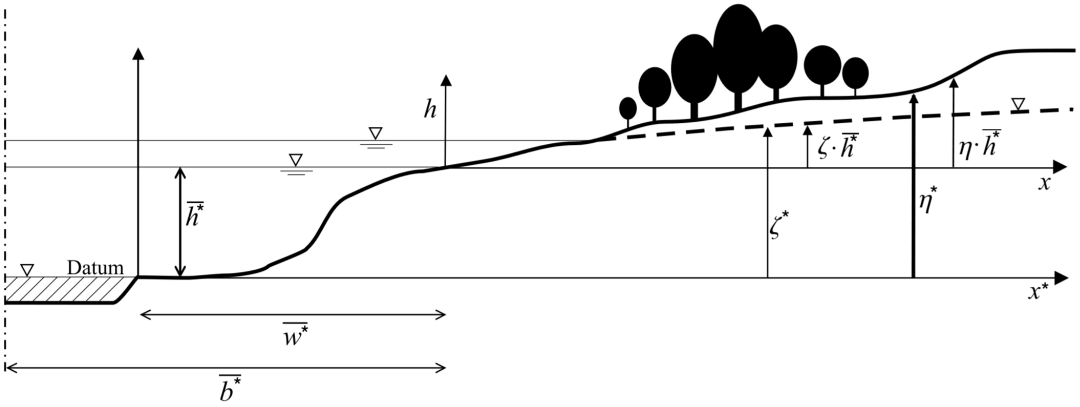

Figure 7.1. Sketch of the riparian transect, including the main variables of the model. Source: Camporeale and Ridolfi [16].

Input parameters

This stochastic model requires a set of input parameters that can be classified in three groups:

- *Hydrometric parameters:* hydrological fluctuations are describes by the water level h, which is a random variable, and its probability density function (pdf) $p(h)$, autocorrelation function (ACF) $\rho_{\mathrm{h}}(s^*)$ and integral scale τ^* constitute part of the necessary input data for running the model. Defining t as time and s as delay, the ACF is:

$$\rho_{\mathrm{h}}(s^*) = \frac{\overline{h(t^*)h(t^*+s^*)}}{h^2(t^*)} \tag{7.5}$$

  The integral scale corresponds to the area of ACF and represents the memory of a signal that, in this case, is the time series of water levels:

$$\tau^* = \int_0^x \rho_{\mathrm{h}}(s^*)ds^* \tag{7.6}$$

- *Geometric parameters:* once the vertical datum has been set, the variables $\eta(x)$ and $\zeta(x,h)$ completely describe the geometric configuration of the scenario the

model refers to. Starting from these two variables, a third one, the depth of the phreatic surface $\delta(x, h)$ can be defined:

$$\delta(x, h) = \eta(x) - \zeta(x, h) \tag{7.7}$$

As already stated, since both ζ and δ depend on the random variable h, they are random variables too.

- *Biologic parameters:* the features of vegetation are completely defined once a set of parameters has been fixed:
  - *n, m, p, α_2:* these are numerical constants depending on the characteristics of vegetation and whose values can be found in literature. A set of α_2 for various tree species was found by setting m=n=p=1 and fitting Equation 7.10 to the logistic Verhulst's model[2].
  - *K:* it is a positive coefficient that must be determined empirically and that depends on vegetation species. It is related to the decay of vegetation and it is necessary to computed the parameter α_1 that describes the rate of decay of vegetation due to damages caused by flooding:

    $$\alpha_1 = K(h - \eta) \tag{7.8}$$

    As the water level h raises, the stress that vegetation has to deal with increases and so the rate of decay. This trend is explained by observing that high water levels, corresponding to inundations, create anoxic conditions that affect vegetation health and mechanically damage or uproot the weakest plants. The assumption at the basis of the definition of this parameter is that the negative effects of flooding exceed the positive ones. K is the parameter whose calibration was part of the objectives of this dissertation.
  - *a, δ_{opt}, δ_1, δ_2:* these parameters lead to the definition of dimensionless carrying capacity V_c. More in detail, δ_1 and δ_2 limit the range of depths of phreatic surface that allows vegetation to survive; a and δ_{opt} determine the values of V_c inside the range [δ1, δ_2]. All these parameters are dimensionless, but they derive from dimensional measures that have been normalized in relation to the average dimensional water level $\overline{h^*}$, as explained at Section 8.1.

Output

Information about the distribution of vegetation along the riparian corridor, in relation to topography, hydrological regime and vegetation features, are the expected outcome of this stochastic model.

[2]Verhulst's logistic model: a sigmoid curve describing the growth of population. According to it, population grows exponentially during the earliest phases, slowing the growth rate in time until an asymptote, the carrying capacity of that population, is reached. In practice, this trend means that the growth rate is inversely proportional to the amount of population.

In practice, the probability distribution function (pdf) of the dimensionless biomass $p(\nu)$ and its first four moments are the model output. These values can be converted in dimensional biomass, if the carrying capacity is known, and further transformed in carbon so that the potential of carbon sequestration in the riparian corridor can be defined. The conversion of vegetation biomass and carbon is a simple proportion as the carbon content is around the 50% of total dry biomass [48].

Formulation

The starting point for formulating the stochastic model that must be calibrated is the couple of equations that constitute a simple model of vegetation dynamics:

$$\frac{d\nu}{dt^*} = -\alpha_1 \nu^{\mathrm{n}} \qquad for\ h \geq \eta \tag{7.9}$$

$$\frac{d\nu}{dt^*} = \alpha_2 \nu^{\mathrm{m}} (V_{\mathrm{c}} - \nu)^{\mathrm{p}} \qquad for\ h < \eta, \tag{7.10}$$

where t* is dimensional time. The first equation of this model describes the response of riparian vegetation to floods: as the negative consequences of inundation are assumed as greater than the positive ones, Equation 7.9 models the decay of vegetation. On the contrary, the second one, Equation 7.10, describes the growth of phreatophyte plants during exposure period. During this growing phases, biomass cannot exceed the carrying capacity that, in turn, depends on groundwater level.

Considering the probabilities of inundation P_{I} and exposure P_{E} along the x-coordinate, the parameters α_1 and V_{c} can be approximated by their weighted average values:

$$P_{\mathrm{I}} = \int_{\eta(x)}^{\infty} p(h)dh \tag{7.11}$$

$$P_{\mathrm{E}} = \int_{-1}^{\eta(x)} p(h)dh \tag{7.12}$$

$$\langle \alpha_1 \rangle = \frac{1}{P_{\mathrm{I}}} \int_{\eta(x)}^{\infty} \alpha_1 p(h)dh \tag{7.13}$$

$$\langle V_{\mathrm{c}} \rangle = \frac{1}{P_{\mathrm{E}}} \int_{-1}^{\eta(x)} V_{\mathrm{c}} p(h)dh \tag{7.14}$$

Equations 7.11, 7.12, 7.13, 7.14 can be inserted in Equations 7.9 and 7.10:

$$\frac{d\nu}{dt^*} = -\langle \alpha_1 \rangle \nu^{\mathrm{n}} \qquad for\ h \geq \eta \tag{7.15}$$

$$\frac{d\nu}{dt^*} = \alpha_2 \nu^{\mathrm{m}} (\langle V_{\mathrm{c}} \rangle - \nu)^{\mathrm{p}} \qquad for\ h < \eta \tag{7.16}$$

Scaling both the equations by α_2 and assuming the following dimensionless parameters t, α and β, Equations 7.9 and 7.10 become:

$$\frac{d\nu}{dt} = -\alpha\nu^{\mathrm{n}} \qquad for\ h \geq \eta \tag{7.17}$$

$$\frac{d\nu}{dt} = \nu^{\mathrm{m}}(\beta - \nu)^{\mathrm{p}} \qquad for\ h < \eta, \tag{7.18}$$

where:

$$t = \alpha_2 t^{*} \tag{7.19}$$

$$\alpha = \frac{\langle \alpha_1 \rangle}{\alpha_2} = \frac{K(h-\eta)}{\alpha_2} = \frac{k(h-\eta)}{\alpha_2} \tag{7.20}$$

$$\beta = \langle V_{\mathrm{c}} \rangle \tag{7.21}$$

As the use of Equation 7.9 instead of Equation 7.10, and vice versa, depends solely on the switching between inundation and exposure, the two equations can be summarized in a single stochastic equation:

$$\frac{d\nu}{dt} = f(\nu) + \mathcal{E}(t)g(\nu) \tag{7.22}$$

$\mathcal{E}(t)$ allows for the switching between inundation and exposure states, while the functions $f(\nu)$ and $g(\nu)$ are formulated in a way to return Equation 7.17 in case of flooding Δ_{I} and Equation 7.18 in case of exposure Δ_{E}:

$$f(\nu) = \frac{\Delta_{\mathrm{I}}\nu^{\mathrm{m}}(\beta - \nu)^{\mathrm{p}} + \alpha\Delta_{\mathrm{E}}\nu^{\mathrm{n}}}{\Delta_{\mathrm{I}} - \Delta_{\mathrm{E}}} \tag{7.23}$$

$$g(\nu) = \frac{\alpha\nu^{\mathrm{n}} + (\beta - \nu)^{\mathrm{p}}\nu^{\mathrm{m}}}{\Delta_{\mathrm{E}} - \Delta_{\mathrm{I}}} \tag{7.24}$$

The solution of Equation 7.22 is the pdf of dimensionless vegetation $p(\nu)$, whose steady state function, obtained for $t \to \infty$, is:

$$p(\nu) = \frac{N}{\alpha}\nu^{\frac{\beta(1-\alpha\tau)-(\alpha+\beta)P_{\mathrm{I}}}{\alpha\beta\tau}}(\beta - \nu)^{\frac{P_{\mathrm{I}}}{\beta\tau}\text{-}1}(\alpha + \beta - \nu) \tag{7.25}$$

N is a normalization constant so that $\int_0^{\infty} p(\nu)d\nu = 1$. Equation 7.25 is valid if:

$$P_{\mathrm{I}} < \frac{\beta}{\alpha + \beta} \tag{7.26}$$

This condition depends on bed elevation $\eta(x)$ and it allows for the determination of the physical limits of the riparian corridors.

Finally, starting from Equation 7.25, the first two moments of dimensionless vegetation μ_1 and μ_2 can be computed:

$$\mu_1 = \int_0^{\beta}(\nu)d\nu = \frac{\beta\rho_1(\beta\varepsilon_2\sigma_1 - \varepsilon_1\rho_0\sigma_2)}{(\beta\varepsilon_1\rho_1 - \rho_0\rho_{21})\sigma_2} \tag{7.27}$$

$$\mu_2 = \frac{\beta^2\rho_1}{(\beta\varepsilon_1\rho_1-\rho_0\rho_2\sigma_1)^2{\sigma_2}^2\sigma_3}[\beta\varepsilon_1\rho_1\sigma_1\sigma_2(\beta\varepsilon_3\sigma_2+\varepsilon_2\rho_0\sigma_3)$$

$$-({\varepsilon_1}^2{\rho_0}^2\rho_1{\sigma_2}^2\sigma_3)-{\sigma_1}^2(\beta\varepsilon_3\rho_0\rho_2{\sigma_2}^2+\varepsilon_2(\beta^2\varepsilon_2\rho_1-{\rho_0}^2\rho_2\sigma_2)\sigma_3)] \tag{7.28}$$

The parameters for computing these moments are defined as:

$$\rho_0 = \alpha + \beta \tag{7.29}$$

$$\rho_1 = \Gamma\left[\frac{1-P_\mathrm{I}}{\alpha\tau}\right] \tag{7.30}$$

$$\rho_2 = \Gamma\left[\frac{1-P_\mathrm{I}}{\alpha\tau}-\frac{P_\mathrm{I}}{\beta\tau}\right] \tag{7.31}$$

$$\sigma_\mathrm{i} = \Gamma\left[\frac{1-P_\mathrm{I}}{\alpha\tau}+i\right] \tag{7.32}$$

$$\varepsilon_\mathrm{i} = \Gamma\left[\frac{1-P_\mathrm{I}}{\alpha\tau}-\frac{P_\mathrm{I}}{\beta\tau}+i\right], \tag{7.33}$$

where $\Gamma[\cdot]$ is the Gamma function.

Equation 7.27 was used to calibrate the stochastic model, while Equation 7.28 to validate the results by comparing the computed second moment of vegetation in the study area and the LiDAR-derived values.

Chapter 8

Vegetation data

Although the stochastic model allows for the definition of the evolution of riparian vegetation at any point of the fluvial corridor, the aim of this dissertation was to calibrate the decay parameter K, or rather its ratio with the growth rate α_2, $k = \frac{K}{\alpha_2}$. Hence, vegetation had to be an input data. More in detail, the required information were the parameters n, m, p, α_2, β and the probability density function of vegetation $p(\nu)$. Values for parameters n, m, p, which depend on vegetation characteristics, were found in literature. By processing the available LiDAR data, values for ν, which is the dimensionless vegetation biomass, were estimated and $p(\nu)$, which describes the trend of vegetation for increasing topographic altitude was defined. The dimensionless carrying capacity V_c was computed by defining values for the parameters δ_1 and δ_2.

As models for estimating biomass from LiDAR acquisitions returned a dimensional value of vegetation ν^*, expressed as Mg·ha^{-1}, the dimensional carrying capacity ${V_c}^*$ was computed and used for normalizing ν^*, obtaining dimensionless biomass ν.

8.1 Vegetation characteristics in the study area

As mentioned at Section 3.2, most of the floodplain is occupied by crop fields, pressing riparian vegetation in a narrow band along the water banks or above the islands, which are sparsely present. Riparian vegetation, is mainly constituted by shrubs, whose dominant species are bitter willow *(Salix Elaeagnos)*, black willow *(Salix Atrocinerea)* and white willow *(Salix Alba)* but also mastics *(Pistacia Lentiscus)*, black poplars *(Populus Nigra)* and white poplars *(Populus alba)*. In the spots where inundations are more likely to occur, grass-like plants, as common reeds *(Arundo Donax and Phragmites Australis)* and bulrush *(Typha sp.)* grow [61].

Due to the mixture of trees in the study area and paucity of literature data about less common species, an equivalent tree population was defined. In this case, a tree species with intermediate features between willows *(Salix spp.*[1]*)* and poplars *(Populus spp.)*

[1] *Spp.: in botany it indicates that the author is referring to all the species of a certain genus. It is*

was considered sufficiently realistic. The calibration of needed parameters for running the stochastic model relied on this assumption.

Parameters m, n, p

Parameters m, n and p are numerical constants and they are related to the characteristics of vegetation. Despite the assignation of values to these parameters may influence the solution of the model, in this dissertation they were fixed equal to 1, according to literature examples [16].

Parameter α_2

Values of α_2, which regulates the intrinsic rate of vegetation growth, were found in literature for *Populus Deltoides* ($\alpha_2 = 13{\cdot}10^5$ d^{-1}) and *Salix Nigra* ($\alpha_2 = 28{\cdot}10^5$ d^{-1}). These values had been set by fitting Equation 7.10 with already validated logistic model for vegetation growth [16].

Parameters δ_1, δ_2, β

Starting from the values of the optimum depth ${\delta_{\text{opt}}}^*$ of water table, δ_1 and ${\delta_2}$[2] were calculated as:

$$\delta_1 = \delta_{\text{opt}} - \sqrt{a} \tag{8.1}$$

$$\delta_2 = \delta_{\text{opt}} + \sqrt{a} \tag{8.2}$$

These dimensionless parameters are the results of normalization of dimensional values in relation to the average dimensional water level $\overline{h^*}$; δ_{opt} is equal to the ratio between dimensional depth ${\delta_{\text{opt}}}^*$ and $\overline{h^*}$, while a is equal to the product between the constant a^* and $\overline{h^*}^2$:

$$\delta_{\text{opt}} = \frac{{\delta_{\text{opt}}}^*}{\overline{h^*}} \tag{8.3}$$

$$a = a^* \overline{h^*} \tag{8.4}$$

Once these parameters are fixed, the dimensionless carrying capacity V_c can be calculated as:

$$V_c = V_c(\delta) = 1 - a(\delta - \delta_{\text{opt}})^2 \qquad for\ \delta_1 \leq \delta \leq \delta_2 \tag{8.5}$$

$$V_c = V_c(\delta) = 0 \qquad for\ \delta < \delta_1\ or\ \delta > \delta_2 \tag{8.6}$$

---

generally used when there is a mixture of plants of the same genus, but the precise amount of elements for each species inside the genus is not known.

[2] ${\delta_{\text{opt}}}^*$ is the depth of phreatic surface that allows for optimal vegetation growth, while δ_1 and δ_2 are respectively the lower and upper limit of the range of phreatic surface depth allowing for vegetation survival.

Literature values of $\delta_{opt}{}^{*}$ for *Populus spp.* and *Salix spp.* are reported in Table 8.1 together with their average value that was considered for the equivalent tree population in the study area. As the study area was split in eleven subareas (i.e. eleven reaches, see Section 8.3), a set of eleven values for $\overline{h}$ was defined and reported in Table 9.1. Regarding a^{*}, a reference value (a^{*}=0.055) was found in literature [16].

Table 8.1. Computed $\delta_{opt}{}^{*}$ for the two tree species and the average equivalent value $\delta_{opt,eq}{}^{*}$ for the study area.

Tree Species	$\delta_{\mathbf{opt}}{}^{*}$ **(m)**	$\delta_{\mathbf{opt,eq}}{}^{*}$ **(m)**
Populus spp.	1.0	0.75
Salix spp.	0.5	

8.2 Definition of the equivalent carrying capacity

Despite the dimensionless carrying capacity V_c was calculated with Equations 8.5 and 8.6, it was necessary to define a value for the dimensional carrying capacity $V_c{}^{*}$ related to willows and poplars, which are the two species constituting the equivalent tree population of the study area. $V_c{}^{*}$ values were required to normalized dimensional vegetation biomass ν^{*}, provided by LiDAR data processing, and obtaining ν.

Literature offers equations to compute the dimensional carrying capacity through allometric relations. As the consistency between the chosen model for LiDAR estimation of biomass (Means et al. 2000) and the allometric formula of Jenkins et al. had been previously tested, this latter was used to compute local values of $V_c{}^{*}$[3]. However, Jenkins et al. formula (Equation 4.1) requires DBH as input data and also another allometric model was necessary in order to relate the available information (i.e. the statistics about vegetation height) to DBH:

$$H_{v,max} = b_1 + b_2 D_{v,max} - b_3 D_{v,max}{}^{2}, \tag{8.7}$$

where $H_{v,max}$ and $D_{v,max}$ refer to the maximum size of the plant for optimal growing conditions; b_1 is the breast height, equal to 137 cm, while b_2 and b_3 are growth rate parameters, whose values were found in literature [47] and reported in Table 8.2. This relation was initially formulated by Trorey (1932) [60] and successively proposed again by Ker and Smith (1955) [31]. In summary, the procedure to obtain $V_c{}^{*}$ consisted of:

- *Computation of the maximum DBH D_{max} :* Equation 8.7 was implemented, once the maximum vegetation height was set according to LiDAR measurements.

[3] *Local* means that the carrying capacity was computed for each topographic band z within all the reaches i, obtaining a $z \cdot i$ matrix of $V_c{}^{*}$ values.

Table 8.2. Values of b_i parameters for the study case.

Tree Species	$\mathbf{b_1}$ **(cm)**	$\mathbf{b_2}$ **(-)**	$\mathbf{b_3}$ **(cm^{-1})**
Populus spp.	137	28.42	0.0388
Salix spp.	137	24.15	0.1348

- *Computation of dry above-ground biomass:* Jenkins et al. model (see Equation 4.1) was used to this purpose. Literature values for input parameters were the ones shown in Section 4.3.

- *Conversion of dry biomass in fresh biomass:* the stem and bark fraction of dry biomass accounts for approximately the 97% of the total above-ground biomass. According to Equation 4.13, the ratio between fresh and dry biomass is equal to $(1 + \theta)$, where θ is the moisture content. Typical moisture values for the two reference species, willows and poplars, were found in literature and reported in Table 8.2.

Table 8.3. Values of θ for the study case.

Tree Species	θ (%)	$\theta_{\mathbf{eq}}$ (%)
Populus spp.	129	105.5
Salix spp.	82	

As this procedure was repeated for each of the 54 topographic bands, the result was an array of 54 $V_c(z)$ values, representing local carrying capacity. The trend of carrying capacity for increasing altitude is shown in Figure 8.1. The maximum value was chosen as reference carrying capacity for the entire area: $V_{c,eq}{}^{*} = 142.5$ Mg·ha^{-1}.

This procedure was included in the MATLAB script reported in Appendix A.

8.3 Distribution of vegetation in the riparian corridor

In order to understand the relationship between biomass and topography, a procedure similar to the one used for testing biomass models was adopted. Through FUSION/LDV package LiDAR data were processed, while the visualization of surface models and their splitting according to altitude was made in GIS environment. Also MATLAB was used with the purpose of regressing the inclined plane that represents the average surface of water level during the LiDAR acquisition.

Reference LiDAR data were the same that were used in Part III, or rather the LiDAR datasets captured on 20th October 2016 and downloaded from CNIG online platform [1].

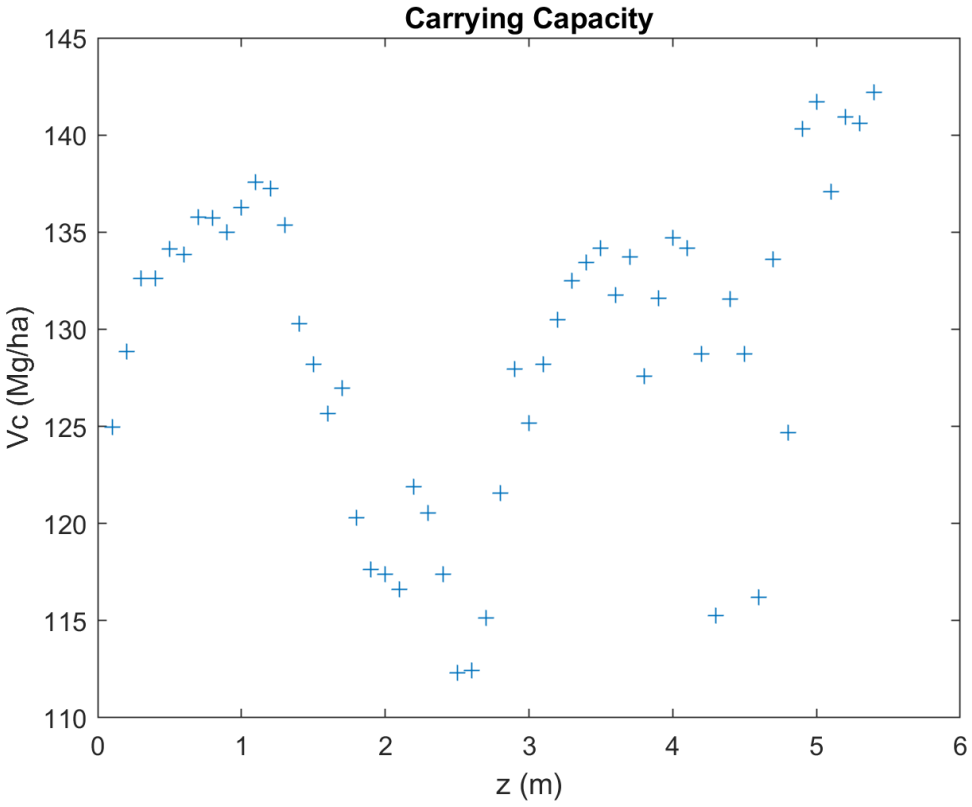

Figure 8.1. Trend of carrying capacity for increasing altitude in the study area.

Creation of the ground surface model and definition of reaches

Similarly to what was described at Section 5.2, *.laz* files related to the sample area were firstly transformed in a *.las* file by running *LDA2LAS* and then clipped by *PolyClipData* in order to eliminate all the data outside the floodplain. Successively, intensity and first return intensity images were created with *Catalog* as well as a DTM through the sequence of *GroundFilter* and *GridSurfaceCreate* command lines. The z-coordinates of points constituting this DTM referred to absolute altitude (i.e. metres above sea level).

The comparison among DTM, intensity images and first return density images allowed for the definition of the water surface boundaries at the time of acquisition. Then, the selected river segment was divided in sections of approximately homogeneous water surface width, defining eleven reaches. These reaches were used for both estimating biomass estimation and processing hydrometric data.

Regression of water surface

As the final goal of defining a LiDAR processing procedure was to understand the relation between vegetation biomass and floodplain topography, the DTM was normalized in relation to the river bed altitude. However, LiDAR provided only the water surface and not the river bed profile, since water completely absorbs LiDAR pulses. Thus, in a first moment, the DTM was normalized in relation to the inclined plane that represented the average water surface at the time of acquisition[4].

The normalized DTM was created by combining GIS and MATLAB functions. A large number of water surface points, constituting a 3D cloud, were selected in QGIS and finally the plane was defined through a least square regression of z-coordinates in MATLAB.

Considering the plane equation:

$$z = a + bx + cy \tag{8.8}$$

the optimal approximate plane was the one that minimized the sum of the N squares of the deviations between the theoretical values z_i^* located on the plane and the empirical values z_i derived from LiDAR measures:

$$\sum_{i=1}^{N}(z_i^* - z_i)^2 = min \tag{8.9}$$

Therefore, the outcome of the regression was the value for the parameters a, b and c. The used MATLAB script is reported in Appendix A.

After the definition of the inclined plane equation, a punctual *.shp* file representing the points that belong to the plane was imported in QGIS and a DTM of the plane was created by interpolating the imported points. Finally, *Raster Calculator* functions allowed for the normalization of the ground surface, by subtracting to each point the altitude of the plane in the same position.

Estimation of biomass for the various reaches

For a better understanding of the adopted procedure, the steps that have been followed for each reach are briefly explained:

- *Split of DTMs:* as the goal was to analyze the trend of biomass with topography, every normalized DTM was split into altitude bands 0.10 m high, by running *Reclass* command in QGIS.
- *Creation of shapefiles:* every altitude bands was converted into a *.shp* file in order to clip the LiDAR dataset with FUSION.

[4] A further correction was made after the calculation of water depth at the time of acquisition.

- *Use of FUSION/LDV functions:* for each altitude band, the sequence of *PolyClipData*, *GroundFilter*, *GridSurfaceCreate* and *GridMetrics* was run to obtain statics about vegetation height and cover.
- *Computation of biomass statistics:* average, standard deviation and pdf were computed for each altitude band by processing FUSION/LDV output files with MATLAB and implementing Means et al. (2000) formula (Equation 4.10).

The implemented script is reported in Appendix B. As explained at Section 6.3, cell size for *GridMetrics* did not affect biomass computed values, not even if altitude bands had complex shapes and small width. The chosen model for estimating above-ground riparian biomass was the one proposed by Means et al. 2000 (Equation 4.10). Equation 4.10 returned the volume of above-ground vegetation V, expressed in $m^3 \cdot ha^{-1}$ and, therefore, it was converted in biomass AB through the equivalent density $\rho_{v,eq}$ expressed in $Mg \cdot ha^{-1}$. This dimensional value of biomass was normalized according to the equivalent carrying capacity ${V_{c,eq}}^*$ and obtaining dimensionless biomass ν.

Chapter 9

Hydrological data

The hydrometric input parameters of the model describe the randomness of hydrological fluctuations in the fluvial segment and consist of statistical functions of water levels h, namely the pdf $p(h)$, the ACF $\rho_{\mathrm{h}}(s^*)$ and the related integral scale τ^*. However, also the average water level in each reach was computed as it was fundamental for the conversion of various dimensional parameters (e.g. $\zeta^*, \eta^*, \delta^*, a^*$) to the related dimensionless ones (e.g. ζ, η, δ, a).

As the solely available data were the Cinca discharges of the historical series that had been provided by *Confederación Hidrográfica del Ebro* [3], water level statistics were computed after having processed these discharges. The first step was to correct the DTM that had been realized for the study area by processing LiDAR data in FUSION/LDV. This correction was done in order to create the river channel and import cross sections in HEC-RAS. Successively, HEC-RAS 1D was run eleven times to compute the water levels in the various reaches for all of the reference flow rates, which were a set of eleven percentiles of the available series of Cinca discharges. After that, the resulting water levels were processed in MATLAB defining a rating curve $h - Q$ for each one of the reach. Finally, all the rating curves were used to create temporal series of water levels starting from the series of flow rates and statistics about these series were computed.

9.1 HEC-RAS modelling for the study case

HEC-RAS overview

The Hydrologic Engineering Center's River Analysis System (HEC-RAS) is an embedded system of software that enables users to simulate 1D and 2D steady and unsteady river flow, compute of movable boundary sediment transport and model water temperature and constituent transport, as explained in its users' manual [46].

In order to collect the hydrometric data that the stochastic model calibration required, the sole 1D steady flow river component was used. The software calculates water surface profiles for steady gradually varied flow by computing water stages throughout the interior cross sections. Water stages are the outcomes of an iterative procedure,

called *standard step method*, that relies on the solution of the mono-dimensional[1] energy equation among two adjacent cross sections:

$$Z_2 + Y_2 + \frac{a_2 {V_2}^2}{2g} = Z_1 + Y_1 + \frac{a_1 {V_1}^2}{2g} + h_e \tag{9.1}$$

Z is river bed elevation in relation to the topographic datum, Y the water surface elevation in relation to the river bed, $\frac{aV^2}{2g}$ is the kinetic height, or rather the height of the energy grade line in relation to the water surface. The coefficient a is the kinetic correction factor, V the flow longitudinal velocity and g the gravitational acceleration. The factor h_e represents the energy head loss from section (2) to section (1).

The kinetic correction factor a can be calculated by comparing the kinetic height $\frac{a\overline{V}^2}{2g}$ and the weighted average of kinetic height according to the flow rate of the individual sections $\frac{\sum_{i=1}^{N} Q_i \frac{{V_i}^2}{2g}}{Q_{tot}}$. Considering as N different sections the main central channel and the floodplain:

$$\frac{a\overline{V}^2}{2g} = \frac{\sum_{i=1}^{N} Q_i \frac{{V_i}^2}{2g}}{Q_{tot}} \longrightarrow a = \frac{\sum_{i=1}^{N} {V_i}^2 Q_i}{\overline{V}^2 Q_{tot}} \tag{9.2}$$

The energy head loss h_e comprises both the contribution of friction and the contribution of fluid expansion/contraction:

$$h_e = L S_f + C \left| \frac{a_2 {v_2}^2}{2g} - \frac{a_1 {v_1}^2}{2g} \right| \tag{9.3}$$

C is the factor of fluid expansion/contraction that ranges between 0 and 1 and L the reach length, weighted on flow rates of central channel ($_{ch}$), left floodplain ($_l$) and right floodplain ($_r$):

$$L = \frac{L_{ch} Q_{ch} + L_l Q_l + L_r Q_r}{Q_{ch} + Q_l + Q_r} \tag{9.4}$$

S_f is the loss of head per unit reach length computed through Chézy's Formula:

$$Q = \frac{1}{n} S R^{2/3} {S_f}^{1/2} \longrightarrow S_f = \left(\frac{Qn}{SR^{2/3}} \right)^2 \tag{9.5}$$

The *standard step method* comprises few phases:

1. Assuming that the boundary condition at cross section (1) has been set and the discharge flowing in the reach is known, a water surface elevation at the upstream cross section (2)[2] ${h_2}^*$ is arbitrarily chosen.

[1] Solely the longitudinal component of velocity is considered.

[2] Cross section (2) is located upstream in case of subcritical regime and downstream in case of supercritical regime.

2. V_2 is calculated by knowing the cross section geometry, the flow rate and ${h_2}^*$.
3. S_f and h_e are calculated with Equations 9.3 and 9.5.
4. Y_2 is calculated by knowing all the other parameters.
5. Equation 9.1 is used to compute h_2
6. The new value for the water level at section (2), h_2, is compared with the one set at phase 1, ${h_2}^*$, and the procedure is repeated until convergence, $|{h_2}^* - h_2| < threshold$.

In case of transition between slow and fast flows, the water surface level pass through the critical depth and the Equation 9.1 loses its validity. Thus, the software uses the Moment Equation:

$$\frac{a_2 {Q_2}^2}{gA_2} + A_2 Y_2 - \frac{a_1 {Q_1}^2}{gA_1} - A_1 Y_1 + \left(\frac{A_1 + A_2}{2}\right) LS_0 - \left(\frac{A_1 + A_2}{2}\right) LS_\mathrm{f} = 0 \qquad (9.6)$$

$\frac{a_\mathrm{i} {Q_\mathrm{i}}^2}{gA_\mathrm{i}}$ is the hydrodynamic component of hydraulic force at the i-th section, $A_\mathrm{i} Y_\mathrm{i}$ is the hydrostatic component, $(\frac{A_1+A_2}{2})LS_0$ is the contribution of inertial force along the longitudinal flow direction and $(\frac{A_1+A_2}{2})LS_\mathrm{f}$ is the contribution of frictional force. S_0 is the river bed slope.

Input data for HEC-RAS modelling comprise the description of the channel (i.e. length and shape of the reach, cross sections geometry and roughness), a set of boundary conditions and the geometry of eventual structures (e.g. bridges, weirs, culverts). Boundary conditions for steady flow analyses can be specified as:

- *Known water surface:* if the reach ends (or starts) in a water basin of known water depth, this water surface elevation can constitute a boundary condition for the system.
- *Critical depth:* choosing this type of boundary condition the software computes the critical depth for each cross section without needing other information.
- *Normal depth:* if the river slope defined by users, the depth that guarantees the uniform flow can be calculated by the software and set as boundary conditions.
- *Rating curve:* this type of boundary condition requires the specification of water stages for different flow rates.

In case of subcritical flow, as in the study segment, a downstream boundary condition is required to run the simulation, while an upstream one is required in case of supercritical flow. For mixed regime, both upstream and downstream must be set.

Import of the geometry

As the impossibility to obtain river cross sections from literature or in field surveys, exploitation of LiDAR data was considered the most suitable solution. However, despite processing LiDAR data led to the creation of a DTM of the study area, bathymetry was not represented in this DTM as the electromagnetic signal emitted by the laser scanning system during the LiDAR acquisition had not the right wavelength to pass through water and capture the river bed, being stopped by the water surface[3].

In order to overcome this limitation, the DTM was corrected by processing it in QGIS. Firstly, the comparison among DTM, intensity images and first return density images allowed for the definition of the water surface boundaries at the time of acquisition so that DTM[4] could be split in *land*, *islands* and *water*. Secondly, *water* was divided in eleven reaches where the water surface width was approximately homogeneous and, according to Chézy's Formula, water depth at the time of acquisition was found for each of the these reaches. Assuming, uniform flow, Chézy's Formula is:

$$Q = SC(Ri)^{1/2} \tag{9.7}$$

Where S is the cross section of the river, C the factor of flow resistance representing the roughness of the channel, R the hydraulic radius of the section and i is the slope of energy line that, in case of uniform flow, is assumed equal to the slope of river bed. Choosing the Manning's formula to explicit the friction factor C:

$$C = \frac{1}{n} R^{1/6} \tag{9.8}$$

Equation 9.7 can be written as:

$$Q = S\frac{1}{n} R^{2/3} i^{1/2} \tag{9.9}$$

Where n is the Manning's coefficient describing the roughness of the channel. Assuming the shape of river cross sections as rectangles having height equal to h and width equal to the water surface width b, the hydraulic radius is $R \simeq$ h and Equation 9.9 becomes:

$$Q = \frac{1}{n} b i^{1/2} h^{5/3} \tag{9.10}$$

In this case, all the parameters were known except for water depth h:

---

[3]As explained at Chapter 4.1, laser wavelength is linked to the target of LiDAR surveys. Data that were used in this dissertation had been captured with a laser light of around 1 μm, while bathymetry requires blu-green light having wavelength equal to 0.5 - 0.6 μm.

[4]The DTM to be modified was the one that had been previously normalized in relation to the plane representing water surface during LiDAR acquisition (see Section 8.3) so that the water depth could be simply subtracted. Successively, in order to import the geometry in Hec-Ras, the corrected DTM was denormalized.

- Q : daily average discharge for the day of LiDAR scanning, 20th October 2016, was found in the available temporal series of discharges [3] being equal to 31.3 $m^3 \cdot s^{-1}$.
- n : Manning's coefficient was set equal to 0.03. This value, consistent with the geological description of the area, was initially found among the reports of a project in the Ebro catchment [8] and successively confirmed by one of the authors of a study about Cinca River flooding in the city of Fraga [43]. For the river banks Manning's coefficient was increased to 0.1, according to both existing projects [8] and literature reference values [18].
- b : water surface width, assumed equal to the bed width in a rectangular geometry, was measured on the created DTM for each reach.
- i : slope of river bed was defined during the creation of a inclined plane that represented the average water surface during the acquisition, as explain at Section 8.3. The resulting value, $i = 0.003$ is consistent to the ones found in literature [62].

Once the water depth for all the reaches was calculated, all the eleven parts composing the *water* DTM were corrected by lowering the altitude of their points and thus artificially creating the river channel. Finally, a more realistic DTM of the study area was created by merging corrected *water*, *islands* and *land* DTMs.

As the chosen GIS software was QGIS, the plugin that was used to import the study geometry in Hec-Ras, was *Q-RAS*. This plugin was able to create cross section profiles along the river segment, once the axis of the river and the position of the cross sections had been set. In this case, twenty-six cross sections were located approximately 200 m apart along the river axis.

Definition of boundary conditions

Mixed flow regime was chosen to compute water surface profiles. Thus, the discharge flowing in the segment had to be set, together with upstream and downstream boundary conditions.

Starting from the available temporal series of discharges, eleven percentiles were computed (1st, 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, 90th,10th, 100th) and set as upstream condition for eleven simulations.

The chosen type of boundary condition was *normal depth*, or rather fixing a slope value i so that the software could compute the normal depth for the reach. As already mentioned, it was set $i = 0.003$ for each simulation.

Processing of HEC-RAS outcomes

After each simulations, water levels in all the cross sections were recorded and imported in MATLAB where thy were processed in order to computed the input parameters for the stochastic model calibration. The process consisted in few steps that were iterated for all the eleven reaches and that can be summarized as:

- *Definition of the rating curve:* the curve that fitted water levels h computed by HEC-RAS versus the reference discharges Q could be described by the equation:

$$Q = Ah^{\mathrm{B}} \tag{9.11}$$

  The chosen values of parameters A and B were the ones that guarantee the best matching between the computed curve and real data.

- *Computation of a temporal series of water levels:* once the rating curve had been defined, it could be applied to the historical series of discharge converting it in a temporal series of water levels.

- *Computation of statistics:* average, standard deviation, probability distribution function, cumulative distribution function, autocorrelation function and integral scale were calculated starting from the temporal series of water levels. The chosen type of distribution was lognormal due to its good fit of available data.

The used MATLAB script is reported in Appendix C.

9.2 Resulting hydrometric data

The average $\overline{h}$ and standard deviation σ_{h} of water levels for all the eleven reaches are reported in Table 9.1, while an overview of the probability density functions is shown in Figure C.2. τ^* was approximately 13 days for every reach.

Table 9.1. Results of hydrometric data processing.

Reach	$\overline{h}$ m	$\sigma_{\mathbf{h}}$ m
1	1.51	0.68
2	1.63	0.66
3	1.09	0.57
4	1.50	0.59
5	1.32	0.60
6	2.08	0.78
7	1.21	0.54
8	1.02	0.62
9	1.54	0.66
10	1.37	0.51
11	1.23	0.53

Chapter 10

Model calibration

Once all the necessary parameters had been set according to literature values and the input data had been collected, the stochastic model was calibrated. To this purpose, a MATLAB script was created in order to define an array of $k(z)$ values, or rather one parameter k for each topographic band z. Those values were set as the ones that minimize the square deviations between computed data of dimensionless biomass $\mu_{1,\mathrm{C}}$ and LiDAR-measured data $\mu_{1,\mathrm{L}}$, according to Equation 7.27:

$$\mu_1 = \int_0^{\beta}(\nu)d\nu = \frac{\beta\rho_1(\beta\varepsilon_2\sigma_1 - \varepsilon_1\rho_0\sigma_2)}{(\beta\varepsilon_1\rho_1 - \rho_0\rho_{21})\sigma_2}$$

The reliability of the calibrated values can be tested by comparing computed and LiDAR-derived values for the two first moments and pdf for each topographic band.

10.1 Considerations and assumptions

The expected objective of calibration was to set a value of the parameter k for each altitude band z, yet considering that the study area was split in 11 reaches having homogeneous hydrological features. Some considerations about the boundaries of the interval of values that k can assume and some simplifications regarding input parameters were made in order to facilitate the process:

- k: as Equation 7.25 is valid if Condition 7.26 is satisfied, this condition can be used to define the upper limit of k for each topographic band. Thus, defining:

$$a = \frac{1}{P_{\mathrm{I}}}\int_{\eta(x)}^{\infty}(h-\eta)p(h)dh, \tag{10.1}$$

  α can be written as:

$$\alpha = ka \tag{10.2}$$

and Condition 7.26 becomes:

$$k < \frac{\beta}{a}\left(\frac{1}{P_{\mathrm{I}}} - 1\right) = \frac{\beta}{\int_{\eta(x)}^{\infty}(h-\eta)p(h)dh}(1 - P_{\mathrm{I}}) \tag{10.3}$$

The lower limit for k was set as 0, considering that the rate of vegetation decay and growth are positive by definition and that negative values have no a physical significance.

- β: having normalized the LiDAR-derived dimensional values of biomass according to the maximum carrying capacity of the area, local variability of the dimensionless carrying capacity β can be neglected, assuming $\beta = 1$. Also, the reference to a unique value of carrying capacity allowed for the comparison among different topographic bands.

- η : as the dimensional topographic heights refer to the average water surface, while the model refers to a vertical datum that is set equal to the minimum water level, the difference between the water level during the LiDAR acquisition and the minimum water level was computed and used for the correction. To this purpose, the two known flow rates $Q_{\mathrm{acq}}, Q_{\mathrm{min}}$ and the coefficients $A_{\mathrm{i}}, B_{\mathrm{i}}$ were used to define the water levels, through the definition of rating curves (Equation 9.11). The corrected dimensional values were then normalized according to Equation 7.4 in order to obtain their dimensionless corresponding ones.

10.2 Computation and results

The proposed solution consisted in creating a MATLAB one-variable function $mu\_dev(k)$, which returns the square deviations among LiDAR-derived dimensionless biomass values and computed ones, and using an already existing function (i.e $fminbnd$). The function $fminbnd$ allows to find the value of k that minimizes the function $mu\_dev(k)$ and it must be run for each altitude band, setting 0 as lower limit and the value provided by the implementation of Condition 10.3 as upper limit.

The process is reported in Appendix D. In practice, firstly, the initial part of function $mu\_dev(k)$ must be run until the upper limit is computed and successively the entire script can be used for the calibration, returning $k(z)$ values.

Upper limits for k

The implementation of Condition 10.3 returned an array of values representing k_{u} that is the upper limit of k for each altitude z. As the probability of inundation P_{I} decreases with altitude, as well as the integral $\int_{\eta(x)}^{\infty}(h-\eta)p(h)dh$, k_{u} has an increasing trend, as it is shown in Figure 10.1. Consequently, the interval of valid k enlarges with z.

Resulting $k(z)$

Figure 10.1 also shows the calibrated values of the parameter $k(z)$, comparing them with the related upper limits $k_{\mathrm{u}}(z)$. Despite also k exponentially increases with altitude, it is always lower than the related upper limit. Considering that k is the ratio between the rate of vegetation decay and growth, it can also be expressed as:

$$k = \frac{K}{\alpha_2} \simeq \frac{T_{\mathrm{g}}}{2T_{\mathrm{d}}}, \tag{10.4}$$

where T_{g} and T_{d} are respectively the time needed by vegetation for an overall growth of 90% (i.e. from ν=0.05 to ν=0.95) and decay by 90% (i.e. from ν=0.95 to ν=0.05) [16]. Thus, the exponential growth of k physically means that the timescales of growth become much longer than timescales for decay when altitude increases. This means that inundations, despite their probability of occurrence decreases for high altitude, more easily affect vegetation growth and survival in the upper topographic bands when they occur. This behaviour can be explained by the lower degree of adaptation to floods for vegetation that has grown far from water surface.

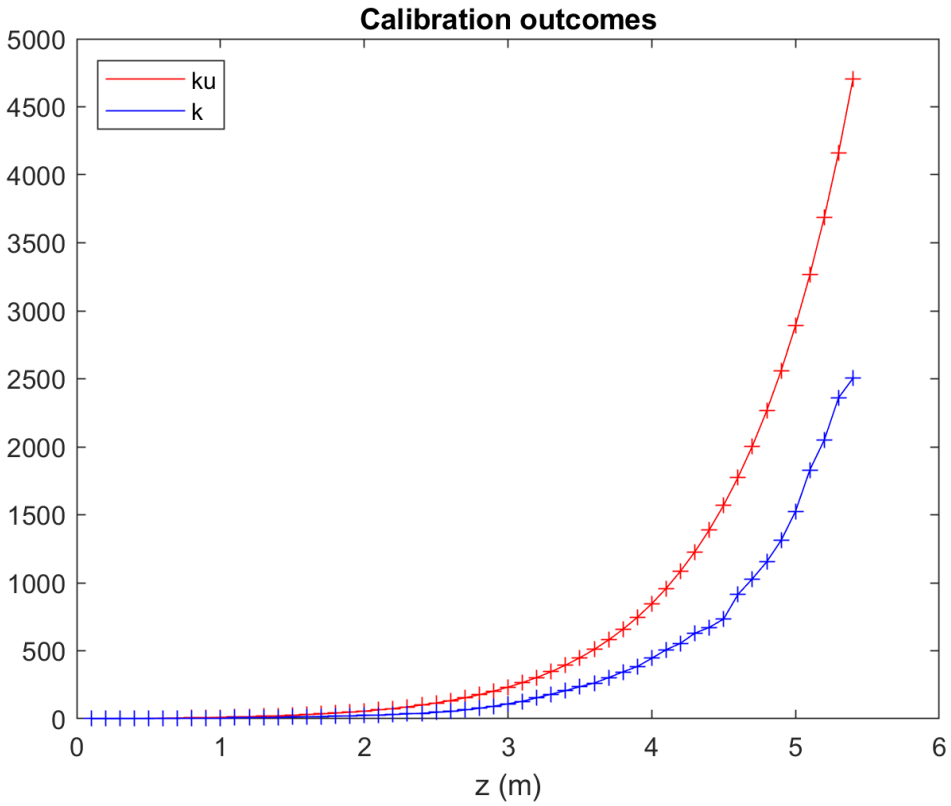

Figure 10.1. Trend of k for increasing altitude z and comparison with the upper limits k_{u}.

Part III

Chapter 11

Conclusion

Referring to the case study, which was a segment of Cinca River in Spain, two procedures were defined during this dissertation. The first one allows to estimate riparian biomass starting from raw LiDAR data, while the second one leads to the calibration of the stochastic model that describes dynamics of riparian vegetation.

After a short analysis about the state of the art of LiDAR technology, particularly focusing on the use of airborne laser scanner in forestry, a procedure for processing LiDAR raw data and return above-ground biomass was delineated. The first step of this method consists in the combined use of GIS software and FUSION/LDV to obtain statistics about vegetation height and coverage and it can be easily implemented once the command-line for FUSION/LDV are properly set. The second step allows to transform these statistics in biomass by implementing Means et al. (2000) model. This model was chosen after having tested its reliability in riparian environments by comparing the results it provides with the ones computed with allometric formulas. In addition, it demonstrated to provide reliable results when the study area has complex and narrow shapes, regardless the chosen grid discretization.

The procedure for stochastic model calibration relies on the use of LiDAR-derived information (i.e. biomass estimates and topography) and hydrological data. LiDAR data are processed with both GIS and FUSION/LDV software in order to return topography, while also Means et al. formula is used to obtain biomass. Hydrological data are processed with HEC-RAS and MATLAB to obtain statistics of water levels. Calibration is done according to a least square process that minimizes the deviations among computed and measured dimensionless first moments of vegetation. The results of calibration is an array of one k value for each topographic classes.

Follow-up activities will consist of a in situ survey for the validation of the method for biomass estimation. Also, the outcomes of calibration will be evaluated and the method will be further improved in order to return consistent second moments and probability distribution functions of vegetation for each location of the riparian corridor.

Appendix A

Regression of water surface plane

```
%MATLAB SCRIPT FOR REGRESSING THE INTERPOLATION PLANE

%xcoord, ycoord and zcoord are vectors containing respectively
%the x-coordinates, y-coordinates and z-coordinates of the points
%in the 3D cloud representing Cinca River water surface.

%exp is the matrix of exponents for the generic surface equation:
%z=coeff(1)*x^m(1,1)*y^m(1,2)+...+coeff(n)*x^m(n,1)*y^m(n,2)
%In the case of a plane, the surface equation is:
%z=coeff(1)+coeff(2)*x+coeff(3)*y
%that corresponds to the generic equation having:
%exp=[0 0; 1 0; 0 1]

%Matrix_X, Matrix_Y are the coordinates of points belonging to
%the regressed plane that will be interpolated in GIS environment
%to create a plane DTM.

%z_new is the expected outcome, or rather the z-coordinates of
%points belonging to the plane.

%O is the matrix containing the x,y,z coordinates of points
%constituting the regression plane.


%WATER SURFACE POINTS ACQUISITION AND VISUALIZATION

x=xcoord;
y=ycoord;
z=zcoord;
```

```
figure(1)
scatter3(x,y,z)
title('Water surface')
xlabel('x')
ylabel('y')
zlabel('z')

%LEAST SQUARE REGRESSION

exp=[0 0; 1 0; 0 1];
i=length(exp);
for s=1:i
    num(s)=sum(z.*x.^exp(s,1).*y.^exp(s,2));
    for r=1:i
        den(s,r)=sum(x.^exp(r,1).*y.^exp(r,2).*x.^exp(s,1).*y.^exp(s,2));
    end
end
coeff=(num/den)';

%as division does not exist for matrices, the symbol /
%indicates the solution to the system of equations: coeff*den=num

%DEFINITION AND VISUALIZATION OF POINTS BELONGING TO THE PLANE

x_new=Matrix_X;
y_new=Matrix_Y;
dim=size(x_new)';
N=dim(1);
M=dim(2);
for j=1:N;
    for k=1:M;
        z_new(j,k)=coeff(1)+coeff(2)*x_new(j,k)+coeff(3)*y_new(j,k);
    end
end

figure(2)
surf(x_new,y_new,z_new)
title('Regression Plane')
xlabel('x')
ylabel('y')
zlabel('z')

O=[x_new(:) y_new(:) z_new(:)];
```

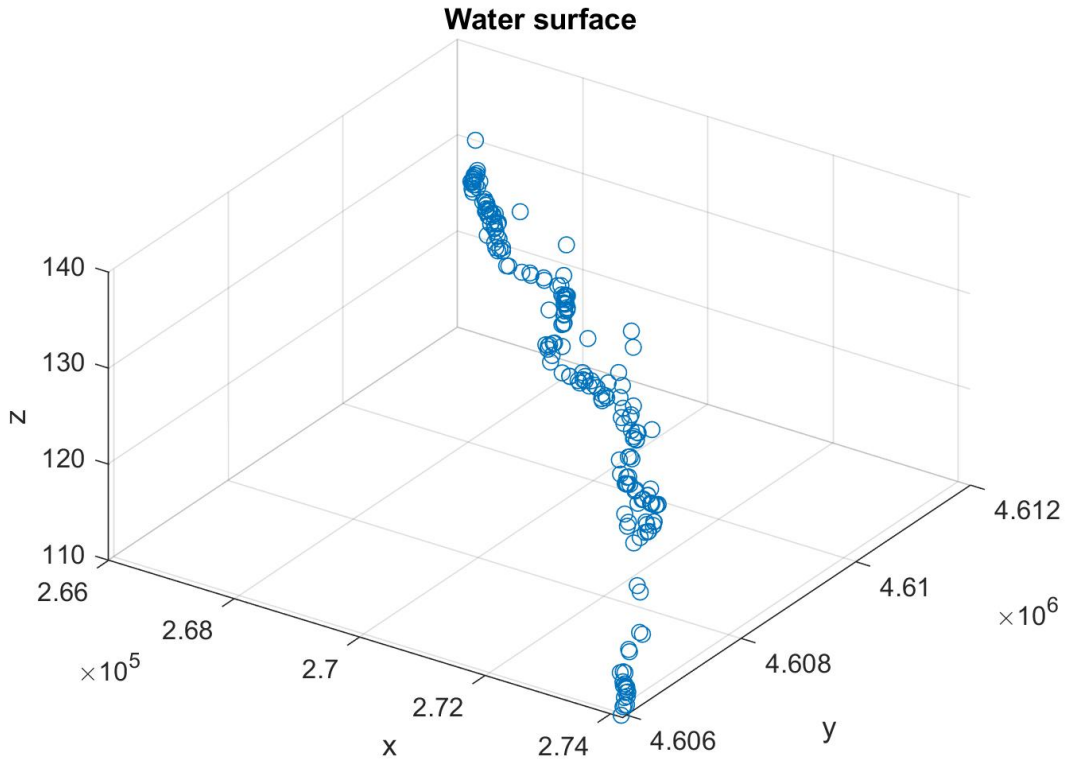

Figure A.1. Points belonging to the water surface at the time of acquisition, selected in GIS environment.

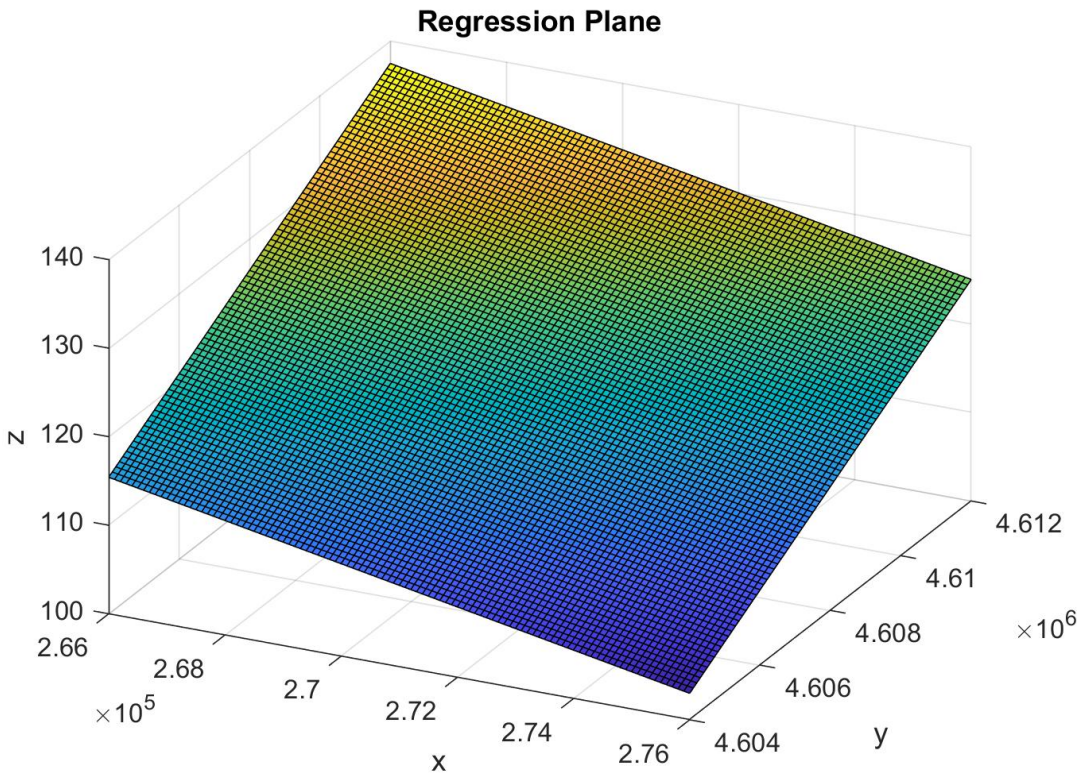

Figure A.2. Regression plane representing the water surface of Cinca River at the time of acquisition.

Appendix B

Computing biomass statistics

```
   %MATLAB SCRIPT FOR COMPUTING VEGETATION CARRYING CAPACITY

%The equivalent carrying capacity (Vc*), expressed in Mg/ha, must be
%computed to convert dimensional above-ground biomass in dimensionless
%values.

%The maximum vegetation height (Hmax) is firstly converted in maximum DBH
%(Dmax) with Trorey (or Ker & Smith) formula:
%Hmax= b1+b2*DmaxH-b3*Dmax^2 --> b3*Dmax^2-b2*Dmax+(-b1+Hmax)=0
%The input parameters are:
b1=137; %cm
b2=(28.42+24.15)/2;
b3=(0.0388+0.1348)/2;

%Once Dmax is known, Jenkins et al. formula can be applied:
%Vc*_dry=exp(beta0+beta1*ln(Dmax))
%The input parameters are:
beta0=-2.2094;
beta1=2.3867;

%As this formula returns the dry above-ground biomass, this is converted in
%fresh above-ground biomass, knowing that the ratio between fresh and dry
%wood is (1+theta), where theta is the moisture content. In this conversion
%solely the wood fraction of above-ground dry biomass is considered (i.e.
%foliage is neglected) --> solely the 97% of Vc*_dry is converted.
theta=1.055;
percentage=0.97;

%Local carrying capacity is computed for each topography band, ranging from
%0 to 5.4 m. Altitude is expressed in relation to the
```

```
%inclined plane that represents the average water surface at the time of
%LiDAR acquisition --> z_max=54
i_max=11;
z_max=54;


%Import of elevation statistics computed with FUSION
delimiterIn=',';  %to import FUSION output in MATLAB
headerlinesIn=1;  %to import FUSION output in MATLAB
for z = 1:z_max;
    Hmax(z)=0;

    for i=1:i_max;
        v_star=[];
        Sum_Average=0;
        Sum_StDev=0;
        Sum_Average_lognorm=0;
        Sum_StDev_lognorm=0;
        Sum_Skew=0;
        Sum_Skew_lognorm=0;
        Sum_Hmax=0;
        Count=0;
        formatSpec = '%s%d%s%d%s%d%s';
        A1 = '...directory...\';
        A2 = i;
        A3 = '\';
        A4 = i;
        A5 = '_DN';
        A6 = z;
        A7 = '_all_returns_elevation_stats.csv';
        filename = sprintf(formatSpec,A1,A2,A3,A4,A5,A6,A7);
        A = importdata(filename,delimiterIn,headerlinesIn);
        l=length(A.data);
        for m=1:l;
            if A.data(m,8)~=-9999
               Sum_Hmax=Sum_Hmax+A.data(m,7);
               Count=Count+1;
            end
        end

       if (Sum_Hmax/Count)*100> Hmax(z)
          Hmax(z)=(Sum_Hmax/Count)*100; %cm
       end
```

```
    end %i


    %Computation of Vc*
    %solely the higher Dmax solution is considered:
    Dmax(z,1)=(b2+sqrt((b2^2)-4*b3*(-b1+Hmax(z))))/(2*b3);
    Vc_dry=exp(beta0+beta1*log(Dmax(z,1)));
    Vc(z,1)=Vc_dry*(1+theta)*percentage/1000; %Mg/ha

end %z
```

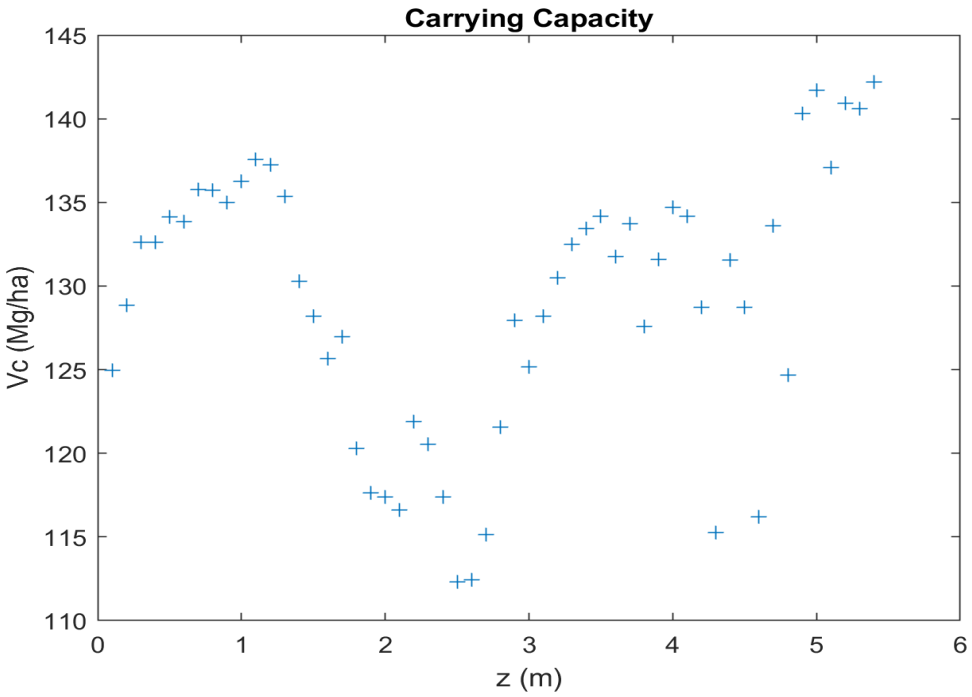

Figure B.1. Trend of carrying capacity for increasing altitude in the study area.

```
%MATLAB SCRIPT FOR COMPUTING VEGETATION STATISTICS

%Above-ground biomass (AB)is computed with the formula provided by
%Means et al. 2000: %ln(V)=2.532+0.05651*h_80+2.355*(CC/100)-0.1581*h_0
%Where: V is above-ground vegetation volume expressed in kg/ha,
%h_80, h_0 are percentiles of vegetation height, expressed in meters,
```

```
%CC is the canopy cover, expressed as percentage.

%In order to convert V in AB, the equivalent biomass density (rho),
%expressed in Mg/m3, is used.
rho=0.74;

%In order to convert AB in dimensionless biomass (v), AB is normalized in
%relation to the equivalent carrying capacity (Vc*), expressed in Mg/ha.
Vc=142.5;

%AB statistics must be computed for each of the 11 reaches --> i_max=11
%For each reach, topography was split in 54 altitude band, every 0.1
%meters, ranging from 0 to 5.4 m. Altitude is expressed in relation to the
%inclined plane that represents the average water surface at the time of
%LiDAR acquisition --> z_max=54
i_max=11;
z_max=54;

%Average and standard deviation values are obtained by a weighted average
%of the statistics of each reach. The weighting factors depends on the
%extension of the altitude band in each reach.

Average_z=zeros(z_max,1);
St_Dev_z=zeros(z_max,1);
Average_lognorm_z=zeros(z_max,1);
St_Dev_lognorm_z=zeros(z_max,1);

%pdf for the various altitude bands and reaches constitute a 2D array
%having 54 rows,
%corresponding to the altitude bands, and 101 columns, corresponding to the
%values of pdf for each step of the graduate axis representing
%dimensionless biomass. This axis ranges from 0 to 1 and has a step every
%0.01.
v_axis=[0:0.01:1];
m_max=101;
pdf_norm=zeros(z_max,m_max);
pdf_logn=zeros(z_max,m_max);


%Import of elevation statistics computed with FUSION
delimiterIn=',';  %to import FUSION output in MATLAB
headerlinesIn=1;  %to import FUSION output in MATLAB
for z = 1:z_max;
```

```
    Count=0;
    Tot=0;
    v_star=[];
    v=[]

for i=1:i_max;

    Count_zi(z,i)=0;
    formatSpec = '%s%d%s%d%s%d%s';
    A1 = '...directory...';
    A2 = i;
    A3 = '\';
    A4 = i;
    A5 = '_DN';
    A6 = z;
    A7 = '_all_returns_elevation_stats.csv';
    filename = sprintf(formatSpec,A1,A2,A3,A4,A5,A6,A7);
    A = importdata(filename,delimiterIn,headerlinesIn);
    l(i)=length(A.data);
    if i==1
        for m=1:l(i);
            if A.data(m,8)~=-9999
                Count=Count+1;
                Count_zi(z,i)=Count_zi(z,i)+1;
                v_star(Count,1)=(exp(2.532+0.05651*A.data(m,35)+
                2.355*(A.data(m,49)/100)-0.1581*A.data(m,6)))*rho;
                %Conversion of v_star (dimensional) in v (dimensionless)
                v(Count,1)=v_star(Count,1)/Vc;
            end
        end

        if Count_zi(z,i)~=0
            %Average
            Average(z,i)=mean(v(1:length(v)));
            Average_lognorm(z,i)=mean(v(1:length(v)));
            %Standard Deviation
            St_Dev(z,i)=std(v(1:length(v)));
            St_Dev_lognorm(z,i)=std(v(1:length(v)));


            ll=length(v);
            Tot=Tot+Count_zi(z,i);
```

```
            %Average
            Average_z(z,1)=Average_z(z,1)+(Count_zi(z,i))*Average(z,i);
            Average_lognorm_z(z,1)=Average_lognorm_z(z,1)
            +(Count_zi(z,i))*Average_lognorm(z,i);
           %Standard Deviation
            St_Dev_z(z,1)=St_Dev_z(z,1)+(Count_zi(z,i))*St_Dev(z,i);
            St_Dev_lognorm_z(z,1)=St_Dev_lognorm_z(z,1)
            +(Count_zi(z,i))*St_Dev_lognorm(z,i);
        end

    else for m=1:length(A.data);
            if A.data(m,8)~=-9999
                Count=Count+1;
                Count_zi(z,i)=Count_zi(z,i)+1;
                v_star(Count,1)=(exp(2.532+0.05651*A.data(m,35)+
                2.355*(A.data(m,49)/100)-0.1581*A.data(m,6)))*rho;
                %Conversion of v_star (dimensional) in v (dimensionless)
                v(Count,1)=v_star(Count,1)/Vc;
            end
        end
                    if Count_zi(z,i)~=0
            %Average
            Average(z,i)=mean(v(ll+1:length(v)));
            Average_lognorm(z,i)=mean(log(v(ll+1:length(v))));
            %Standard Deviation
            St_Dev(z,i)=std(v(ll+1:length(v)));
            St_Dev_lognorm(z,i)=std(log(v(ll+1:length(v))));

        ll=length(v);

        Tot=Tot+Count_zi(z,i);
        %Average
        Average_z(z,1)=Average_z(z,1)+(Count_zi(z,i))*Average(z,i);
        Average_lognorm_z(z,1)=Average_lognorm_z(z,1)
        +(Count_zi(z,i))*Average_lognorm(z,i);
       %Standard Deviation
        St_Dev_z(z,1)=St_Dev_z(z,1)+(Count_zi(z,i))*St_Dev(z,i);
        St_Dev_lognorm_z(z,1)=St_Dev_lognorm_z(z,1)
        +(Count_zi(z,i))*St_Dev_lognorm(z,i);
        end%if
    end
end
```

```
     %Average
    Average_z(z,1)=Average_z(z,1)/Tot;
    Average_lognorm_z(z,1)=Average_lognorm_z(z,1)/Tot;
    %Standard Deviation
    St_Dev_z(z,1)=St_Dev_z(z,1)/Tot;
    St_Dev_lognorm_z(z,1)=St_Dev_lognorm_z(z,1)/Tot;
    %pdf
    pdf_norm(z,:)=normpdf(v_axis,Average_z(z,1),(St_Dev_z(z,1)));
    pdf_logn(z,:)=lognpdf(v_axis,Average_lognorm_z(z,1),St_Dev_lognorm_z(z,1));

end%z
```

Appendix C

Computing hydrometric statistics

```
%MATLAB SCRIPT FOR COMPUTING HYDROMETRIC STATISTICS

%Reference discharges through which computing the rating curves
Q_ref=Reference_Q; %m3/s, imported

%Number of reaches
b=11;

%Water levels h computed with HEC-RAS for the reference discharges
h_ref=h_hecras';


%CREATION OF A RATING CURVE FOR EACH REACH

%Assuming the relation between h and Q is Q=A*h^B:
    %log10(Q)=log10(A)+B*log10(h)
    %logq=log10(Q), logh=log10(h), a=log10(A) --> q=a+Bh
% Coeff is the matrix of coefficients A and B
% Corr is the R factor for the computing rating curves

%Conversion of h and Q variables in their logarithms
for j=1:b;
    for k=1:length(Q_ref);
    logh(k,j)=log10(h_ref(k,j));
    end
end
logq=log10(Q_ref);

%Fitting of the curves
for j=1:b;
```

```
    p=polyfit(logh(:,j),logq,1);
    Coeff(1,j)=10^p(1,2);
    Coeff(2,j)=p(1,1);

    for k=1:length(Q_ref)
        h_calc(k,j)=(Q_ref(k)/Coeff(1,j))^(1/Coeff(2,j));
    end
    Corr(1,j)=corrcoef(h_ref(j), h_calc(j));
end




%COMPUTATION OF A TEMPORAL SERIES OF WATER LEVELS (h) FOR EACH REACH

Q=Q_series;
%h=rand(length(Q),b);
for j=1:b;
    for k=1:length(Q);
        h_star(k,j)= (Q(k)/Coeff(1,j))^(1/Coeff(2,j)); %meters

    end

    %scaling water level in relation to the vertical datum, set as the
    %minimum water level
    h_min(1,j)=min(h_star(:,j));
    h_star(:,j)=h_star(:,j)-h_min(1,j);

end


%Q_series is the historic series of Cinca discharges




%COMPUTATION OF STATISTICS FOR EACH REACH

    %mean and standard deviation
for j=1:b;
    Average_star(1,j)=mean(h_star(:,j));
    Std_star(1,j)= std(h_star(:,j));
end
```

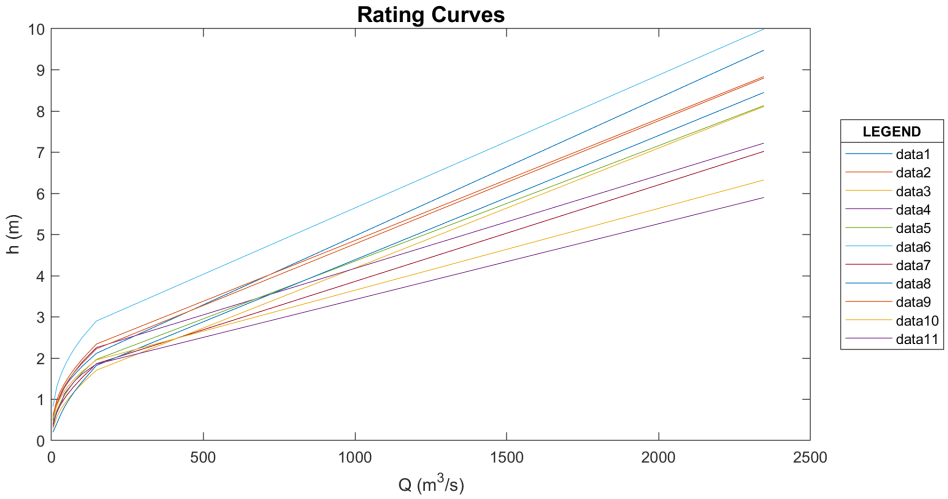

Figure C.1. Rating curves for the eleven reaches.

```
for j=1:b;
    %lognormal distribution
    h_axis=[0:0.1:10]';
    ln_h_star(:,j)=log(h_star(:,j));
    ind=0;
    sum_av=0;
    sum_ds=0;
    for k=1:length(ln_h_star);
        if ln_h_star(k,j)~= -Inf
            ind=ind+1;
            sum_av=sum_av+ln_h_star(k,j);
        end
    end
    ln_Average(1,j)=sum_av/ind;

    for k=1:length(ln_h_star);
        if ln_h_star(k,j)~= -Inf
            sum_ds=sum_ds+(ln_h_star(k,j)-ln_Average(1,j))^2;
        end
    end
    ln_St_Dev(1,j)=sqrt(sum_ds/ind);
    p_h_logn_star(:,j)=lognpdf(h_axis, ln_Average(1,j), ln_St_Dev(1,j));
    cdf_logn_star(:,j)=cdf('logn',h_axis, ln_Average(1,j), ln_St_Dev(1,j));

    %Normalization of dimensional h_star in h and visualization of distributions
```

```
     h_ax= (h_axis-Average_star(1,j))/Average_star(1,j);
    figure(2)
    plot(h_ax,p_h_logn_star)
    title('LogNormal probability distribution function of the water level')
    xlabel('h (m)')
    ylabel('p(h)')
    axis([-1 10 0 1])
    hold on
    figure(3)
    plot(h_ax,cdf_logn_star)
    title('LogNormal cumulative distribution function of the water level')
    xlabel('h (m)')
    ylabel('P(h)')
    axis([-1 10 0 1])
    hold on
    figure(6)

end
```

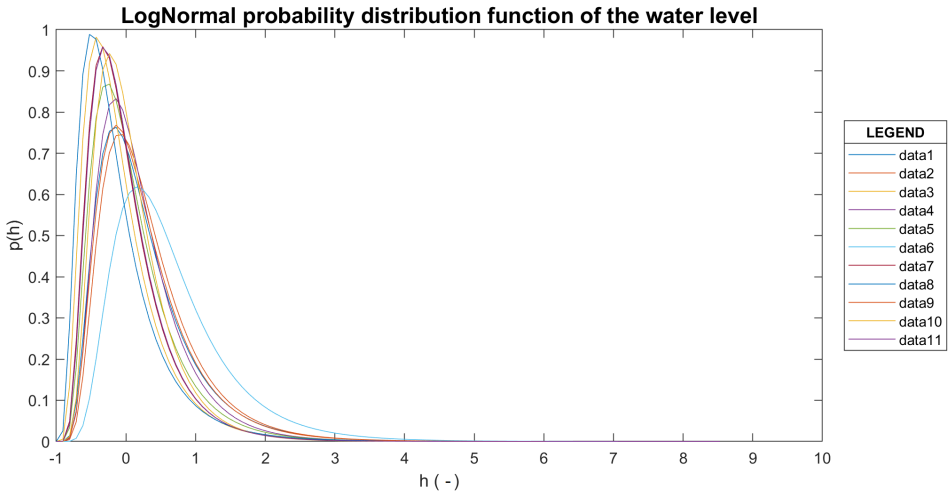

Figure C.2. Computed pdf for the eleven reaches.

```
%Autocorrelation and integral scale
tau=ones(1,b);
lags=1000; %values fixed after some attempts
Q_tot=Q_complete';%import data from SAIHEbro
```

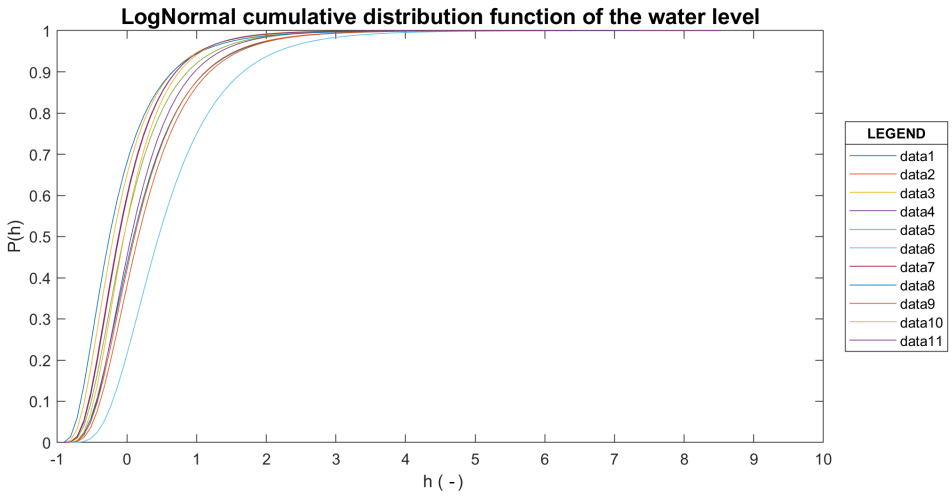

Figure C.3. Computed CDF for the eleven reaches.

```
for j=1:b;
    for k=1:length(Q_tot);
        h_tot(k,j)= (Q_tot(k)/Coeff(1,j))^(1/Coeff(2,j));
    end
end

for j=1:b;
    xmax=0;
    figure(3+j) %to create a figure for each reach
    autocorr(h_tot(:,j),'NumLags',lags,'NumSTD',2);
    title('LogNormal probability distribution function of the water level')
    xlabel('s* (d)')
    ylabel('rho(s*) (d)')
    y(:,j)=autocorr(h_tot(:,j),'NumLags',lags,'NumSTD',2);
    k=1;
   while k<lags
          if y(k,j)<0.2 %threshold for computing the integral scale
             xmax=k;
             k=lags;
          else
              k=k+1;
          end
    end
    for k=1:xmax;
        tau(1,j)=tau(1,j)+y(k,j);
```

```
    end
end
```

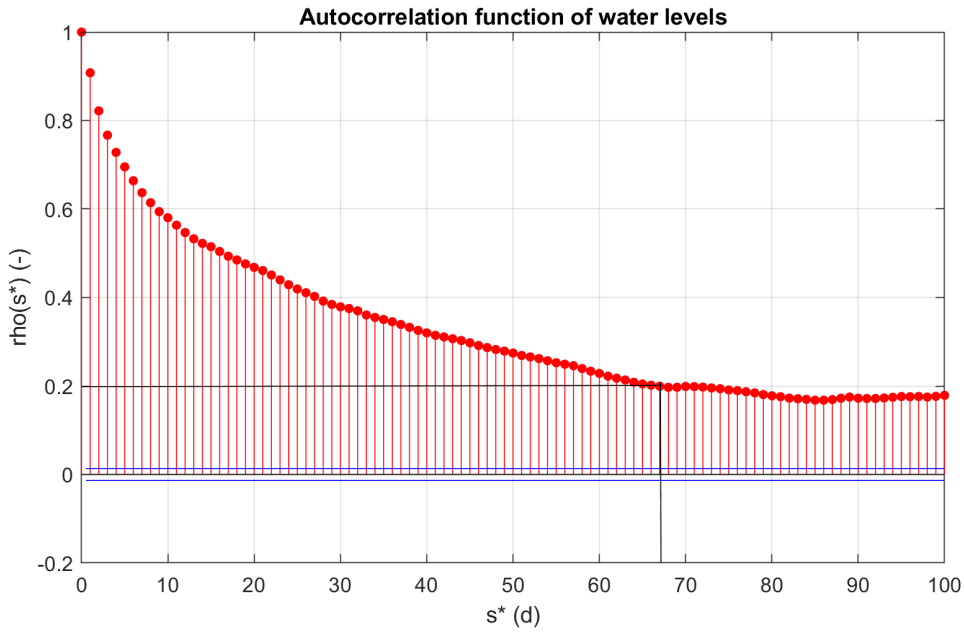

Figure C.4. Computed autocorrelation function for the first reach.

Appendix D

Stochastic model calibration

```
%SCRIPT FOR THE STOCHASTIC MODEL CALIBRATION

%Once the altitude z is set, the MATLAB function @fminbnd is used to define
%the value of k that minimizes the function @min_mu, or rather that
%minimizes the deviation among LiDAR-derived values and computed ones.

%import of upper limit of k
lim=importdata('...directory...\condition_ku.mat');

for zz=1:54; %amount of altitude bands
    save('..directory..\zz.mat','zz')
    k(zz,1)=fminbnd(@min_dev,0,lim(zz));
end



%SCRIPT OF THE FUNCTION min_dev(k)

function [mu_z]=min_dev(k)

%study area comprises 11 subreaches and 54 altitude bands
i_max=11;
z=importdata('...directory...\zz.mat');

ind_j=0;

%INPUT VEGETATION

%import of the outcomes of "MATLAB SCRIPT FOR COMPUTING VEGETATION
%STATISTICS"
mu_lidar=importdata('...directory...\Average_veg.mat');
```

```
%parameters linked to vegetation features
W=importdata('...directory...\W.mat'); %weighting factors for
%vegetation according to the extension of the z altitude band in
%each reach
alpha2=20.5*10^(-5);
beta=1;

%INPUT HYDROLOGY

%The mean value (h_mean), pdf (p_h) and integral scale (tau) of
%dimensionless water levels have been computed with
%"SCRIPT FOR COMPUTING HYDROMETRIC STATISTICS"
W_idr=importdata('...directory...\W_idr.mat');%weighting factors for
%hydrological data according to the reach extension

%pdf
load('ph');
load('tau_star');
for l=1:length(ph);
    ph(l,1)=0;
    for i=1:11
        ph(l,1)=ph(l,1)+Pesi_idr(1,i)*ph(l,i);

    end
end
%integral scale
tau=tau_star*alpha2;
Tau=0;
for i=1:11;
    Tau=Tau+tau(1,i)*Pesi_idr(1,i);
end
tau=tau;

%h_axis
Q_acq=31.3; %m3/s
Q_min=1; %m3/s
load('h_mean);
load(h_axis);
shift_average=0;
h_m=0;
for i=1:i_max;
    water_level_acq(1,i)=(Q_acq/Coeff(1,i))^(1/Coeff(2,i));
```

```
    water_level_min(1:i)=(Q_min/Coeff(1,i))^(1/Coeff(2,i));
    shift(1,i)=(water_level_acq(1,i)-water_level_min(1,i))*Pesi_idr(1,i);
    shift_average=shift_average+shift(1,i);
    h_m=h_m+h_mean(1,i)*Pesi_idr(1,i);%weighted averaged water level
end%i

%INPUT GEOMETRY
altitude=z*0.1; %meters above average water surface
eta_star=(z*0.1+shift_average);
eta=(eta_star-h_m/h_m);

%COMPUTATION
dev=0; %deviation

%Probability of inundation (Pi)
Pi=0;
for n=1:length(h_axis)
    if h_axis(n)<eta
        ind=n;
    else  n=length(h_axis+1);
    end
end
for n=(ind+1):length(h_axis);
    Pi=Pi+ph(n,1)*0.1;
end

%Decay rate (alpha1/k)
alpha=0;
sum2=0;
for n=(ind+1):length(h_axis);
    g=(h_ax(n,1)-eta)*ph(n,1);
    sum2=sum2+g*0.1;
end
alpha=sum2/Pi;

condition_ku(z,1)= (beta/alpha)*(1/Pi-1);

%mu1 and deviation
if Pi < (beta/(k*alpha+beta))
    rho0=k*alpha+beta;
    rho1=gamma((1-Pi)/(k*alpha*Tau));
    rho2=gamma((1-Pi)/(k*alpha*Tau)-Pi/(beta*Tau));
    sigma1=gamma((1-Pi)/(k*alpha*Tau)+1);
```

```
    sigma2=gamma((1-Pi)/(k*alpha*Tau)+2);
    epsilon1=gamma((1-Pi)/(k*alpha*Tau)-Pi/(beta*Tau)+1);
    epsilon2=gamma((1-Pi)/(k*alpha*Tau)-Pi/(beta*Tau)+2);

    mu1(z,1)=((beta*rho1*(beta*epsilon2*sigma1-epsilon1*rho0*sigma2))/
    ((beta*epsilon1*rho1-rho0*rho2*sigma1)*sigma2));

    dev=((mu_lidar(z,1)-mu1(z,1)))^2;

else dev=9999;

end

mu_z=scarto;

end
```

0.9

Bibliography

[1] Centro de descargas - cnig . [accessed on may 2018]. http://centrodedescargas.cnig.es/CentroDescargas/index.jsp.

[2] The cloud-aerosol lidar and infrared pathfinder satellite observation (calipso) [accessed on august 2018]. https://www-calipso.larc.nasa.gov/.

[3] Confederación hidrografíca del ebro. [accessed on may 2018]. http://www.saihebro.com/saihebro/index.php?url=/principal.

[4] Icesat cryospheric science lab code 615 [accessed on august 2018]. https://icesat.gsfc.nasa.gov/icesat/glas.php.

[5] Lunar laser ranging instrument (llri) [accessed on august 2018]. https://nssdc.gsfc.nasa.gov/nmc/experimentDisplay.do?id=2008-052A-03.

[6] The mars orbiter laser altimeter [accessed on august 2018]. https://attic.gsfc.nasa.gov/mola/.

[7] Plan nacional de ortofotografía aérea. [accessed on may 2018]. http://pnoa.ign.es/.

[8] Proyecto de adecuación de ribera y margen del ebro en el ámbito: U9, echegaray pilar, 2005. anejo n. 8. https://www.zaragoza.es/cont/paginas/grandesproyectos/pdf/riberasu9/Anejo8.pdf.

[9] Qgis a free and open source geographic information system. [accessed on may 2018]. hhttps://qgis.org/en/site/index.html.

[10] J. B. Abshire. Nasa's space lidar measurements of earth and planetary surfaces. technical report. [accessed on august 2018].

[11] D. Andújar, A. Escolá, J. R. Rosell-Polo, C. Fernández-Quintanilla, and J. Dorado. Potential of a terrestrial lidar-based system to characterise weed vegetation in maize crops. *Computers and Electronics in Agriculture*, 92:11–15, 2013.

[12] T. J. Battin, L. A. Kaplan, S. Findlay, C. S. Hopkinson, E. Marti, A. I. Packman, D. Newbold J, and F. Sabater. Biophysical controls on organic carbon fluxes in fluvial networks. *Nature Geoscience*, 1:95–100, 2008.

[13] T. J. Battin, S. Luyssaert, L. A. Kaplan, A. K. Aufdenkampe, A. Richter, and L. J. Tranvik. The boundless carbon cycle. *Nature Geoscience*, 2:598–600, 2009.

[14] R. Brennan and T. L. Webster. Object-oriented land cover classification of lidar-derived surfaces. *Canadian Journal of Remote Sensing*, 32(2).

[15] S. Brown, J. Sathaye, M. Cannell, and P. Kauppi. *Management of forests for mitigation of greenhouse gas emissions. In: Watson, R.T., Zinyowera, M.C., Moss,*

R.H. (Eds.), Climate Change 1995. Impacts, Adaptation and Mitigation of Climate Change: Scientific-technical Analyses. Contribution of WG II to the Second Assessment Report of the IPCC. Cambridge University Press, Cambridge, pp. 773-797.

[16] C. Camporeale and L. Ridolfi. Riparian vegetation distribution induced by river flow variability: a stochastic approach. *Water Resources Research*, 42(W10415), 2006.

[17] A. P. Charaniya, R. Manduchi, and S. K. Lodha. Supervised parametric classification of aerial lidar data. *In CVPRW−04, Proceedings of the IEEE 2004 Conference on Computer Vision and Pattern Recognition Workshop, 27 June â 2 July 2004, Baltimore, Md.*, 3:1–8, 2004.

[18] V. T. Chow. *Open-channel hydraulics*. McGRAW−HILL, New York, 1959.

[19] J. J. Cole, adn N. F. Caraco Y. T. Prairie, W. H. McDowell, L. J. Tranvik, R. G. Striegl, C. M. Duarte, P. Kortelainen, J. A. Downing, J. J. Middelburg, and J. Melack. Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems*, 10:171–184, 2007.

[20] J. Estornell Cremades, and B. Velázquez Martí LÁ. Ruiz Fernández, and T. Hermosilla. Analysis of the factors affecting lidar dtm accuracy. *International Journal of Digital Earth*, 4(6):521–538, 2011.

[21] P. J. Edwards, J. Kollmann, A. M. Gurnell, G. E. Petts, K. Tockner, and J. V. Ward. A conceptual model of vegetation dynamics on gravel bars of a large alpine river. *Wetlands Ecology and Management*, 7(3):141â153, 1999.

[22] J. U. H. Eitel, T. S. Magney, L. A. Vierling, T. T. Brown, and D. R. Huggings. Lidar based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status. *Field Crops Research*, 159:21–32, 2014.

[23] G. Fiocco and L. Smullin. Detection of scattering layers in the upper atmosphere (60-140 km) by optical radar. *Nature*, 199:1275–1276, 1963.

[24] M. Flood. Laser altimetry: from science to comercial lidar mapping. *Photogrammetric Engineering Remote Sensing*, 67(11):1209–1217, 2001.

[25] New Zealand Ministry for the Environment. *New Zealand Greenhouse Gas Inventory 1990-2015*. New Zealand Government, Wellington, 2017.

[26] M. E. Galvez and J. Gaillardet. Historical constraints on the origins of the carbon cycle concept. *Comptes Rendus Geoscience*, 344(11-12).

[27] M. García, D. Riaño, E. Chuvieco, and F. M. Danson. Estimating biomass carbon stocks for a mediterranean forest in central spain using lidar height and intensity data. *Remote Sensing of Environment*, 114:816–830, 2010.

[28] H. Hasegawa. Evaluation of lidar reflectance amplitude sensitivity towards land cover conditions. *Bulletin of the Geographical Survey Institute*, 53(6).

[29] Q. He, E. Chen, R. An, and Y. Li. Above-groundbiomass and biomass components estimation using lidar data in a coniferous forest. *Forests*, 4:984–1002, 2013.

[30] J. C. Jenkins, D. C. Chojnacky, L.S. Heath, and R. A. Birdsey. National-scale biomass estimators for united states tree species. *Forest Science*, 49(1):12–35, 2003.

[31] J. Ker and J. Smith. Advantages of the parabolic expression of height-diameter relationship. *The Forestry Chronicle*, 31:235–246, 1955.

[32] V. Klemas. Beach profiling and lidar bathymetry: an overview with case studies. *Journal of Coastal Research*, 27(6):1019–1028, 2011.

[33] K. Kraus and N. Pfeifer. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53:193â203, 1998.

[34] M. A. Lefsky, W. B. Cohen, D. J. Harding, G. G. Parker, S. A. Acker, and S. T. Gower. Lidar remote sensing of aboveground biomass in three biomes. *Photogrammetry and Remote Sensing*, XXXIV(3):155–160, 2002.

[35] M. Lemmens. *Terrestrial Laser Scanning. In: Geo-information. Geotechnologies and the Environment, vol 5.* Springer, Dordrecht, 2011.

[36] Y. Li, H. E. Andersen, and R. McGaughey. A comparison of statistical metods for estimating forest biomass from light detection and ranging data. *Western Journal of Applied Forestry*, 23.

[37] T. H. Maiman. Stimulated optical radiation in ruby. *Nature*, 187(4736).

[38] R. T. McGaughey. Fusion/ldv: Software for lidar data analysis and visualization. *USDA - United States Department of Agriculture. Forest Service. Pacific Northwest Research Station*, 2016.

[39] M. McGill. *LiDAR remote sensing - Submitted to the Encyclopedia of Optical Engineering.* NASA Goddard Space Flight Center - Laboratory for Atmospheres, Greenbelt (MD)., 2002.

[40] J. E. Means, S. A. Acker, B. J. Fitt, M. Renslow, L. Emerson, and C. J. Hendrix. Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering Remote Sensing*, 66(11):1367–1371, 2000.

[41] J. E. Means, S. A. Acker, D. J. Harding, J. B. Blair, M. A. Lefsky, W. B. Cohen, M. E. Harmon, and W. A. McKee. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western cascades of oregon. *Remote Sensing of Environment*, 67:298–308, 1999.

[42] B. Mitchell, M. Walterman, T. Mellin, C. Wilcox, A. M. Lynch, J. Anhold, D. A. Falk, J. Koprowski, D. Laes, D. Evans, and H. Fisk. *Mapping Vegetation Structure in the Pinaleño Mountains Using LiDAR Phase 3: Forest Inventory Modeling.* RSAC-10007-RPT1, Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center, 2012.

[43] J. Murillo, P. Brufau, and P. Garcia-Navarro. Aplicación del modelo bidimensional guad-2d para la determinación de zonas inundables en el t.m. de fraga (huesca). *Universidad de Zaragoza e INCLAM, S. A.*, 2008.

[44] E. Naesset. Estimation of above- and below-ground biomass in boreal forest ecosystems. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI(8/W2).

[45] National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center. *Lidar 101: An Introduction to Lidar Technology, Data, and Applications.* NOAA Coastal Services Center, Charleston, SC, 2012.

[46] US Army Corps of Engineers. *HEC-RAS. River Analysis System. Users Manual.* Institute for Water Resources. Hydrologic Engineering Center, Davis (CA), 2016.

[47] L. Pearlstine, H. McKellar, and W. Kitchens. Modelling the impacts of a river diversion on bottomland forest communities in the santee river floodplain, south carolina. *Ecological Modelling*, 29:283–302, 1985.

[48] J. Penman, M. Gytarsky, T. Hiraishi, T. Krug, D. Kruger, R Pipatti, L. Buendia, K. Miwa, T. Ngara, K. Tanabe, and F. Wagner. *Intergovernmental Panel on Climate Change - Good Practice Guidance for Land Use, Land-Use Change and Forestry.* Institute for Global Environmental Strategies (IGES) for the IPCC, Kanagawa (Japan), 2003.

[49] W. M. Post, T. H. Peng, W. R. Emanuel, A. W. King, V. H. Dale, and D. L. DeAngelis. The global carbon cycle. *American Scientist*, 78:310–326, 1990.

[50] N. H. Ravindranath and M. Ostwald. *Carbon Inventory Methods Handbook for Greenhouse Gas Inventory, Carbon Mitigation and Roundwood Production Projects.* Springer, Dordrecht, 2008.

[51] S. E. Reutebuch, R. J. McGaughey, H. E. Andersen, and W. W. Carson. Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing*, 29(5):527–535, 2003.

[52] D. M. Richardson, P. M. Holmes, K. J. Esler, S. M. Galatowitsch, J. C. Stromberg, S. P. Kirkman, P. Pysek, and R. J. Hobbs. Riparian vegetation: degradation, alien plant invasions, and restoration prospects. *Diversity and Distributions*, 13:126–139, 2007.

[53] A. L. Schawlow and C. H. Townes. Infrared and optical masers. *Physical Review*, 112(6):1940–1949, 1958.

[54] D. M. Smith, D. M. Finch, C. Gunning, R. Jemison, and J. F. Kelly. Post-wildfire recovery of riparian vegetation during a period of water scarcity in the southwestern usa. *Fire Ecology Special Issue*, 5(1):38–55, 2009.

[55] S. Solberg and E. Naesset. Mapping defoliation with lidar. *Proceedings of the SilviLaser Conference, 2007, Espoo, Finland, 12-14 September, International Archives of Photogrammetry and Remote Sensing,*, XXXVI(3-W52):379–382, 2007.

[56] J. H. Song, S. H. Han, K. Y. Yu, and Y. I. Kim. Assessing the possibility of land-cover classification using lidar intensity data. *In PCVâ02: Photogrammetric Computer Vision, ISPRS Commission III Symposium, 9â 13 September 2002, Graz, Austria*, page B259, 2002.

[57] C. Straub, J. Tian, R. Seitz, and P. Reinartz. Assessment of cartosat-1 and worldview-2 stereo imagery in combination with a lidar-dtm for timber volume estimation in a highly structured forest in germany. *Forestry*, 00:1–11, 2013.

[58] M. T. Ter-Mikaelina and M. D. Korzukhin. Biomass equations for sixty-five north american tree species. *Forest Ecology and Management*, 97:1–24, 1997.

[59] L. M. Tritton and J. W. Hornbeck. Biomass equations for major tree species of the northeast. *USDA Forest Service General Technical Report NE-69*, 1982.

[60] L.G. Trorey. A mathematical method for the construction of diameter height curves based on site. *The Forestry Chronicle*, 18(2):3–14, 1932.

[61] URS. *Estudio de la calidad ecológica integral de los tramos fluviales más importantes del Río Cinca - Atlas del río - Vol. III.* 2002.

[62] URS. *Estudio de la calidad ecológica integral de los tramos fluviales más importantes del Río Cinca - Informe de Síntesis - Vol. II.* 2002.

[63] Forest Products Laboratory (U.S.). *Wood handbook: wood as an engineering material.* U.S. Dept. of Agriculture, Forest Service, Forest Products Laboratory, Madison (WI)., 1999.

[64] M. Vastaranta, T. Kantola, and M. Holopainen P. Lyytikäinen-Saarenmaa, and J. Hyyppä V. Kankare, and M. A. Wulder, and H. Hyyppä. Area-based mapping of defoliation of scots pine stands using airborne scanning lidar. *Remote Sensing*, 5:1220–1234, 2013.

[65] R. Vesipa, C. Camporeale, and L. Ridolfi. Effect of river flow fluctuations on riparian vegetation dynamics: Processes and models. *Advances in Water Resources*, 110:29–50, 2017.

[66] A. R. Weiskittel, D. W. Hann, J. A. Kershaw, and J. K. Vanclay. *Forest Growth and Yield Modeling.* Wiley-Blackwell, John Wiley Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK, 2011.

[67] C. J. Willmott and K. Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research*, 30(?):79â82, 2005.

[68] M. A. Wulder, J. C. White, R. F. Nelson, E. Naesset, H. O. Orka, N. C. Coops, T. Hilker, C. W. Baker, and T. Gobakken. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, 121:196–209, 2012.