POLITECNICO DI MILANO

SCUOLA DI INGEGNERIA CIVILE E AMBIENTALE
Corso di studi in Ingegneria per l'Ambiente e il Territorio

ASSESSMENT OF THE USE OF PARTICLES' MORPHOMETRY FOR SEDIMENT FINGERPRINTING

Autore: Alessia Gurini (matr. 899773)
Relatore: Prof. Francesco Ballio

Correlatori: Prof. Mário J. Franca
PhD/ Lecturer Alessandro Cattapan

Accademico 2019/2020

Contents

Table of Figures

List of Tables

Abstract

In recent years, several models have been proposed to describe sediment transport at basin scale. Some use connectivity concepts to allow the assessment of the impacts associated with alternative water management scenarios. However, since connectivity cannot (to date) be directly measured in the field nor remotely sensed, the identification of a measurable property allowing the inference of some transport characteristics, e.g. source locations, travel distances, residence time, would represent a significant improvement for model calibration.

Recent studies on sediment attrition claim the existence of a "universal" relation between particles relative mass-loss (μ) and their circularity (C) and suggest that this relation might be useful for inferring the distance travelled (L) by individual particle (Novák-Szabó et al., 2018).

Since collecting data about the mass-loss of individual particles in the field is not feasible using techniques available to date, it would be advantageous to assess whether it would be possible to establish a theoretical relationship between μ and L.

Sternberg in 1875 suggested the exponential decrease of mass with the distance (Krumbein, 1941). Nevertheless, the coefficient which describe the speed with which the mass varies is unknown. Since the initial mass of the particles at the source points are unknown, an alternative could be the study of the exponential increase of μ with L. In this research, samples of metabasalts have been collected within the Sarzana River basin, located in North-East Italy. The digital images collected were then processed for deriving a series of shape characteristics. In order to overcome the limitations with respect to accuracy of segmentation in presence of shadows of available methods, a new segmentation model has been developed. This allowed the removal of the main effect produced by the shadows, reaching an accuracy of the model higher than 90%.

Finally, the abrasion coefficients (k) that describe the particles' mass degradation were estimated through a sequence of functions that relates C, μ and L.

A sensitivity analysis has also been performed to infer the influence of the uncertainty on the estimation of C and k on the estimated values of L.

This showed that, under the hypothesis of an accuracy of the estimations of circularity and abrasion coefficient of ~10^{-2}, the accuracy of travelled distance cannot fall below 20%. Furthermore, the application of a sequence of two increasing function revealed a relevant dispersion of the distances, leading to errors of more than one order of magnitude.

Estratto

Negli ultimi anni sono stati proposti diversi modelli per descrivere il trasporto dei sedimenti su scala di bacino. Alcuni utilizzano il concetto di connettività per valutare gli impatti associati a diversi scenari legati alla pianificazione e progettazione di interventi antropici. Tuttavia, poiché la connettività non può (ad oggi) essere misurata direttamente sul campo né percepita a distanza, l'identificazione di una proprietà misurabile che consenta l'inferenza di alcune caratteristiche di trasporto, ad es. luoghi di origine, distanze di viaggio, tempo di residenza, rappresenterebbe un miglioramento significativo per la taratura del modello.

Studi recenti sull'attrito dei sedimenti affermano l'esistenza di una relazione "universale" tra la perdita di massa relativa delle particelle (μ) e la loro circolarità (C) e suggeriscono che questa relazione potrebbe essere utile per dedurre la distanza percorsa (L) dalle singole particelle (Novák-Szabó et al., 2018). Dal momento che la raccolta di dati sulla perdita di massa di singole particelle in campo non è fattibile utilizzando tecniche finora disponibili, sarebbe vantaggioso valutare se sia possibile stabilire una relazione teorica tra μ e L. Sternberg nel 1875 suggerì la diminuzione esponenziale della massa con la distanza (Krumbein, 1941). Tuttavia, il coefficiente che descrive la velocità con cui la massa varia è sconosciuto. Poiché la massa iniziale delle particelle nei punti di origine è sconosciuta, un'alternativa potrebbe essere lo studio dell'aumento esponenziale di μ con L. In questa ricerca sono stati raccolti campioni di metabasalti all'interno del bacino del fiume Sarzana, situato nel Nord-Est Italia. Le immagini digitali raccolte sono state poi elaborate per ricavare una serie di caratteristiche di forma.

Al fine di superare i limiti legati alla precisione dei metodi di segmentazione ad oggi disponibili, è stato sviluppato un nuovo modello. Questo ha permesso la rimozione dell'effetto principale prodotto dalle ombre, raggiungendo un'accuratezza del modello superiore al 90%.

Infine, i coefficienti di abrasione (k) che descrivono l'erosione di massa delle particelle sono stati stimati attraverso una sequenza di funzioni che si riferiscono a C, μ e L. È stata inoltre effettuata un'analisi di sensibilità per dedurre l'influenza dell'incertezza sulla stima di C e k sui valori stimati di L. Ciò ha dimostrato che, nell'ipotesi di una precisione delle stime della circolarità e del coefficiente di abrasione di ~10^{-2}, la precisione della distanza percorsa non può scendere al di sotto del 20%. Inoltre, l'applicazione di una sequenza di due funzioni crescenti ha dimostrato una rilevante dispersione delle distanze stimate, portando ad errori di più di un ordine di grandezza.

1. Introduction

A river basin is a complex and dynamic system, whose morphology changes continuously as a result of differences of sediment transport rates in time and space.

Each basin is defined as the surface of the land that is drained by a river and its tributaries: this means that theoretically every drops falling inside the basin, if moving along the direction of steepest descent, is expected to exit the basin through its closing section.

It is important to consider that the response of the basin, in terms of peak discharge and sediment transport, is affected, to different degrees, by all environmental changes due to anthropic pressure, as the urbanization of rural areas or the construction of engineering works. All these variations, in addition to the climatic and land use change, can affect the hydraulic response and the sediment transport at different time scale, bringing to:

- changes in the frequency and intensity of floods and debris flow along the hillslopes;
- modification of the ecosystems, modifying habitats suitability for establishment and growth of different species;
- changes in the river morphology.

Under the hypothesis of existence of a long-term steady state condition for sediment transport rates at basin scale, the river basin can be considered as a network made by nodes and edges and the sediment transfer can be analyzed through the concept of connectivity, explained in the next paragraphs, which becomes a property of the system, expressed in terms of transfer and deposition rate.

A robust estimation of the transport process, can be useful for underlying how a change in the conditions of a reach, can impact on the transfer of sediments within the basin.

To date, some simplified models have been proposed. These schematize the river basin as a network, and they aim to define the connectivity expressed in term of sediment fluxes that come from different sources and deposit in each downstream sink. One of the limits of these models is represented by the impossibility to calibrate and validate them because of the lack of measurements of transport connectivity.

This research aims to find a sediment property, measurable in the field or remotely (e.g. size, shape, chemical composition, or other), that allows to infer transport characteristics (e.g. the distance travelled by sediments) which could then be used for the calibration and validation, against the surveyed data from the fieldwork, of the connectivity models.

This, as previously said, can be useful for the management and planning of engineering works, important because of their impact on different aspects, e.g. morphology and environment, and even on the socio-economic asset of a specific region.

In this context, the UN introduced the Sustainable Development Goals, aiming to ensure the socio economic development of nations while taking into account ongoing global changes and climate-related issues.(Nations, 2030) The SDG are 17 and they cover a broad range of problems whose addressing, and solution are crucial in order to gain a more sustainable future. Even this research, can contribute in the achievement of some of the SDG:

- SDG7: Ensure access to affordable, reliable, sustainable and modern energy;
- SDG13: Take urgent action to combat climate change and its impacts;
- SDG15: Protect, restore, promote a sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss.

About the SDG7, it is important to underline that, as reported by the Global Status Report, in 2018 the global renewable power capacity was 2378 GW; out of which, the share of hydropower was 1132 GW (i.e. 48%).

These plants with the climate change and the sea level rise can have a great impact on rivers' morphology(Anderson et al., 2018)(Manh et al., 2015)(Vauclin et al., 2019)(Kummu et al., 2010), because dams, acting like barriers typically induce erosion of the riverbed located downstream of the structure and an aggradation in the upstream side.

Therefore, this research, can potentially help planning infrastructures taking into account, on top of several other fundamental aspects like energy production and safety, but also the impact on the land and water ecosystems at basin scale.

SDG13 and SDG15 stress the necessity of protecting the landscapes and the associated ecosystems and raise awareness about the importance of the climatic issues in the design and management of infrastructures. This research can help in the evaluation of the effects produced on river morphology which constitute the supporting framework for any ecosystem.

Sediments are one of the most common pollutants in water bodies such as streams, rivers and lakes (Davis and Fox, 2009; Guzmán et al., 2013). Human activities that reduce or remove vegetation typically increase soil erosion rates. Sources of sediment include upland areas, as well as the water bodies themselves (Davis and Fox, 2009; Guzmán et al., 2013). Human activities that reduce or remove vegetation typically increase the amount of soil eroded (Davis and Fox, 2009; Guzmán et al., 2013). Possible sediment sources include tilled crop fields, grazed pastures, construction sites, and timber harvesting areas. Along water bodies, the beds and banks erode due to the force of moving water, increasing the amount of solid mass that comes to the outlet of the basin (Davis and Fox, 2009; Guzmán et al., 2013). The sediment transport is an important process that has been investigated for decades since the last century.

2. State of the art

2.1. Sediment transport within the river network

Shield's theory has been one of the most widely used criterion for evaluating the critical condition of incipient motion. Considering a particle characterized by a specific diameter d_{50}, it will remain in a static equilibrium condition only if the drag force exercised by the water flow, does not reach and overcome the resistant one. This state has been evaluated by Shield, defining a non-dimensional threshold coefficient θ_C deriving from the equilibrium of forces acting on the particle and function of the water flow regimes (laminar, transition, or turbulent) (Equation 2-1).

$$\theta_C = \frac{\rho u_c^2}{(\gamma_s - \gamma) d_{50}} \qquad 2\text{-}1$$

The criterion, valid only under restrictive hypotheses, describes the incipient motion condition in a simplified way, since it considers particles as spheres for determining the volume and the area impacted by the flow, neglecting the actual complex geometry of the particles.

Considering a specific reach, the maximum sediment mass that can be transported by the water flow per unit time is a function of its hydraulic and geomorphologic conditions and of the sediment properties (density, volume and shape) and can be quantified through deterministic or stochastic approaches. The main difference regards the way in which these models analyze the process of motion. Indeed, while the first consider the transfer as a continuous process, the latter account for the possibility of the particles to rest for a certain time period.

Among the stochastic sediment transport models, some of the most relevant ones are the ones by Einstein (1942), Kalinske (1947), Ancey et. al. (2008), Furbish et. al. (2012), Armanini et. al. (2018) while, among the empiric ones it is worth mentioning those developed by Meyer-Peter & Müller (1948), Schoklitsch (1962), Engelund & Hansen (1967), Wong & Parker (2006) (Ancey, 2020a, 2020b).

Although the former might be more representative of the actual transport processes taking place within a natural system, they almost inevitably depend on parameters which are difficult to estimate. For this reason, in modelling the sediment transport, the deterministic approaches are used more frequently.

In order to forecast the effects produced by the anthropic pressure on natural processes within a river basin, in recent years the concept of connectivity has been borrowed by other disciplines and applied to geomorphology (Czuba and Foufoula-Georgiou, 2015).

Many have tried to formally define sediment connectivity. According to Bracken "it represents the connected transfer of sediment from a source to a sink in a system via sediment detachment and sediment transport, controlled by how the sediment moves between all geomorphic zones in a landscape" (Czuba and Foufoula-Georgiou, 2014). In general, the connectivity quantifies the degree with which the sediments are able to move within the river basin and can be considered as a property of the system useful for describing the transfer of sediment.

2.2.Sediment connectivity models

To date, several models describing sediment connectivity at river basin scale have been developed, but none of them has been calibrated yet. For this and other reasons their accuracy in representing the actual sediment transport rates within a river basin is still unclear. As explained later, in order to improve these models, it is necessary to find a property which can be easily estimated from the field data, so that the uncertainties regarding the process can be reduced and the connectivity map can be more significant.

The first model that will be discussed is the one introduced by Czuba et. al.in 2014, which has been reviewed and improved in 2015. The model describes the transport of parcels supplied to the river network as instantaneous and uniform inputs.

The output of the initial version of the model, developed in 2014, aims to define the sedimentological response at the outlet, i.e. probability distribution of arriving time of the sediment mass to the outlet (Czuba and Foufoula-Georgiou, 2014).

The main limitation of the model is the impossibility to assess the degree with which the sediments move within the basin, i.e. the connectivity of the system. This model evaluates the connectivity between the reaches only in terms of the parcels' travel time for arriving to the closing section. Although some of the initial issues have been solved, the main limit that still characterizes the model is that it does not give back the explicit map of the connectivity of the network.

Indeed, the output of this model does not explicitly represent the connection between sources and sinks, since the parcels are not subjected to the local transport capacity limitations and depositions (Schmitt, Bizzi and Castelletti, 2016). The model only represents the time needed for travelling from a point to another one, the number of sources connected to each link and the number of links connected to the output at the same travel time distance. Furthermore, the model can evaluate the evolution of transport mechanism with time.

The other model, considered for the research project, is the Cascade model, developed by Schmitt et. al. in 2016. The innovative idea is to consider the river network as a graph composed by nodes and reaches. The model describes the transfer of sediment from each input zone as a separate cascade process. This means that each source point is connected to its possible sinks through a single path, called cascade, so that it is possible to compute a unique connection between a source and a sink and represent a connectivity map in term of sediment deposition rate, and time needed for depositing (Schmitt, Bizzi and Castelletti, 2016). The process of sediment transport along each cascade is a function of the transport capacity of each reach, which is related to the mean diameter of the pebbles inside the reach and its hydraulic and geometrical conditions.

The model wants to reproduce a hypothetic long term steady state of the river network in terms of transport and does not consider morphological changes and their feedback to transport capacity. Moreover, the model considers continuous sediment supply and rectangular cross-section for each reach.

As already mentioned, each cascade is characterized by a single uniform particle size. One sediment source might see the input of more than one size of particles, but these will be treated independent, as separated cascades (Figure 2-1 D). In a confluence of two or more cascades, the interplay of sediments coming from different paths is accounted reducing, through a so-called competition factor, the total available transport capacity.

Since the transport capacity of a reach varies with the grain size, when particles with different dimensions income in the same reach, the available energy is redistributed between all incoming cascades through the competition factor (Schmitt, Bizzi and Castelletti, 2016).

More energy for transferring the sediments downstream will be given to those cascades characterized by pebbles of smaller sizes. The smaller are the particles and the greater is the sediment supply, the more energy will be available to the cascade for transporting sediments downstream. (Figure 2-1F).

Once defined all characteristic of sediment flow in each reach and under the hypothesis of steady flow condition, the model computes the time spent by pebbles for depositing in a certain river reach as the sum it took them to travel through each reach that compose their path. The model produces as a result the connectivity maps expressed in terms of mass deposited within the river network for each cascade (Figure 2-1 G, H). When all or most of the initial sediment supply has all deposited upstream, e.g. a dam force the deposition upstream, or if the transport capacity is not sufficient to transport the sediment, e.g. the sediment particle is too big and heavy, the sink and the source are disconnected, and the cascade is interrupted in the upstream reach.

One of the limitations related to these frameworks concerns the impossibility of calibration with respect to the connectivity property. This is itself due to the lack of a measurable sediment property that could be directly linked to the distance they travelled.

Furthermore, these frameworks do not account for the evolution along the pathways of the particles size and shape characteristics due to collision with other particles, either in motion or standing on the river bed, or due to friction related to sliding motions.

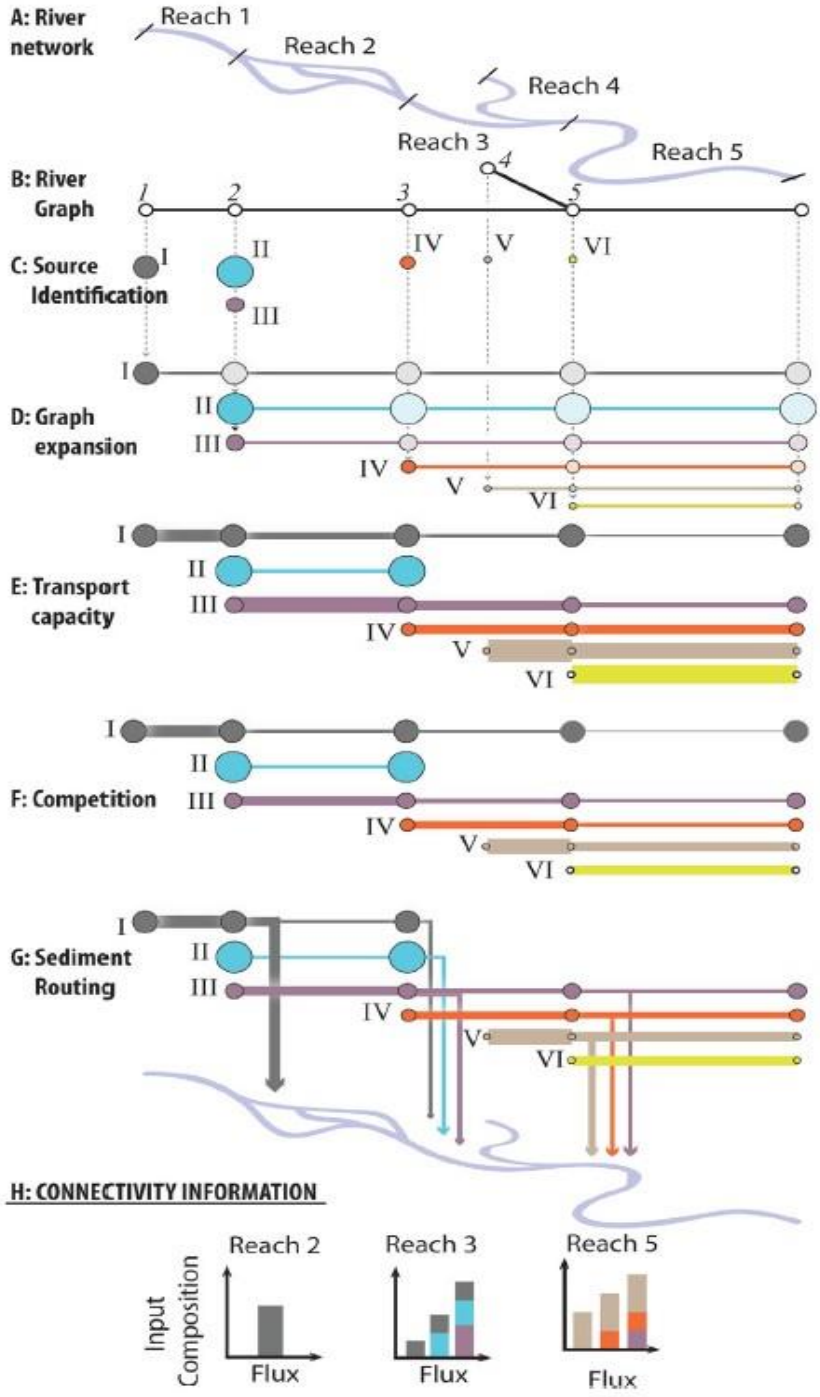

Figure 2-1 Schematic representation of the Cascade framework. Figure a) shows the river network, in b) the network is schematized as a graph, while in c) sources points, characterized by specific diameters of the particles, are shown. In Figure d) individual cascades are tracked for each source. In e), the transport capacity is represented through the line width. In f), the competition in confluence points has been considered reducing the transport capacity. In g), sediment fluxes are represented, so the deposition is accounted for. In h), the connectivity information are summarized in terms of fluxes, provenance and time distance of the reach to the source.

2.3.Sediment fingerprinting

Many sediment fingerprinting techniques have been developed in order to estimate the position of the sources (Davis and Fox, 2009; Guzmán et al., 2013). These techniques use physical, chemical and/or biological tracers to distinguish between the types of sediment sources in a watershed and estimate how much sediment each source provides to the stream (Davis and Fox, 2009; Guzmán et al., 2013). To date, since these methods are deeply affected by the uncertainties linked to the tracers' evolution, the inference of the position of the sources cannot be robust (Davis and Fox, 2009; Guzmán et al., 2013).

2.3.1. Sediment attrition models

Geometrical characteristics of sediments have been studied since Aristotle (Krynine, 1960), since their changes provide indications about their history and help to characterize the depositional environment and to mark the sediment sources at basin scale (Cassel *et al.*, 2018).

Recent studies on abrasion present a model that relates sediment shape properties (e.g. circularity, roundness) and mass loss (Figure 2.3.1-1) (Novák-Szabó et al., 2018). This study assesses the possibility of finding universal characteristics of the shape evolution of the particles with the mass loss when particles collide with similarly-sized particles or with the bed, and collision energy is small enough that chipping is the dominant mechanism (Novák-Szabó et al., 2018).

In this sense, Bloore proposed a theory for describing the shape evolution of the particles when they are transported by bed-load saltation and when the conditions just mentioned are valid. According to his theory, abrasion rate (ν [/]) is a function of the particles' local curvature (κ [1/m]) and the perimeter of the impactor particle c [m] (Novák-Szabó et al., 2018).

$$\nu = 1 + \kappa c \qquad 2\text{-}2$$

This geometric relation is a partial differential equation and it determines the rate at which the particle erodes. This model describes abrasion rate as dependent on two different components. When impacting particles are small, the dominant component is $\nu = 1$, while if particles are of similar size the relation is said curvature-driven abrasion, since the main influent component is kc (Miller et al., 2014). In Figure 2-2, it is possible to see how the shape of a particle evolves according to this theory. Indeed, in those points characterized by low curvature, the particle is abraded slowly. Furthermore, the evolution of the particle is divided in two phases: first, sharp edges with high curvature rapidly round off without major changes in the global axis dimensions, then, the axis dimensions start to decrease slowly and the pebble becomes more spherical (Miller et al., 2014) (Figure 2-5).

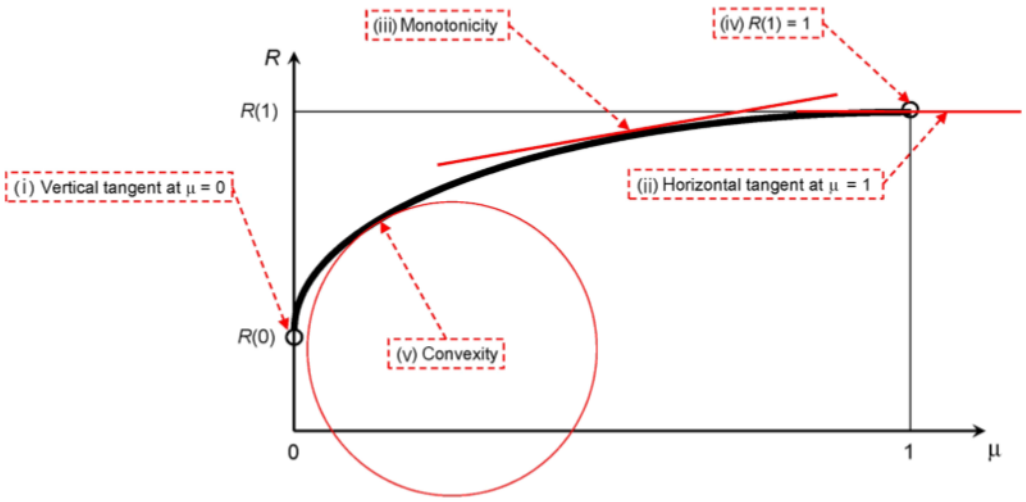

Figure 2-2 Expected evolution of circularity (R) as a function of the mass loss μ.

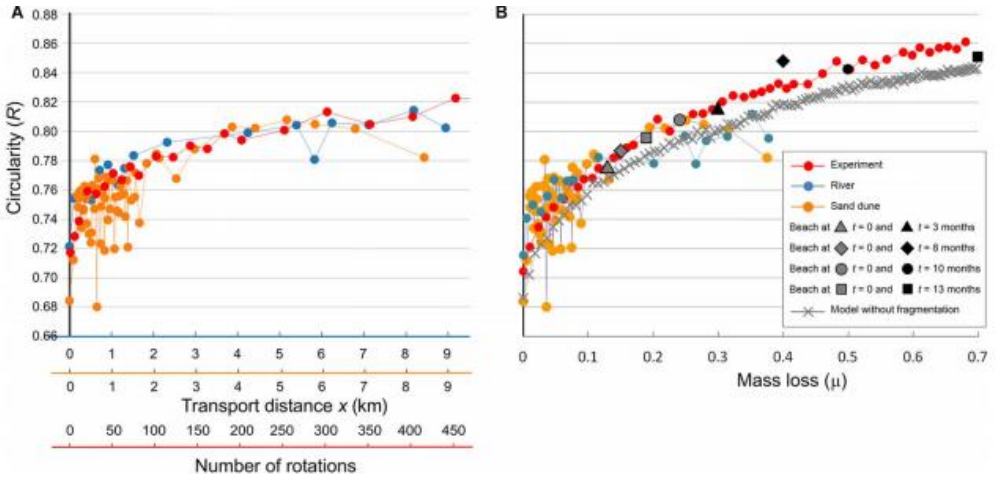

Figure 2-3 Field, laboratory experiences and model data showing the evolution of circularity with relative mass-loss (B) and distance travelled (A).

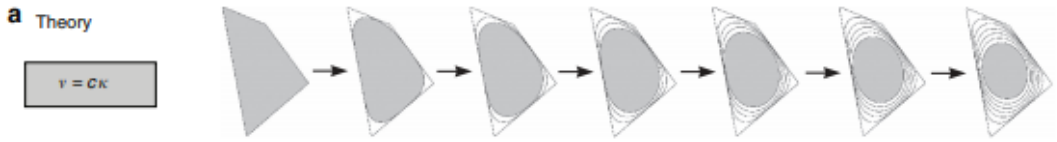

Figure 2-4 Szabo et. al., 2015. Theoretical evolution of the particle shape according to the Bloore's model restricted to the case of big particle.

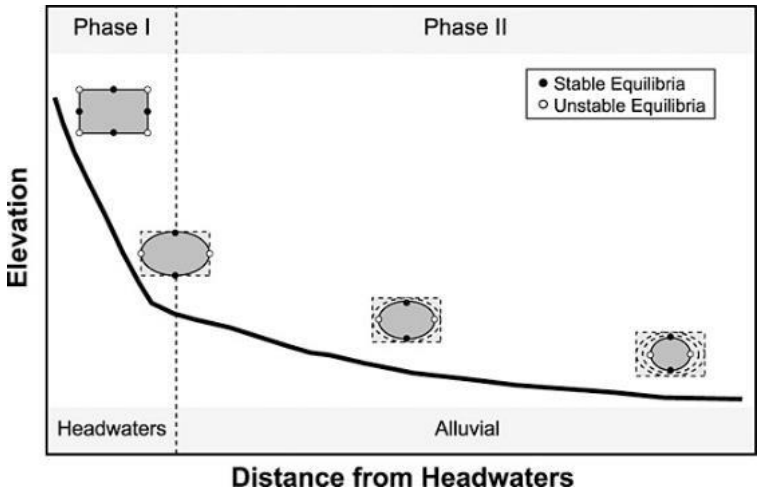

Figure 2-5 2-Phase evolution of the particles shape

From this theory, Novák-Szabó et al. have tried to predict the qualitative trend of circularity evolution with mass loss. Considering an initial polyhedral particle, characterized by a low circularity, and the restrictive hypothesis of same-sized particles collisions, the initial rate of rounding is infinite because the particle is sharp, and the curvature is infinite. Then, travelling downstream, it becomes more and more round, decreasing the local curvature and therefore, the rounding rate. Therefore, the expected curve is a continuous monotonic increasing function, defined in the range [0, 1] and varying between R_0 and 1, while its derivative is a continuous decreasing function. The theoretical trend of the curve $R(\mu)$ is shown in Figure 2-2 (Novák-Szabó et al., 2018).

It is necessary to underline that since the evolution of circularity with the relative mass-loss is strictly dependent on the boundary and initial conditions, only the geometrical characteristics of the evolution have been claimed universal. To support these assumptions, Figure 2-3 B shows the evolution of particles' with relative mass-loss analysed for three different systems: cobble-bedded river, sand dune, laboratory experiences with volcanic clasts, and pebble beach. All the curves R versus relative mass-loss present all the qualitative expected features, but their trends are numerically different.

2.3.2. Sternberg Law

Since measuring mass loss of individual particles in the field is not possible with available technologies, several studies have been conducted aiming at the definition of an empirical relationship to describe its evolution as a function of the distance travelled by particles (Kuenen, 1956; Rice, 1994; Pizzuto, 1995; Novák-Szabó *et al.*, 2018). In the study of Novak-Szabo (Novák-Szabó et al., 2018), this relation has been supposed being linear although the rate of mass loss is expected to be affected by a series of interconnected variables including: transport conditions, which depend on discharge, morphology and particles' sizes and on particles' properties (e.g. shape, lithology).

Nevertheless, past studies on abrasion found a weak influence of the flow velocity, considered for sake of simplicity as a surrogate of the grain velocity, on the abrasion rate (Lewin and Brewer, 2002).

Sternberg, in 1875, suggested that particles' mass decreases exponentially with their distance travelled (Krumbein, 1941; Kuenen, 1956; Brewer and Lewin, 1993; Lewin and Brewer, 2002; Frings, 2008). This relation depends on the abrasion coefficient k [km^{-1}] which represents the rate with which the mass is abraded (Equation 2-3).

$$\frac{M}{M_0} = \exp(-kx) \qquad 2\text{-}3$$

This relation was determined empirically considering only the effect of collisions among particles. Nevertheless, in real systems the mass decrease is caused by both size-selective transport, for which lighter particles will travel more frequently, and abrasion, due to collisions and dependent on several characteristics, e.g. lithology, grain size, grain shape *et cetera* (P., 1879; Krumbein, 1941; Szabó *et al.*, 2015) (Equation 2-4).

Both these effects contribute to the value of the fining coefficient k_W that can be estimated from field measurements (Equation 2-4). It remains nonetheless difficult to estimate the relative contribution of abrasion and selective transport to downstream fining, which is the evidence of decrease of average particle size as one moves from the mountain part of a basin to the lowland (Szabó et al., 2015). To date, several authors have tried to estimate the speed at which the size of sediments varies going downstream accounting only the abrasion process $k_{A,D}$, but different results have been obtained. The values reported generally vary between 10^{-5} and 10^{-1} [km^{-1}] (Morris and Williams, 1999). Only Kodama reported a variability between 10^{-2} and 10^{-1} [km^{-1}] (Kodama, 1994). Size reduction coefficient can then be used to calculate the mass reduction coefficient. Since the weight of the particle is considered proportional to D^3 (Mikos, 1995), the mass abrasion coefficient is, generally, expressed by as $k_A = 3k_{A,D}$.

However, since the k_A values were estimated during laboratory experiments using either an abrasion tank or a tumbling barrel, they underestimate the effective coefficient k_W with which mass decreases since they do not account for other processes, e.g. the selective transport, that influence the abrasion process (Morris and Williams, 1999).

$$k = k_A + k_S \qquad 2\text{-}4$$

Moreover, in a study of Sklar et al., 2016 they suggested a decreasing evolution of the abrasion rate with the distance from the source.

2.4. Image processing applied to geosciences

Roundness and shape properties of pebbles can be estimated using one of the many available tools that have been developed through the years for this purpose. Two relatively recent and well known in the geomorphological/sedimentological settings are the ones developed by Roussillon *et al.*, 2009 and Zheng and Hryciw, 2015. These models work on images representing samples of particles collected and positioned on a board (Roussillon et al., 2009; Zheng and Hryciw, 2015). Only the former model includes both a tool for pebbles' segmentation and one for the computation of shape parameters (e.g. Circularity, Elongation, Roundness). The second model requires digital images to be in gray-scale with already segmented particles.

A digital image is a matrix made of pixels (*picture element*), each of which contains a value called digital level, a whole number that represents the energy captured by the camera sensor, i.e. the light intensity, and that varies between 0 and 255 (in the case of a 8-bit per pixel representation).

Image visualization can be performed using:

- Monochromatic system, where each pixel contains a single digital level and the image is said to be a gray-scale representation;
- Colour system, where image is a 3-dimensional matrix and each pixel contains three light intensities, each one relative to each of the three components or channels. In the case of a true-colour image, the colour correspondent to each pixel, is the result of a combination of the digital levels relative to the red, green and blue channels. For this reason, the image is said represented in RGB space. If, for instance, in a pixel the light intensity relative to the second channel (Green) is zero, the pixel colour is purple.

A linear combination of the three components converts the true-colour image into a monochromatic one. The function used in the Matlab™ tool *rgb2gray* is showed in Equation 2-5.

$$0.2989 * R + 0.5870 * G + 0.1140 * B \qquad 2\text{-}5$$

In order to have a robust estimation of the parameters characterizing each particle, before implementing these algorithms, it is necessary to improve the quality of the pictures and then, through the segmentation process, identify the outline of each particle within the sample.

In general, the edge of an object represents a visible and sharply discontinuities in the color or in the light intensity. In the field of digital image processing, this is translated in terms of the light intensity gradient. Considering a monochromatic image, the gradient represents the variation of the light intensity values along a specific direction. Considering two adjacent pixels, if the discontinuity is sharp, the difference in the light intensity is relevant as the image gradient.

Several methods for edge detection, based on image gradient, have been implemented (e.g. Canny algorithm, Sobel, Robert and Prewitt operators, Laplacian of Gaussian, Watershed algorithm (Salman, 2006; Maini and Aggarwal, 2009; Gupta and Ghosh Mazumdar, 2013) .

The segmentation method used in this research, is the one implemented by John F. Canny in 1986 (Canny, 1986).

This method aims to enhance the already existent approaches (e.g. Sobel and Laplace). For this reason, Canny followed three criteria (Canny, 1986):

- All the edges have to be detected;
- All the edges have to be well located, which means that the distance between the pixel belonging to the edge detected and the one located along the actual edge, has to be the minimum;
- The edge of an object has to be single.

This approach can be schematized as a six-step chain process that starts transforming the true-color (RGB) image into a monochromatic one (gray-scale), obtained as a linear combination of the three plans of the RGB space. Considering a colored im carries in each pixel only the information about light intensity, i.e. the energy . The process continues with the application of a Gaussian filter that smoothens the image. This is achieved convoluting the greyscale image with a square matrix and is therefore a linear filter. Its objective is to blur the image and to reduce the noise caused by random fluctuations of light intensity. The convolution process works as follows: for each pixel of the grayscale image the code computes the weighted average of the values of the pixels surrounding it using as weights the values of a square matrix that is centred on the pixel of interest. In case of a Gaussian filter, the matrix of weight is made of valued distributed according to a Gaussian distribution, as a function of their distance from the centre of the square. The grater the standard deviation of the Gaussian distribution, the larger the size of the square matrix and the smoother the resulting image.

Nevertheless, a too large Gaussian filter may produce a significant smoothing of the image, even reducing the contrast along the particles' outline that instead, for our analysis, must be as enhanced as possible.

On the contrary, a too small filter my not detect the Gaussian noise that can disturb the search of the particles' outline.

The Canny algorithm continues with the application of the 2-D first derivative Sobel operator to the smoothed image to highlight regions of the image with high first spatial derivatives.

This operator performs the spatial gradient evaluation in both x and y directions using two 3x3 convolution Equation 2-6. Since the edges give rise to high local image contrast, peaks in the image gradient along the edges are expected.

$$G_y = \begin{matrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{matrix} \qquad G_x = \begin{matrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{matrix} \qquad \text{2-6}$$

The following step is the evaluation of the edge direction and then its adjustment in order to relate it to a direction that can be traced in the image. This means that 4 principal directions are defined, i.e. horizontal and vertical directions (0 and 90 degrees), and along the positive and negative diagonals (45 and 135 degrees), and the angles between 0 and 180 degrees are grouped around these principal directions as showed in Figure 2-6.

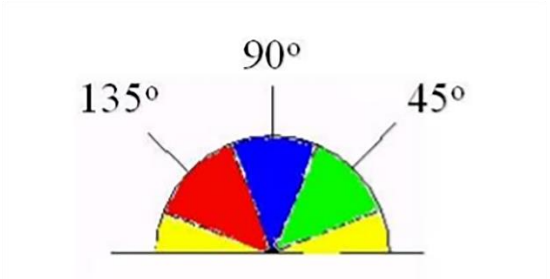

Figure 2-6 Angles grouped around the principal edge direction

Then, in order to give a thin line in the output, all pixels that are not actually on the ridge top, are set to zero value through a process known as non-maximal suppression. Finally, in order to avoid the breaking up of the detected contour, the tracking process makes use of two thresholds: T1 and T2, with T1 > T2.

If a single threshold, e.g. T1, would be defined, all the pixels with an edge strength lower than T1 would be set to zero, although they actually could belong to the edge.

This would result in a discontinuous particle's outline, similar to a dashed line. On the contrary, the definition of two thresholds (T1< T2) allows all the pixel with edge strength greater than T2 to be detected as belonging to an edge, while weaker pixels characterized by edge strength lower than T1 would be disregard.

Pixels having an edge strength comprised between T1 and T2 that are connected with the stronger ones are would be classified as edges while the ones that are not connected to pixels having edge strength above T2 would be discarded.

This selection process improves the edge detection accuracy and, in many cases, allows the definition of a continuous outline.

2.4.1. Review of two computational geometry models

The shape of the particles can be a very important property for the inference of the distance of travel and the river conditions. The shape characterization can be conducted using parameters that can detect the shape variations at different scales: Form (e.g. Elongation and Circularity), Roundness (related to sharpness of the particle) and Texture (e.g. Roughness). The shape metrics here analyzed are: Roundness, evaluated through the Wadell's formula, Circularity and Elongation (Equation 2-7).

The first one defines how sharp the corners of the pebble are, while the last two parameters describe the shape of the pebble and give a measure of how similar to a circle the particle is.

The Zheng's model is implemented in Matlab and it evaluates the roundness through Wadell's formula (Wadell, 1932).

This roundness measure is defined as the ratio between the averages of the radii of curvature of the corners to the radius of the maximum inscribed circumference in the particle. The radii accounted for evaluating the mean value must be smaller than the radius of the largest inscribed circumference.

$$r_W = \frac{1}{N}\frac{\sum_{i=1}^{N} r_i}{R}$$

$$C = \frac{4\pi A_s}{P_s^2} \qquad 2\text{-}7$$

$$e = \frac{b}{a}$$

Where N and r_i [m] are respectively the number of corners composing the particle's outline and their radius of curvature, while R [m] is the radius of the maximum inscribed circumferences, A_s and P_s are respectively the area and the perimeter of the particle.

For defining corners, the algorithm developed by Zheng et. al. (2015) first removes the roughness associated to the particle's texture defining a smoothed mean surface, then it approximates the silhouette with a set of connected segments, defining a distant threshold value δ_0. The corners are then defined are those laying on the convex stretches of the silhouette.

Finally, the algorithm tries to fit iteratively the circle to each set of closest corner points Figure 2-7.

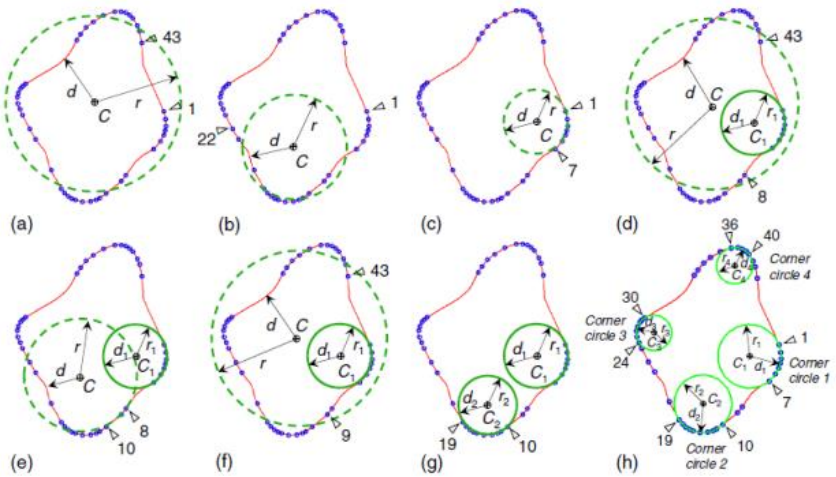

Figure 2-7 Roundness algorithm proposed by Zheng

It is important to underline that the output of the model is based on a threshold value on which depend the definition of the corners δ_o, and consequently the evaluation of the roundness. It has been observed that the threshold value, and therefore the roundness and the sphericity inferences, are a function of the PCD, i.e. the resolution of the image expressed in terms of pixel per circumscribed circle diameter.

The results of this analysis assess that, in order to accurately capturing the roundness value, at least 200 pixels for PCD are necessary and the δ_{MAX}, i.e, the maximum value of the parameter δ for the determination of the corners, has to be 0.03% of PCD (Zheng and Hryciw, 2015).

The other algorithm proposed by Roussillon et al. (2009), aims to the evaluation of the shape parameters and roundness as well, approaching in a different way.

The framework first applies the k-means segmentation algorithm and then it applies on the binarized image a second algorithm through which it evaluates shape parameters. Because the parameters depend on the area and perimeter, the algorithm first tracks the contour as a polygonal individuated by a sequence of 8-connected pixels, and then it calculates the geometrical characteristics of the pebbles (Figure 2-8 left).

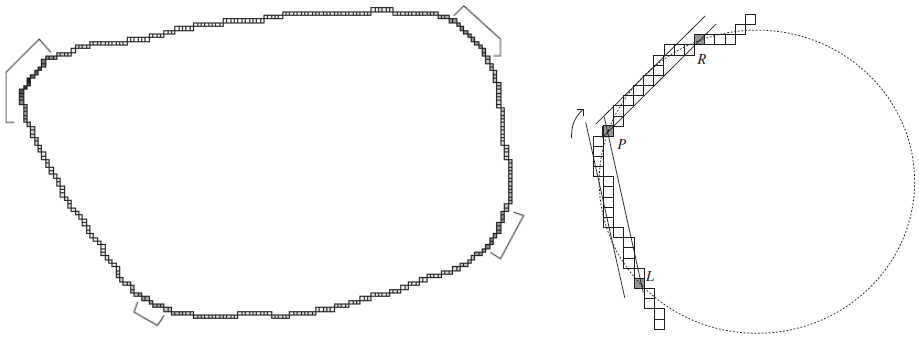

Figure 2-8 Computation of the edge of the particle (left) Computation of the radius for each pixel belonging to the silhouette of the particle (right)

Through this framework, r_w has inferred without defining the corners but drawing for each pixel, belonging to the silhouette, the circle obtained joining the pixel with two others located on the right and on the left of it through the longest sequence of 8-connected pixels lying between two straight parallel lines (Figure 2-8 right).

The analysis performed by Roussillon et. al. (2009), showed that the resolution of the image influences the result: in order to avoid its effect, the perimeter of the pebble should be greater than 150 pixels. In this way, the error on the inference of the shape parameters should reduce to less than 10% (Roussillon et al., 2009).

In order to have a robust estimation of the parameters characterizing each particle, before implementing the algorithms, it is necessary to improve the quality of the pictures and then, through the segmentation process, individuate the edges of each particle of the sample.

Once the algorithms are applied, information about axes length, roundness and circularity are gained.

2.5. Gaps of knowledge

The sediment transport at basin scale is a complex process which is difficult to be analyzed because of the uncertainties in the hypothesis and the parameters on which each model is based. Their accuracy in representing the actual response of a system is therefore not clear. The objective of the research is to reach a robust estimation of the sediment transport connectivity, e.g. the distance travelled by sediment. To do this, the shape characteristics are suggested as possible key property for the sediment fingerprinting. Nevertheless, to date, a formal function that relates the shape metrics, in particular the Circularity, to the distance has not been given yet.

Past studies suggest two relations that may be useful for the purpose of the research:

1. The universal relation between Circularity and Relative mass loss;
2. The exponential increasing in mass loss with distance.

Nevertheless, both the relations present some limits. The universal abrasion curve has not been defined analytically, while the empirically observed exponential degradation function is dependent on a coefficient, the abrasion rate which is not *a priori* known and which is expected to depend on the lithology, particles' size and size-selective transport.

Regarding the image processing for the acquisition of necessary data for improvement of the model, the main defect is related to the problem of pre-processing of digital images. The disturbances on the image do not allow to reach an accurate segmentation of the pebbles.

Indeed, the presence of shadows and of complicate texture of the pebbles, make impossible to apply a mask to the image and separate the pebbles from the background.

3. Research Questions

To date, a series of gaps of knowledge are still present with respect to the possibility of finding a relationship between the shape morphometric characteristics and the distance travelled by the particles. The research aims to overcome some of the existent limits and, in particular to answer the following research questions:

- How can the segmentation of images of pebbles collected in the field can be improved with respect to the effects produced by the shadows?
- How can accurate can be the distance travelled by particles within the river basin, when inferred from the morphometric characteristics of the pebbles?

4. Methodology

In this section the different phases of the research work will be explained:

The methodology used in this research to answer the previously stated questions is shown in the following scheme and divided in two main phases (Figure 4-1).

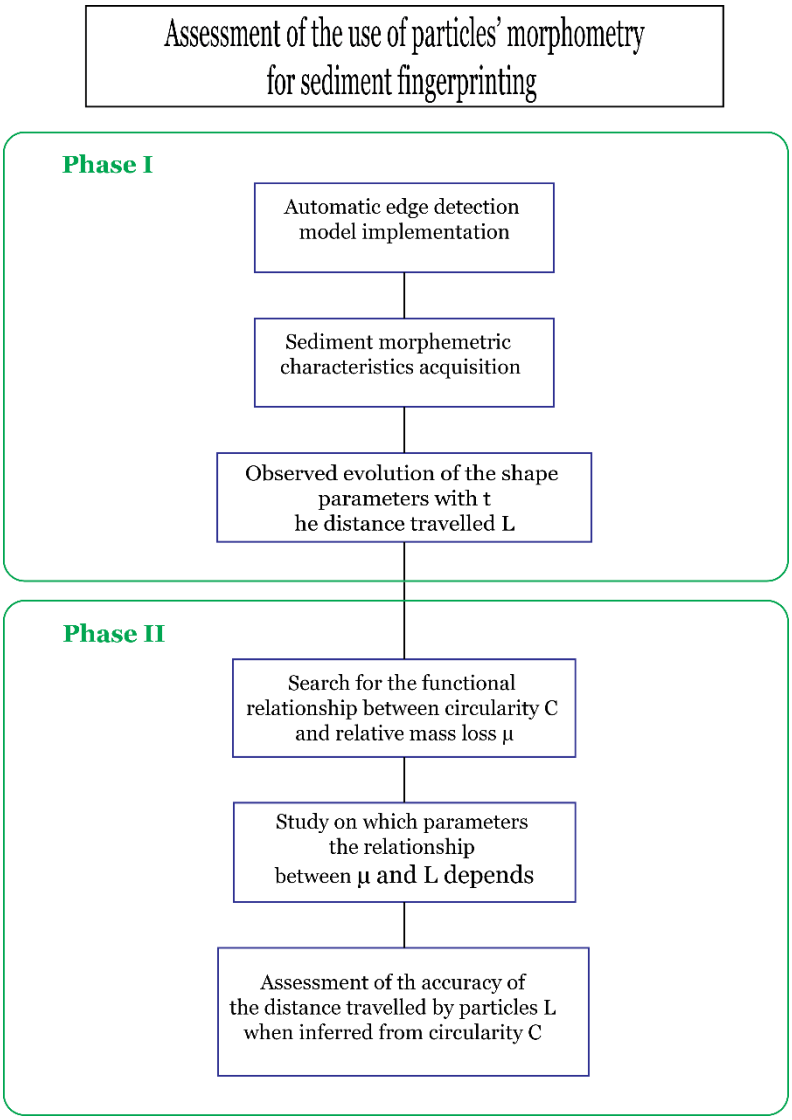

Figure 4-1 Work phases

4.1. Image Processing

As mentioned in the previous sections, changes in size and shape can be useful for the description of the sediment transport which, to date, is still constrained. The shape characterization of the particles based on image is the cheaper solution. During the fieldworks, images of the sampled particles are taken positioning the pebbles on a flat surface to give a uniform colored background.

Sediment images collected in the field are generally characterized by irregular shadowing and by intra-granular variations in texture and colour Figure 4-2. Since in many cases shadows reduce the local contrast near the edge of a particle, common automated methods for edge detection do not provide accurate contours segmentation (e.g. Sobel operator(Gupta and Ghosh Mazumdar, 2013), Canny algorithm (Canny, 1986) and Watershed algorithm (Zheng and Hryciw, 2015)). The model here proposed, comprises two phases. It first identifies the area of interest, meaning the board on which particles are lied during the field survey. The next phase involves the identification of each individual particle and its analysis using the probabilistic Canny method (Canny, 1986). This iterative approach allows the operator to visually control the quality of the segmentation obtained for each pebble and to decide whether to continue with shape parameters extraction or to erase the pebble from the sample.

Figure 4-2 Example of an image of pebbles sample collected during a field survey.

4.1.1. Phase I: background board identification

The first phase of the model aims at removing the outer material from the initial image Figure 4-2. For this purpose, the true-color image is first transformed into gray-scale. Given the sharp change in color and texture between the outer material and the background board, its edge represents a discontinuity in the light intensity, i.e. these points are characterized by a significant gradient in gray-scale. The image contrast is therefore initially enhanced, increasing the highest light intensity (white board) and reducing the lowest ones (background and pebbles of the sample). The contrast of the image represents the ratio between the maximum value and the minimum one. Enhancing the contrast means, enhancing the ratio, reducing the window of light intensities between the maximum and the minimum value, and removing the intermediate intensity Figure 4-3 A. The light intensity adjustment is therefore used for scaling the image intensity values from the window [0 255] to a new range. The *imcontrast* Matlab™ tool works on the histogram of pixel values and allow to enhance the contrast and the brightness of an image by modifying manually the interval of the light intensity values of the image. Stretching the interval reduces the contrast. Shrinking the range enhances the contrast. The extremes selected as those that sufficiently enhance the contrast are [100, 120]. This means that in the output image all the pixel, with an intensity values below 100, are black, while the pixels, with values above 120, are white. The light intensities between 100 and 120 are transformed linearly to the values between 0 and 255 Figure 4-4.

The contrast enhancement increases the probability of the localization of the maximum absolute value of the intensity gradient on the edges of the white board. This gradient is computed both horizontally and vertically by the convolution of the Sobel operator(Engel, 2006; Gupta and Ghosh Mazumdar, 2013) and the adjusted image.

The intensity gradient can be positive or negative, depending on whether there is an increase, a decrease in the light intensity. The material outside the background board can be characterized by color and texture variations that might produce relatively high intensity gradients.

Moreover, the pebbles on the board also create gradient changes, especially located along the pebbles' contour. The model partially reduces these effects by removing the edges, defined as connected components using the 8 direction connectivity, with fewer than 500 pixels. This is done under the hypothesis that the outline of the table is strong enough to be characterized by a number of connected pixels greater than 500. The binary filtered image is multiplied by the original representation of the intensity gradient so that each pixel, belonging to the filtered connected component, contains the positive values of intensity gradient Figure 4-3 B.

This method for segmenting the background board searches for its edges in the vertical and horizontal directions separately and proceeds sequentially respectively by column and row. Here is reported the steps of the process adopted for the identification of the top horizontal edge, while the identification of the other edges will follow an analogue process.

The search for the pixels composing the top horizontal edge starts from the first column. The code searches for the pixel along the first column having the highest positive gradient and then evaluates the following conditions:

(1) If none of the pixel in the i^{th} column has a positive gradient, then the code moves to the column i+1.

(2) If the pixel with the maximum positive gradient in the vertical direction is located in the upper half of the image and if its distance from the edge identified in column i-1 is less than 10 pixels, then this will be the next edge pixel and the code will move to column 1+1.

(3) If the pixel with the maximum positive gradient in the vertical direction is located either at a distance of more than 10 pixels from the edge on column i-1 or below the middle of the vertical size of the image, this means that the top edge along that column is not characterized by the maximum gradient due to the presence of inside pebbles. The upper edge along that column is therefore approximated by the same row value as in column i-1 and the analysis will move to column 1+1.

The result is a binary image where, to each column, is assigned a value of 1 to all pixels located below the detected edge and a value of 0 elsewhere. A similar process is repeated to identify the three remaining edges.

The final four binary images are then overlapped resulting in a binary mask with a value of 1 representing the pixels corresponding to the whiteboard background.

If two or more regions of pixels with value 1 are identified, they are ordered according to the number of pixels that belong to them. Under the hypothesis that the board is the biggest region identified, all the other regions are removed. Moreover, if part of the white board is not detected as such, the model changes the values of the pixels, marked as external to the white board, from 0 to 1.

The resulting white board is showed in Figure 4-3 C.

At the end of this process, the output image is a true-color representation of the white board and of the pebbles positioned on it. The outer material is replaced by a uniformly colored region. In order to reduce the contrast between the panel and the outer region, the color of the latter was selected to be as similar as possible to the color of the board Figure 4-3 D.

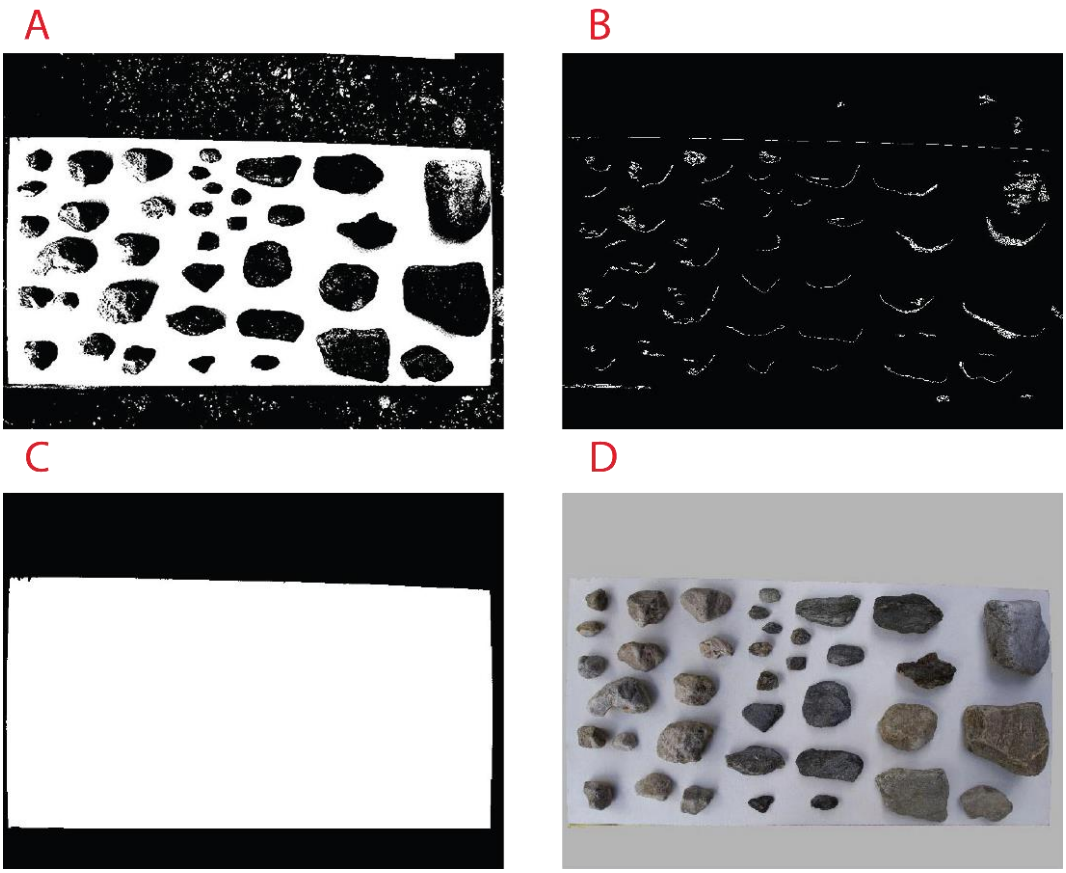

Figure 4-3 A) Result of image enhancement. This process increases the difference between height and low intensity pixels. B) Image of the filtered positive light intensity gradient representation. C) The binary mask representing the white board. D) Result after elimination of the outer material: a RGB image where outer material is replaced by a uniform region.

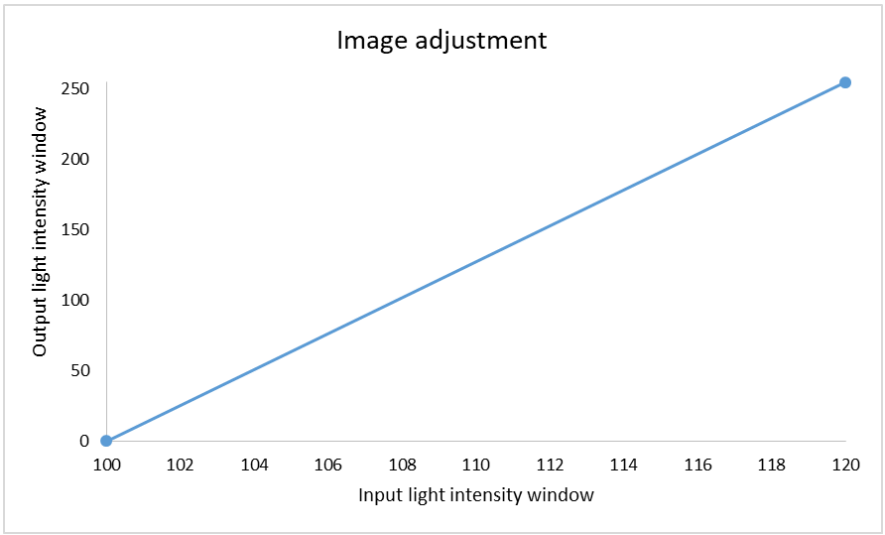

Figure 4-4 Enhancement of the image contrast. Linear transformation of the input light intensity into the output light intensity window.

4.1.2. Phase II: Removal of the shadowing effect on particles' segmentation

Although through Phase 1 the outer material is removed, the shadowing around the pebbles placed on the white board is still present and might bring to inaccurate final segmentation.

The second phase of the model aims to track the contour of each pebble that is cropped through a window built around it. The window is built around the centroid of the analysed particle. The position of the centroid of a body is a function of its volume and how it is distributed. Since the analysis of the image is bi-dimensional, the position of the centroid is dependent on the area of the particle, which is represented by the number of pixels that belong to it. Moreover the size of the cropping-window is related to the size of the bounding box, which represents the smallest rectangle that contains the particle. Nevertheless, the great variability of the light intensity values of the grey-scale image between 0 and 255 makes challenging the possibility of identifying the correct area of the particles when the image is binarized. As showed in Figure 4-5 A, the binarization of the image in grey-scale, leads to identify the shadow like part of the section of the particle, carrying in some cases to the union of two adjacent pebbles and consequently to a wrong positioning of the centroid. For this reason, the image showed in Figure 4-3 D, is enhanced by changing the minimum and maximum values between which the light intensity of the image varies, as done in Section 4.1.1. The extreme values selected as representative of the new range of values are 40 and 200. This means that the light intensities greater than 200 correspond in the new image to 255, while the light intensities lower than 40 are transformed into 0 in the new image. The enhanced image is binarized and the identified black regions represent the part of the particle that have been detected after the enhancement of the contrast Figure 4-5 B. The centroid and the bounding box are then evaluated for each white region representing the particle.

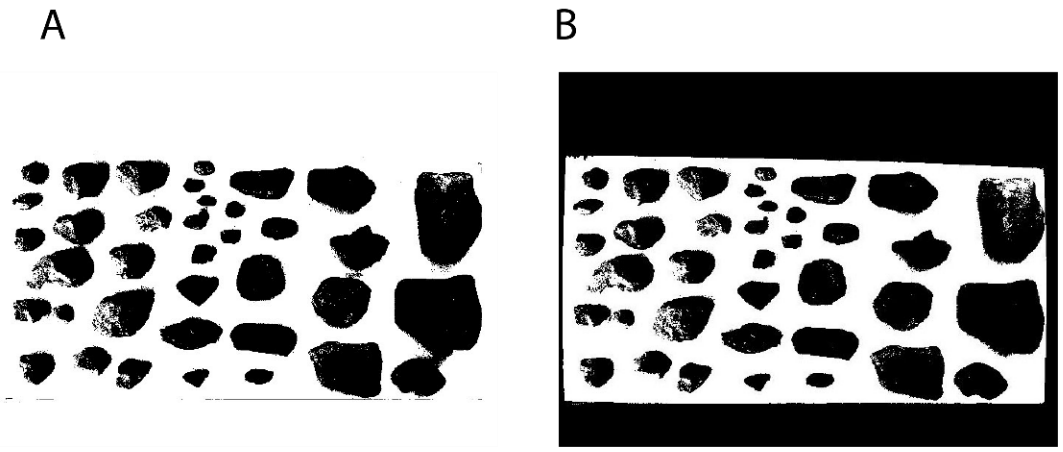

Figure 4-5 A) Image binarized without enhancing the contrast. Some of the pebbles are merged due to the presence of the shadow. B) Image binarized after the enhancement of the contrast

As previously said, the vertical and horizontal sizes of the window are dependent on the size of the bounding box. The window sizes' are enlarged by a factor of 1.6 to the sizes of the bounding box, to consider the possibility for which not all the pixels belonging to the particle are detected and to account for the shadow that in the adjustment process becomes part of the particle area moving its centroid. Moreover, the coordinates of the top-left corner of the bounding box have been shifted horizontally and vertically for a quantity equal to 30% of the horizontal and vertical sizes of the bounding box. If the particle is too close to one of the boundaries of the image, the window is enlarged only on the opposite side of the closest image boundary by a factor of 1.3. Moreover, if the top-left pixel of the bounding box is close to the left boundary of the image, the pixel is simply shifted until reaching the column 1 of the image.

The first step for particle contour tracking uses the Canny algorithm (Canny, 1986), which is already implemented in Matlab™, considering the default values for its controlling parameters. These are the standard deviation of the Gaussian filter applied (σ_C) (the default value is equal to the$\sqrt{2}$) and two threshold for the strength of the edges identified: all edges with strength below the lower threshold (C_l) are neglected, while all edges with strength above the upper threshold (C_u) are preserved. The edges with strength between C_l and C_u are identified only if connected with the strongest edges.

The resulting image is a binary one where the edges are in general not limited to the particle's contour but might comprise internal discontinuities due to variations in color or texture. Moreover, the presence of significant shadows might worsen the quality of the segmentation, producing artificial edges not associated with the shape of the physical object. In order to assess the presence of these effects, the model analyzes the resulting edges as connected components using an 8 direction connectivity and identifies the longest connected component. If the shadow and the intra-granular variations in color and texture are negligible, the outer border is characterized by a strong contrast with the background and therefore the longest connected component properly approximates the pebble's outline. On the contrary, when the shadowing is relevant, the longest component properly identifies only the part of the pebble's contour exposed to the light where the image contrast is significant and where the shadow is negligible. The two different cases can be discriminated in terms of endpoints, i.e. the two extremes of the longest detected edge. If the longest edge tracks the exact pebble's contour, it is closed and therefore the number of endpoints equals zero and the process ends. If, instead, the vector containing the endpoints' coordinates is not empty, the process continues with the second step of the model, trying to identify the closed shape of the particle Figure 4-6 A.

The second step of Phase 2 is based on the hypothesis that, applying the Canny algorithm using different values of the controlling parameters, more significant edges (therefore having higher strength) will be identified more frequently. The parameters that the model varies to assess the relative frequency with which edges are detected are: the standard deviation of the Gaussian filter and the lower and upper thresholds on edges strength. Each combination of these three parameters produces a binary image of edges. The relative frequency, defined as number of times with which each pixel is detected an edge, can therefore be computed. The output representation is a gray-scale image, where pixels detected more frequently as an edge appear of lighter color Figure 4-6 B.

This edge frequency map is used to iteratively attach new edges to the longest connected component previously discussed to close the pebble's outline.

The process is implemented as follows: the model starts from one endpoint and searches for the new endpoint in a square window of size l pixels around this endpoint. The selection criteria for the new endpoint is based on the values of the edge frequency map and on the relative position of the pixels within the search window with respect to the current endpoints.

A series of variables that represent the pixels' relative position, expressed in polar coordinates evaluated with respect to the centroid of the longest connected component, have been introduced. $d_{C\text{-}P}$ and α are the polar coordinates (distance and angle respectively) of a pixel within the search window with respect to the centroid of the actual longest connected component Figure 4-6 C. δ is the angle formed between a pixel in the search window and the opposite endpoint. $d_{E\text{-}P}$ is the distance between a pixel in the search window and the endpoint at the center of the window. ∇f is the gradient of the frequency map computed between each pixel in the window and the current endpoint around which the search window has been built. Δ is the angular distance between the current endpoints. Pixels within the search window are initially filtered according to three intuitive conditions:

(1) $f > 0$

(2) $\delta < \Delta$

(3) $d_{E-P} \neq 0$

The first condition states that the new pixel must be searched among those that have the possibility to be classified as edge ($f>0$); the second condition states that it must be searched among the pixels that contribute to decrease the angular distance between the endpoints of the longest connected component ($\delta>\Delta$), while the third condition states that the selected pixel cannot be the current endpoint, i.e. its distance from the center of the window must be positive ($d_{E\text{-}P} > 0$). The first and third conditions must always be met while, if the second one is not respected, the model selects the pixel with the minimum value of δ (even if $\delta \geq \Delta$).

If multiple pixels respect these conditions, a set of criteria are added to filter the pixels and choose the new endpoint of the particle's outline.

Each additional condition is set only if, within the search window, there are more pixels that meet the conditions just imposed.

The list of additional criteria is:

(4) $\max(\nabla f)$;

(5) $\min(d_{E-P})$;

(6) $\min(\delta)$;

The intuitive meaning of these criteria is: the new edge is selected based on the maximum gradient of the probability for it of being an edge, its distance from the current endpoint should be the minimum in order to guarantee the continuity of the object's outline, and its position should be such that the angle between the current endpoints will decrease. The pixel resulting from such selection criteria become the new endpoint and it will be connected to the current one with a linear interpolation.

At this point the model moves to the other endpoint and the process is repeated. The angular distance between endpoints iteratively decrease bringing them closer together alternatively from both sides. The process stops when either it is not possible to identify more edges to connect to the two endpoints or these are close enough to fulfill one of the two following conditions:

$$\Delta < 6°;$$

$$d_{E-E} < \frac{(l-1)}{2}\sqrt{2}$$

The first conditions states that if the angle between the two endpoint is less than 6°, and the one of the endpoint is located within the search window of the opposite one, they are close enough. At this moment the model stops and the two endpoints are linearly interpolated and the particle's outline is closed.

At the end of the segmentation process, the identified contour is plotted on top of the true color image Figure 4-6 D.

The operator can visually control the quality of the result and decide whether to keep the resulting 2D binary outline or to discard it.

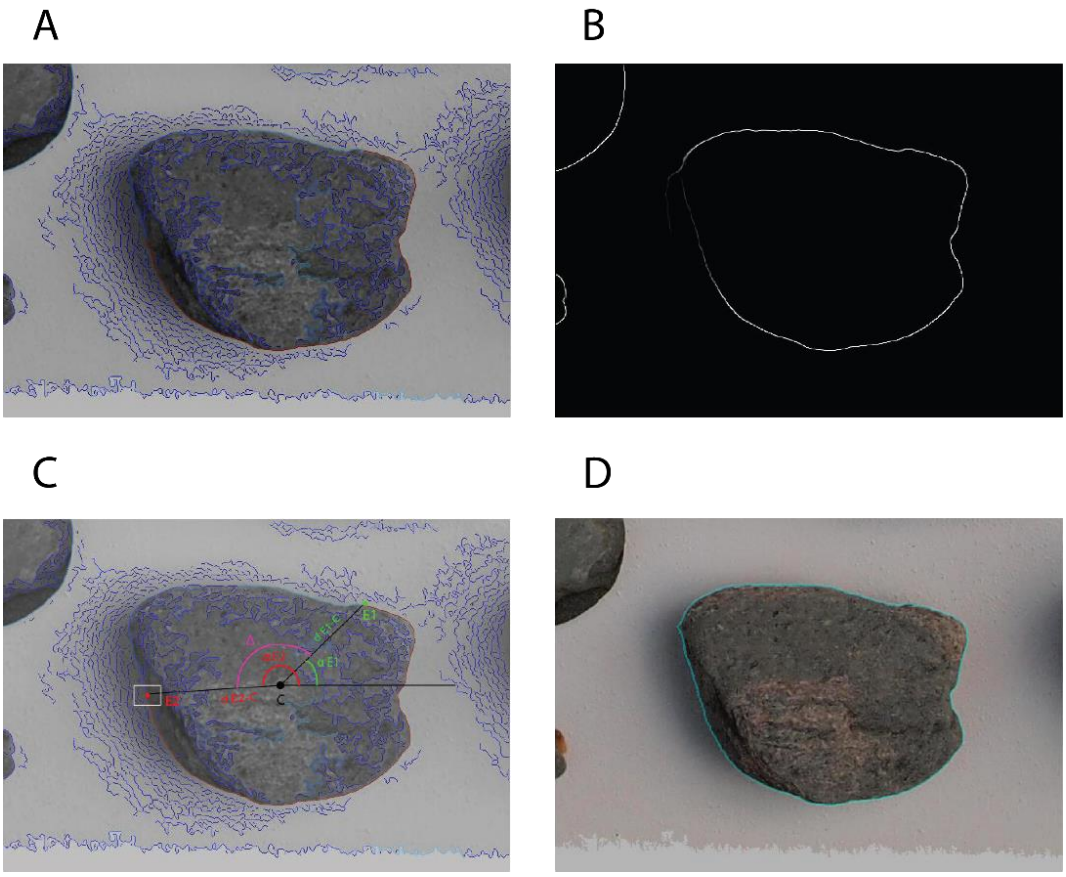

Figure 4-6 A) The edge identified as the longest connected component (red curve) at the end of the first phase of the particles' outline tracking. B) Geometrical coordinates of the two endpoints of the identified longest connected component. C) The frequency map represented as a grey-scale image. D) Final particle's outline overlapped to the particle's representation for the visual inspection.

4.1.3. Calibration and Validation

The model result dependent on a set of parameters: the size of the window built around each endpoint, the standard deviation of the Gaussian filter (σ_C) and the lower and upper thresholds on edges strength (C_l and C_u respectively). To calibrate the model, these parameters have been varied in this range:

$$0 \leq C_l, C_u \leq 1$$

$$0.2 \leq \sigma_C \leq 5$$

$$9 \leq \sigma_C \leq 21$$

The segmented outlines resulting by the application of the model with each combination of the calibration parameters have been processed using the toolboxes developed by Roussillon et al., 2009 and Zheng and Hryciw, 2015 for the evaluation of shape parameters.

Its results have been compared with those obtained by the application of the same toolboxes to pebbles segmented using the "Quick Selection tool" available in Adobe Photoshop CC, which have been considered as the reference values (AP). We used the relative error of elongation, circularity and roundness, three common shape descriptors, as accuracy metrics:

$$\varepsilon = \frac{\beta_{AP} - \beta_{PC}}{\beta_{AP}} \cdot 100 \qquad 4\text{-}1$$

Where β_{AP} is the value of the shape descriptor obtained using the Adobe Photoshop CC segmented outline (AP) and β_{PC} is the value of the shape descriptor obtained using the probabilistic Canny segmentation algorithm (PC). We also assessed model performances using an overall error metric defined as:

$$E = \sqrt{\varepsilon_e^2 + \varepsilon_C^2 + \varepsilon_{R_W}^2} \qquad 4\text{-}2$$

Where ε_e is the relative error with respect to particles' elongation, ε_C is the relative error with respect to particles' circularity and ε_{R_W} is the relative error with respect to particles' roundness. The definitions of these three shape descriptors are reported in Section 2.4.

5 pebbles of the sample have been used for the model calibration, selected to be different in shape and color. The Root Mean Square Error of the overall error metrics E, each one evaluated for each particle, has been used to identify the combination of the calibration parameters that guarantee the highest accuracy.

The calibration parameters selected as those that minimize the Root Mean Square Error are:

(1) The minimum and the maximum values, between which C_l varies, respectively equal to 0.1 and 0.2;
(2) The value selected as C_u to 0.3;
(3) The minimum and the maximum values between which σ_C varies, respectively equal to 0.2 and 5;
(4) l equal to 17 pixel for each side.

The calibrated model has been then applied to 32 pebbles for its validation and the accuracy of the model on the estimation of Circularity, Roundness and Elongation has been evaluated. The results are shown in Section 5.2 where the influence of both the segmentation methods and computational geometry toolboxes have been discussed.

4.2. Sediment fingerprinting using particles morphometry

The second phase of this work aims to assess to what degree particles shape data, estimated with the model implemented in the first phase of the research, could be useful for inferring the distance travelled by particles.

According to one of the studies on sediment attrition a "universal" relation between particles relative mass-loss and their circularity exists. Moreover, it suggests that this relation might be useful for inferring the distance traveled by individual particles (Novák-Szabó et al., 2018).

Since the formal equation has not been given to date, we assumed an elliptical function between the relative mass loss (μ) and circularity *(C)* that meets the conditions summarized in Section 2.3.1 and that interpolates the results obtained during the laboratory experiences conducted by Domokos et al. (Equation 4-3). The unknown parameters of the curve, i.e. C_{max}, C_o, μ_{max}, and α (Table 4-1), have been estimated fitting the experimental curve by minimizing the root mean square error between the estimated and experimental values of circularity (Figure 4-7).

$$C = C_0 + (C_{max} - C_0)\left[1 - \left(1 - \frac{\mu}{\mu_{max}}\right)^{\alpha}\right]^{1/\alpha} \qquad 4\text{-}3$$

Interpolating curve parameters	
C_{MAX} [-]	0.88
C_0 [-]	0.7
μ_{MAX} [-]	1
α	1.75

Table 4-1 Estimated parameters of the abrasion curve

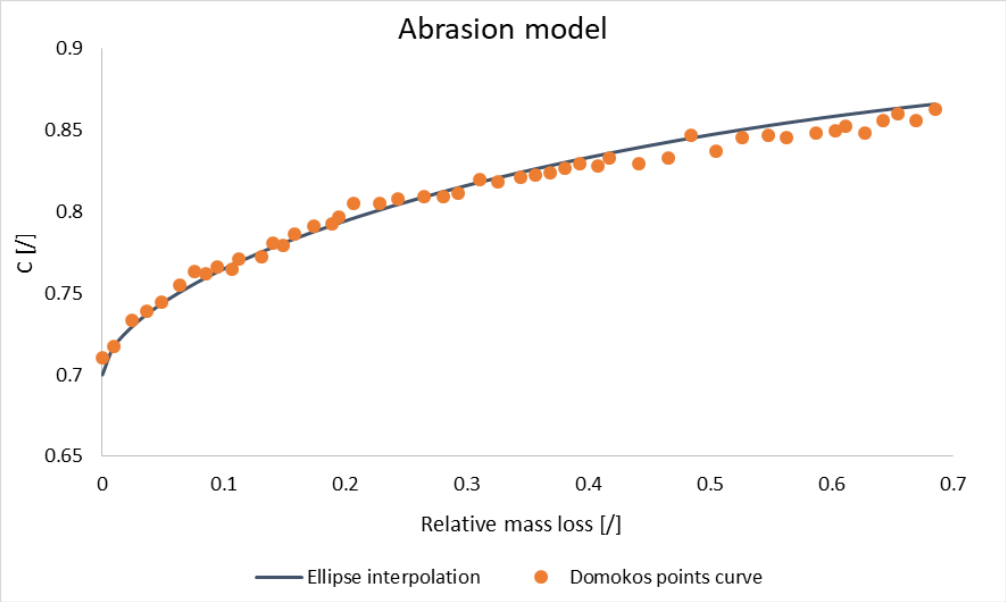

Figure 4-7 Abrasion curve

The interpolating curve has been used for the inference of the mean relative mas-loss from the measured circularity. Since the circularity value, corresponding to a zero relative mass loss, derived from the curve shown in Figure 4-7, differs from the mean value calculated from the field data collected at the points identified as the source, the curve has been shifted to the left (Equations 4-4 4-5).

This discrepancy between the laboratory data, obtained from the experiments with a tumbling barrel (Novák-Szabó *et al.*, 2018), and field data can be justified assuming that the sediments taken from the outcrops, located along the creeks that flow into the Sarzana river, are subjected to a so called *abrasion in place* (Schumm and Stevens, 1973). The abrasion of the sediments at rest, produced by vibratory motions to which the particles are subjected, has been suggested as possible factor that can cause mass loss (Schumm and Stevens, 1973).

$$\mu(C) = \mu_{max} - \frac{\mu_{max}}{(C_{max} - C_0)} * [(C_{max} - C_0)^{\alpha} - (C - C_0)^{\alpha}]^{\frac{1}{\alpha}} \quad 4\text{-}4$$

$$\Delta\mu = \mu(C) - \mu(C = 0) \quad 4\text{-}5$$

Nevertheless, the curve does not allow to infer directly the distance travelled by particles. In order to achieve the final objective of estimating the distance travelled from the shape metric, it is required to assess whether it would be possible to establish a theoretical relationship between the inferred relative mass-loss and travel distance.

Sternberg, in 1875, suggested an exponential the mass of the particles with distance, dependent on the unknown abrasion coefficient *k* (Krumbein, 1941) (Section 2.3.2).

Since the initial mass of the particles at the source points are unknown, an alternative could be the study of the exponential increase of the relative mass-loss of the particles, related to their initial mass, with their travel distance 4-6.

$$\mu = 1 - \frac{M}{M_0} = 1 - \exp(-kx) \quad 4\text{-}6$$

This abrasion coefficient has been evaluated from the increment of relative mass-loss, inferred through the Equation 4-4 and 4-5 for each sampling station located at a certain distance L from the source point (Equation 4-7).

$$k = -\left(\frac{1}{L}\right) * \ln\{1 - \mu\} \qquad 4\text{-}7$$

As will be demonstrated in Section 5.4 and as already mentioned in Section 2.3.2 the speed with which the particles abrade is expected to depend on different feature of the basin and of the pebbles, e.g. grain size *et cetera*.

The accuracy of the estimation of the distances travelled by particles, deduced as a function of circularity and abrasion coefficient, is related to the dispersions of the sampled particles' circularity, linked to the internal variability of the sample, and to the accuracy of the measures of circularity.

Should the uncertainties related to the transformation function from circularity to travel distance be too wide, this could lead to the paradox of estimating travel distances several times higher than the distance between the farthest sediment source for a given lithology.

At the same time, since the distance is a function of circularity *(C)* and abrasion coefficient *(k)* (Equation 4-8), its accuracy is dependent on the accuracy in the measures or estimation of these parameters. Their effects can be assessed studying the derivative of L with respect to *C* and *k* (Equations 4-9 and 4-10). To make a general analysis about the sensitivity of L to variations f *C* and *k*, it has been considered the non-shifted abrasion curve for determining *μ(C)*.

$$L = -\left(\frac{1}{k}\right) * \ln[1 - \mu(C)] \qquad 4\text{-}8$$

$$\frac{\partial L}{\partial C} = \frac{1}{k}\frac{\mu_{max}}{(C_{max} - C_0)}\frac{(C - C_0)^{\alpha-1}[(C_{max} - C_0)^{\alpha} - (C - C_0)^{\alpha}]^{\frac{1}{\alpha}-1}}{1 - \mu(C)} \qquad 4\text{-}9$$

$$\frac{\partial L}{\partial k} = -\frac{1}{k^2}\ln[1 - \mu(C)] \qquad 4\text{-}10$$

The results are shown later in Sections 5.5.1 and 5.5.2, where the influence of the accuracy of the estimation of both the abrasion coefficient and circularity on the distance travelled inference are discussed.

5. Results

5.1. Sarzana River Basin

The study area is the Sarzana River basin, a 25 km² located in the North-East of Italy. The Sarzana River is a tributary of the Cordevole River and it has been chosen since, according the Veneto Region lithologic map ('Veneto Region Geoportal https://idt2.regione.veneto.it/') it is characterized by the presence of localized and well separated sources of metabasalts (Figure 5-1) and arenites. In this research, only the metabasalts will be analysed and considered as tracers.

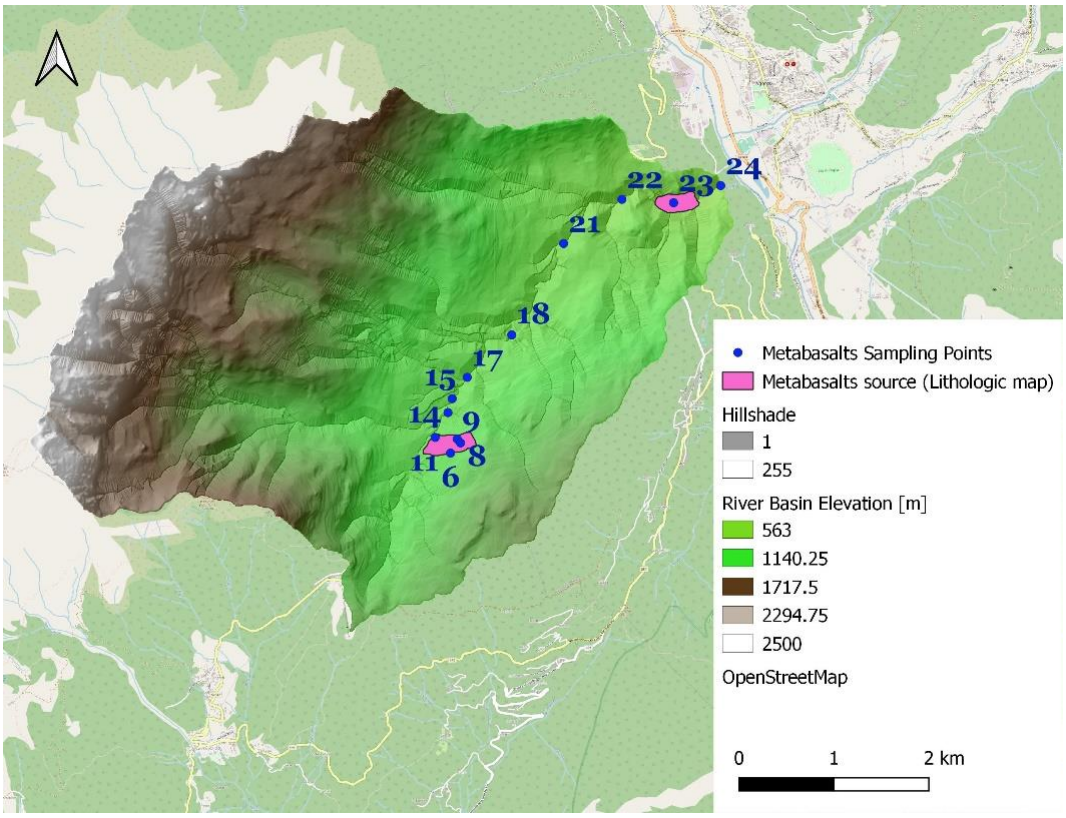

Figure 5-1 Sarzana River Basin

5.1.1. Field Data Collection

The fieldwork experience showed that, unlike what suggested by the lithologic map, most of the contribution of this rock type is due to punctual outcrops of metabasalts placed within small creeks draining the hillslopes (Figure 5-2).

In particular, we found two outcrops along the creeks that cross the region indicated by the lithological map, both responsible for the production of fresh fragments of metabasalts (Figure 5-3).

Sediment particles have been collected along the river network from exposed bars. Their size ranges from cobbles to fine gravel. Before placing them on a 0.5x1 m^2 white board to enhance the contrast, the pebbles have been washed, to remove soil and vegetation deposits that might affect their shape assessment and dried using normal towels. Pebbles were therefore placed spaced from each other, avoiding potential shadow clustering and pebble their projections' overlapping, which might bring to misclassifications. Pictures of the white board were taken using an Olympus® E-PL1 digital camera (with an average resolution of 0.26mm/px).

Unfortunately, due to the low relative presence of metabasalts deposited in some of the sampling stations, with respect to the rest of the lithology composing the bed of the Sarzana River, it has been challenging to achieve large sample sizes.

Figure 5-2 Two observed outcrops of metabasalts inside the region identified as a source by the lithologic map

Figure 5-3 Fresh fragments took from the detected outcrop

5.2. Pebbles' segmentation model results and discussion

The images taken during the fieldwork were then processed with the implemented model for edge detection (Section 4.1).

At the end of the segmentation model, all the images of the particles that passed the visual inspection were analyzed using the computational geometry toolboxes mentioned in Section 2.4.1. Table 5-1 summarizes the mean relative errors ε obtained comparing the shape metrics, estimated for the images segmented through the implemented model $\beta_{(PC)}$, with the reference values $\beta_{(AP)}$ (Equation 4-1).

	ZHENG	ROUSSILLON
ε_C	0.81%	0.77%
ε_E	0.59%	0.88%
ε_R	5.65%	6.22%

Table 5-1 Relative errors evaluated for each shape parameter through each computational geometry models.

Roundness is the parameter that, irrespectively of the shape estimation method used, showed the highest relative error in its estimation. Circularity and elongation results showed instead good agreement with reference values below 1%. Table 5-2 shows the relative difference in the estimation of each shape descriptor depending on the toolbox used (Equation 5-1). The index evaluates the relative error between the values of a generic parameter β inferred using the toolbox implemented by Roussillon et al., 2009 (R) and those obtained using the toolbox of Zheng and Hryciw, 2015 (Z):

$$\varepsilon_{Z/R\_\beta} = \frac{\beta_{Rous.} - \beta_{Zheng}}{\beta_{Rous}} \cdot 100 \qquad 5\text{-}1$$

	AP	PC
ε_{Z/R_C}	0.33%	0.33%
ε_{Z/R_E}	3.24%	3.43%
ε_{Z/R_R}	15.15%	14.35%

Table 5-2 Relative errors between the two computational geometry models (Z and R) evaluated for each parameter both for manually segmented (PC) and automatically segmented (AP) pebbles.

These results show that, for all the shape descriptors considered, the differences between the estimation of the two models is not affected by the segmentation method used, whose accuracy in pebbles' segmentation can be considered to be equivalent to the manual "Quick selection tool" available in Adobe Photoshop CC.

Figure 5-4 show six scatter plots that allow the comparison of the effect of segmentation method and of the shape analysis toolbox used on the accuracy of the estimation of the shape metrics.

The dispersion of data points when comparing the two toolboxes is similar, irrespectively of the segmentation method used (Figure 5-4 A-C-E).

When data points are compared according to the segmentation method, they show similar dispersion independently on which computational geometry toolbox is applied (Figure 5-4 B-D-F). As observed, the most error prone shape parameter is the roundness. Nevertheless, the estimation of this parameter seems to be more dependent on which toolbox is used for the shape characterization than the segmentation method adopted. Indeed, in Figure 5-4 F, the roundness data are more distributed along the bisector than in Figure 5-4 E. Regarding circularity and elongation, data show comparable dispersion when grouped according to both criteria. The shape analysis toolbox adopted and the segmentation method used therefore had comparable effects in terms of their contribution to the accuracy of shape values estimation.

This model represents a step forward in the shape characterization. Indeed, the existing computational geometry toolboxes cannot automatically detect the pebbles' outline. In particular, the Roussillon et al. (2009) model is the only computational geometry toolbox, among those considered for this work, that involves both segmentation of the pebbles and the computation of the shape parameters. Nevertheless, it resulted to be less accurate in the segmentation process than the automatic segmentation model here implemented Figure 5-5. Comparing the results obtained for each shape metrics in terms of mean relative error calculated with respect of the references values (AP), the new automatic segmentation model (PC) represents a great improvement Figure 5-5.

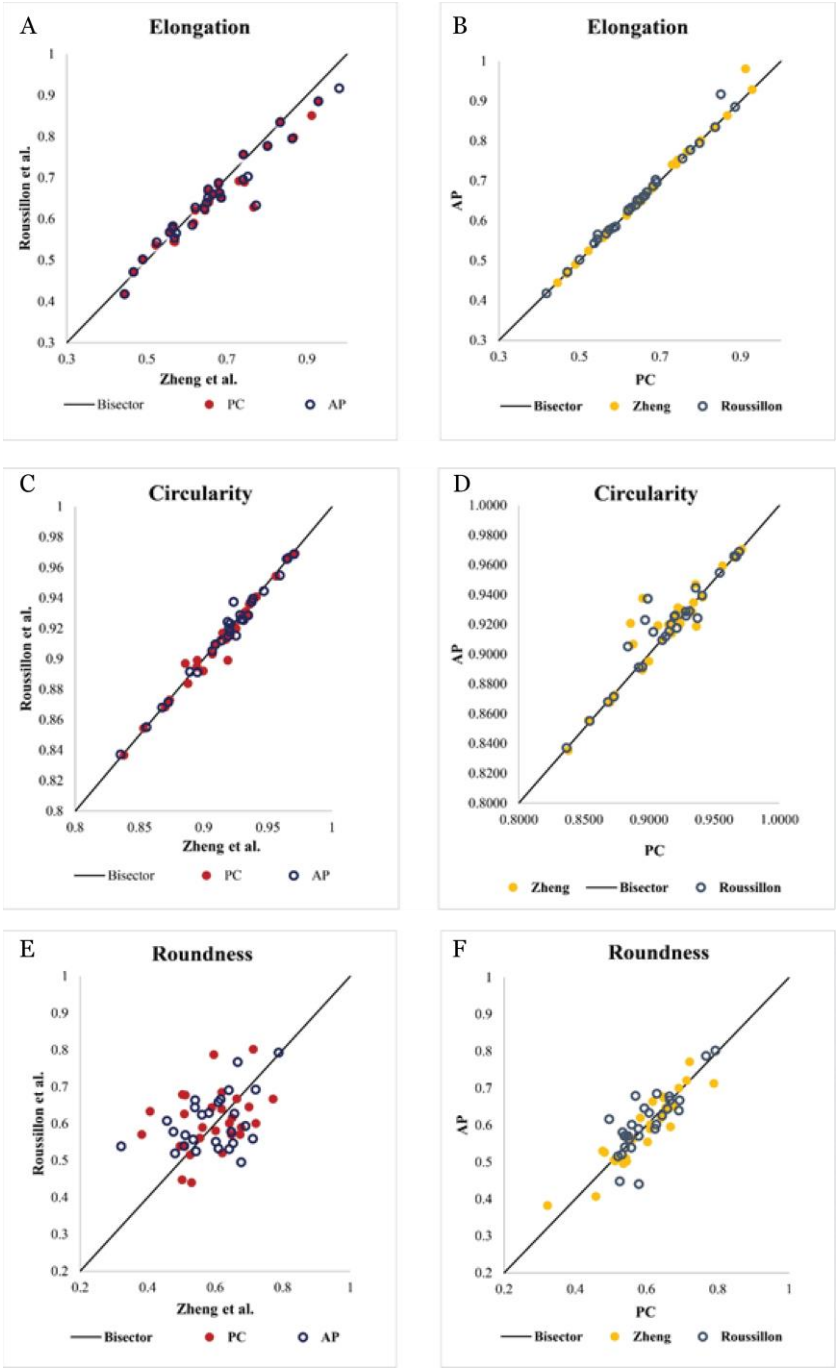

Figure 5-4 A-C-E) Graphical representation of the discrepancy, between the two computational geometry toolboxes, on shape parameters estimation. The blue empty dots represent the shape parameters evaluated for the pebbles manually segmented (AP), while the filled red ones represents the shape descriptors for particles automatically segmented and considered as reference values(AP). B-D-F) Graphical representation of the influence of the adopted segmentation technique on the evaluation of shape parameters using the toolbox of Roussillon et al., 2009 (grey empty dots) and the one of Zheng and Hryciw, 2015 (yellow filled dots).

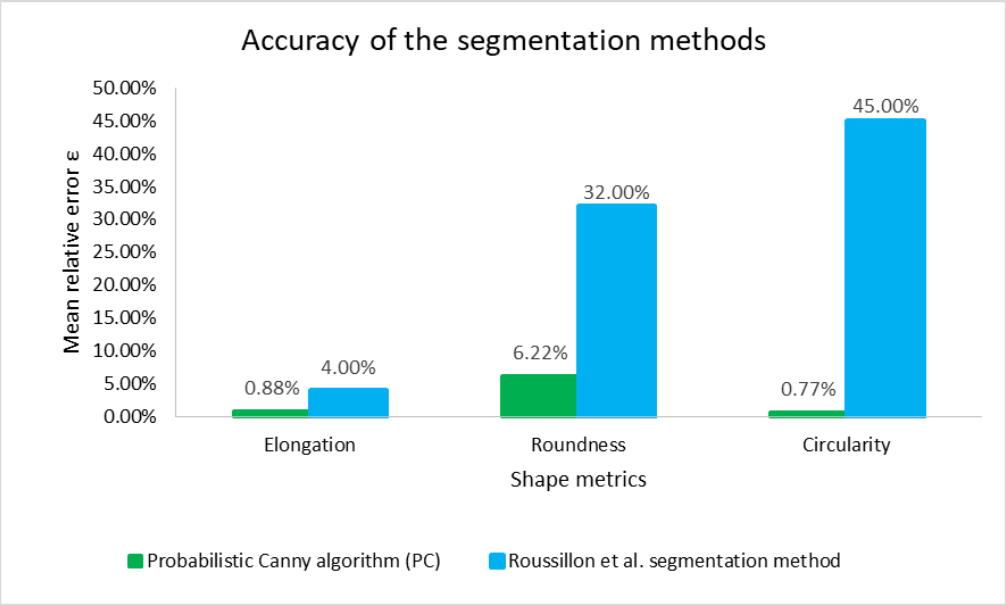

Figure 5-5 Comparison of the accuracy (expressed as mean relative error ε) of the new probabilistic Canny segmentation model (PC) in the estimation of shape parameters, with that relative to Roussillon et al. segmentation method.

5.3. Empirical evolution of circularity with distance

During the fieldwork, the coordinates of the sampling locations were measured using a mobile phone app: Mobile Topographer ('Mobile Topographer') . The distances of the sampling locations with respect to the outlet of the basin have been estimated as follows.

The input data was the DEM provided by the Veneto Region Geoportal ('Veneto Region Geoportal https://idt2.regione.veneto.it/'). Through an application of the Software ArcGIS, it has been possible to define, for each point of the basin, the flow direction and the flow distance with respect to the outlet of the basin. The distance travelled for each source point was obtained by clipping the raster representing the flow distances for each basin pixel with the shape file of the region identified as the source area by the lithologic map Figure 5-1 ('Veneto Region Geoportal https://idt2.regione.veneto.it/', no date).

The empirical evolution of mean circularity with the distance has been tracked for each of the two paths connected to the two outcrops identified as possible metabasalts sources (Points 6 and 8 in Figure 5-6).

Since the mean value of circularity has been evaluated from the collected data, its accuracy could be compromised by the size of the sample.

The smaller the sample, the less accurate the estimation of the mean circularity. As reported in the Table 5-3 and in Table 5-4 and shown in Figure 5-7 and Figure 5-8, where the trends of the mean circularity and the extremes of the confidence interval are drawn, some of the sampling stations do not reach 30 pebbles, making the inference less accurate.

Both trends seem similar to the expected evolution from visual inspection. Under the assumption that the dominant abrasion mechanism is chipping (Section 12.3.1), the main variation in particles' circularity is attended when pebbles have travelled a few kilometres (Miller et al., 2014). This is justified by the fact that near the source particles are more angular and, therefore, abrasion rate is higher. On the contrary, after a certain amount of distance travelled, they have been already rounded and a decrease in the rate with which the circularity varies is expected. This shape parameter is therefore expected to reach a steady value and field observations qualitatively agree with the theoretical expectation.

Despite what one could expect, around the first kilometre, near the point where the creeks coming from the hillslopes flow into the Sarzana River, circularity decreases (Figure 5-7 and Figure 5-8). Nevertheless, the uncertainty associated to the estimated mean value, i.e. the extension of the confidence interval evaluated at 95% confidence, is greater than the variation of the mean circularity measured between two consecutive stations. This result can be justified by the paucity of the available sample in some of the locations.

Furthermore, at around the fifth kilometre from the upstream source points, the trend undergoes a rapid decrease due to the presence of a downstream metabasalts source, showed in Figure 5-6, that produce less abraded and sharper fragments.

Besides, we wondered whether it would be possible to set a boundary condition in the process of fragmentation at the source.

We found that the distribution of the shape properties (circularity, roundness, and elongation), estimated for the two outcrops, were similar, suggesting that the breakage mechanisms of metabasalt rocks in such environments tends to produce fragments with similar shape distribution. A Kolmogorov-Smirnov test, performed on the two samples at 95% confidence, confirmed the null hypothesis of compatibility of the two distributions, of the metabasalts outcrops Figure 5-9. This result allowed us to simplify the analysis and to refer to a unique source obtained combining data referred to the two sampled outcrops (stations 6 and 8) and called in the next Sections as point 4 Table 5-5.

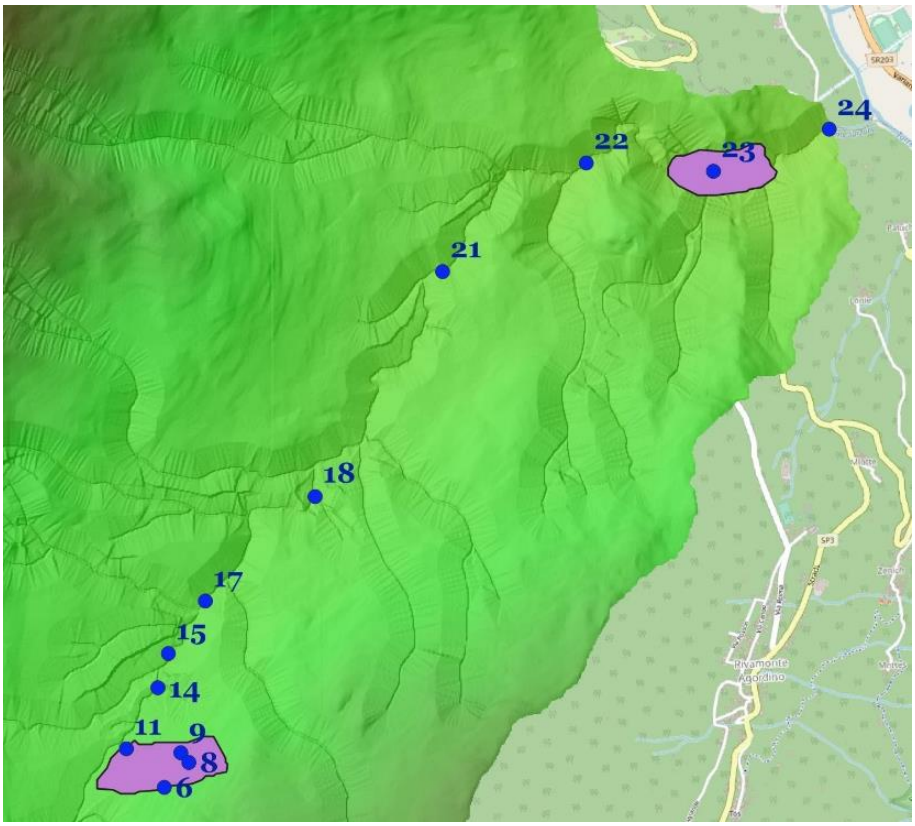

Figure 5-6 Representation of the sampling locations. The two lines represent the two creeks that flow into the Sarzana River.

Smapling Location	N Samples	Distance from the outlet [m]	Distance from the source [m]	mean circularity [-]
6	41	5714.8	0	0.771
11	35	5483.12	231.68	0.814
14	57	5141.38	573.42	0.858
15	32	4960.66	754.14	0.853
17	50	4653.88	1060.92	0.848
18	27	3837.53	1877.27	0.850
21	84	2436.12	3278.68	0.860
22	68	1474.06	4240.74	0.876
23	41	624.778	5090.022	0.851
24	54	0	5714.8	0.866

Table 5-3 Mean circularity data relative to the stations located along the path 1

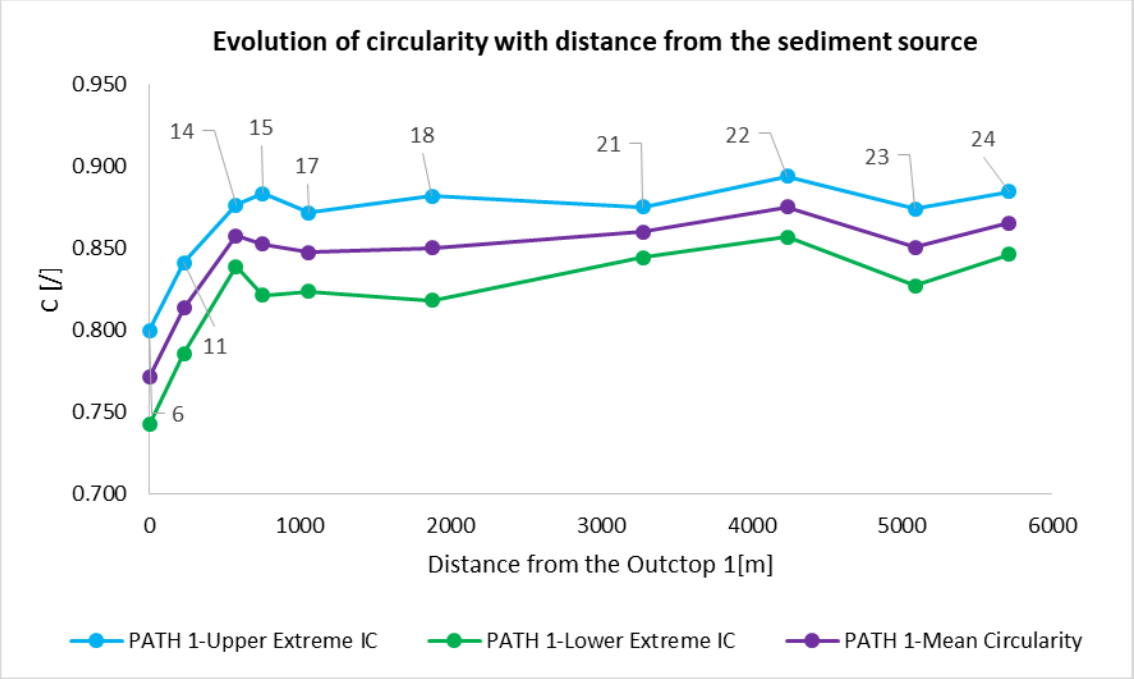

Figure 5-7 Evolution of the mean circularity with the distance travelled by particles along the path 1

Smapling Location	N Samples	Distance from the outlet [m]	Distance from the source [m]	mean circularity [-]
8	55	5612.87	0	0.752
9	25	5546.3	66.57	0.815
14	57	5141.38	471.49	0.858
15	32	4960.66	652.21	0.853
17	50	4653.88	958.99	0.848
18	27	3837.53	1775.34	0.850
21	84	2436.12	3176.75	0.860
22	68	1474.06	4138.81	0.876
23	41	624.778	4988.092	0.851
24	54	0	5612.87	0.866

Table 5-4 Mean circularity data relative to the stations located along the path 1

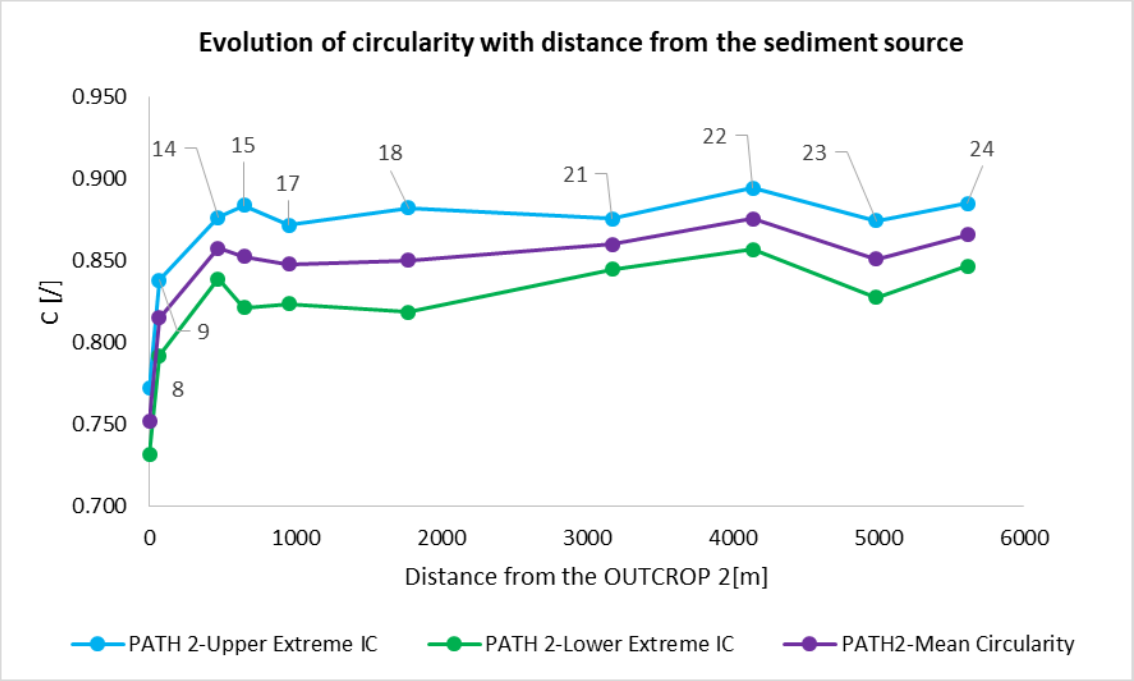

Figure 5-8 Evolution of the mean circularity with the distance travelled by particles along the path 2

Smapling Location	N Samples	Distance from the outlet [m]	Distance from the source [m]	mean circularity [-]
4	96	5663.84	0.00	0.760
9	25	5546.30	117.54	0.815
11	35	5483.12	180.72	0.814
14	57	5141.38	522.46	0.858
15	32	4960.66	703.18	0.853
17	50	4653.88	1009.96	0.848
18	27	3837.53	1826.31	0.850
21	84	2436.12	3227.72	0.860
22	68	1474.06	4189.78	0.876
23	41	624.78	5039.06	0.851
24	54	0.00	5663.84	0.866

Table 5-5 Mean circularity relative to the downstream sections and to the source point 4 obtained by the combinations of the data relative to the sampled outcrops located at points 6 and 8.

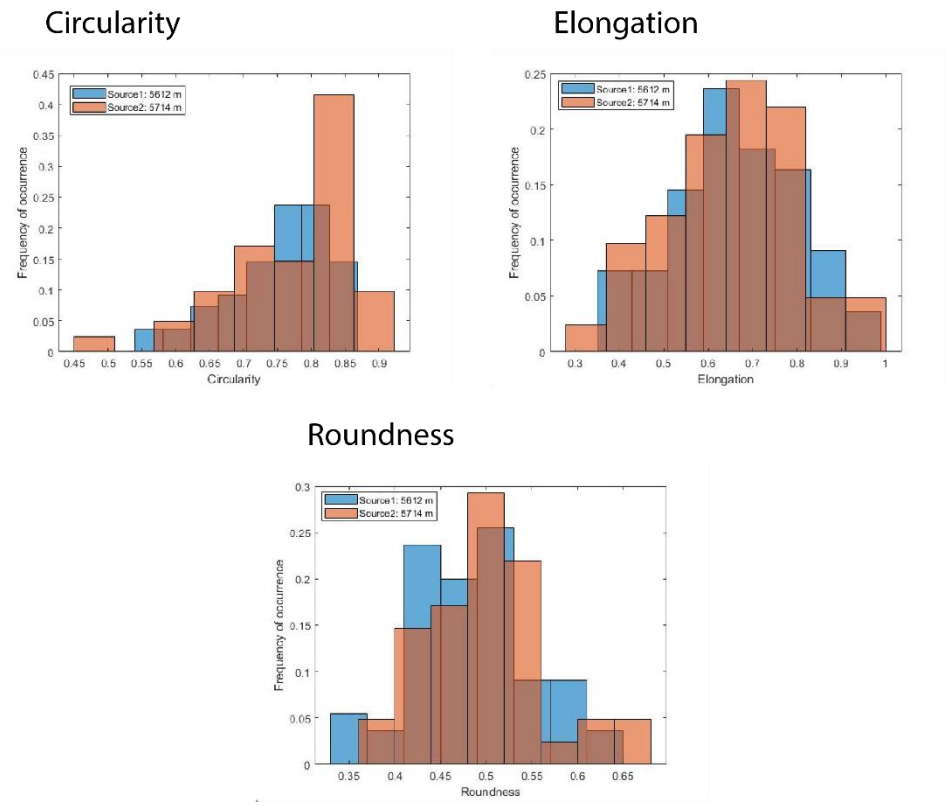

Figure 5-9 Distributions of circularity, elongation and roundness data collected in the two sampled outcrops

5.4. Estimation of the abrasion coefficient

Sternberg suggested an exponential increase of particles' relative mass loss during the travel. The coefficient that determines the speed with which the mass varies and therefore the particle is abraded, is *a priori* unknown. The inference of the abrasion coefficient from the field data can be obtained relating the relative mass loss and the travel distance. The relative mass loss is deduced the circularity data to which is related through the abrasion curve interpolated on the experimental data Section 4.2. As already mentioned in Section 4.2, the curve has been shifted since the circularity data obtained during the experiments carried out on a tumbling barrel proved to be underestimated compared to the circularity values obtained in the field at the points identified as a source (Sampling point 4 in Table 5-5. This, indeed, leads to have analytically, from the Equation 4-4, a positive value of mass loss at the source. The deduced increment of relative mass-losses $\Delta\mu$, evaluated with respect to the mass loss at the source (Equation 4-5), are reported in Table 5-6.

Only the sampling points located along the Sarzana River have been considered for this analysis. Furthermore, the location 18 has been neglected, since as mentioned in the previous Section 5.3, the paucity of the sample makes the analysis less accurate.

Sampling Points	Distance from the outlet [m]	mean circularity [-]	Relative mass loss [-]	Δ Rel. mass loss [-]
4	-	0.7602	0.087	
14	522.5	0.858	0.593	0.506
15	703.2	0.853	0.545	0.458
17	1010.0	0.848	0.504	0.417
21	3227.7	0.860	0.617	0.530
22	4189.8	0.876	0.833	0.746
23	5039.1	0.851	0.531	0.444
24	5663.8	0.866	0.681	0.594

Table 5-6 Table containing data about mean circularity, relative mass loss, found from the abrasion curve, and the increment of relative mass loss with respect to the mass loss estimated from the interpolating curve

Therefore, the inferred increment of relative mass-losses, Δμ in Equation 4-5, have been related to the measured distance through the Sternberg law to estimate the abrasion coefficient that represents the rate with which the sediment mass decreases with the distance Equation 4-7. This process has been repeated for all the sections located along the Sarzana River. The estimated mean abrasion coefficients represent the global rates with which the particles are abraded from the source to each point located downstream. Table 5-7 collects the abrasion coefficient (k), which is not constant along the river course but it decreases going farther downstream.

Loc_Name	Distance from the US source [m]	k [km^{-1}]
Source	-	-
14	522.455	1.351
15	703.175	0.872
17	1009.955	0.535
21	3227.715	0.234
22	4189.775	0.327
23	5039.057	0.117
24	5663.835	0.160

Table 5-7 Mean abrasion coefficient k inferred for each location

Figure 5-10 shows the curve of abrasion coefficient *(k)* descending steeply at first and then more gradually as the distance from the source increases. Kuenen, in 1956, classified different types of abrasion, which affect the mass loss of the particles during the different phases of transport (Kuenen, 1956). The dominant abrasion process, in gravel-bed rivers, is chipping, which causes the loss of small flakes from the sharp edges, increasing the particles' circularity and roundness (Krumbein, 1941; Kuenen, 1956). Its contribution in the evaluation of the abrasion coefficient is relevant especially near the source where the pebbles are more angular and they are more prone to be abraded by chipping. Past studies on abrasion suggest the size of the particles as another effect on the rate with which mass decreases (Krumbein, 1941; Kuenen, 1956; Bradley, 1970). Indeed, it is expected that the bigger particles abrade faster.

The weight of the particles affects their kinetics and, therefore, assumed that the main transport mechanism for bed-load is saltation, the jump lengths of the particles (Krumbein, 1941; Kuenen, 1956; Bradley, 1970). The larger the size, the smaller the jump length. Therefore, the bigger particles impact more frequently with the bed and are more prone to collide with the transported particles than the smaller ones (Krumbein, 1941; Frings, 2008). Moreover, since the smaller the particles are, the farther downstream they travel, the effect of size on the abrasion rate, added to the one produced by the irregular shape, is expected to enhance the values of the coefficient in the most upstream part of the river and reduces them downstream.

It is, therefore, expected a decreasing abrasion rate with distance as a result of the just mentioned effects (Frings, 2008). Figure 5-11 shows both the evolutions of the mean size and abrasion coefficient with distance traveled. The abrasion coefficient seems to decrease with the mean size, suggesting a strong correlation between the two quantities. The trends are quite parallel until the 5th kilometer where the presence of the downstream source makes the mean size greater but, at the same time, the production of fragments reduces the circularity and, therefore, the abrasion coefficient.

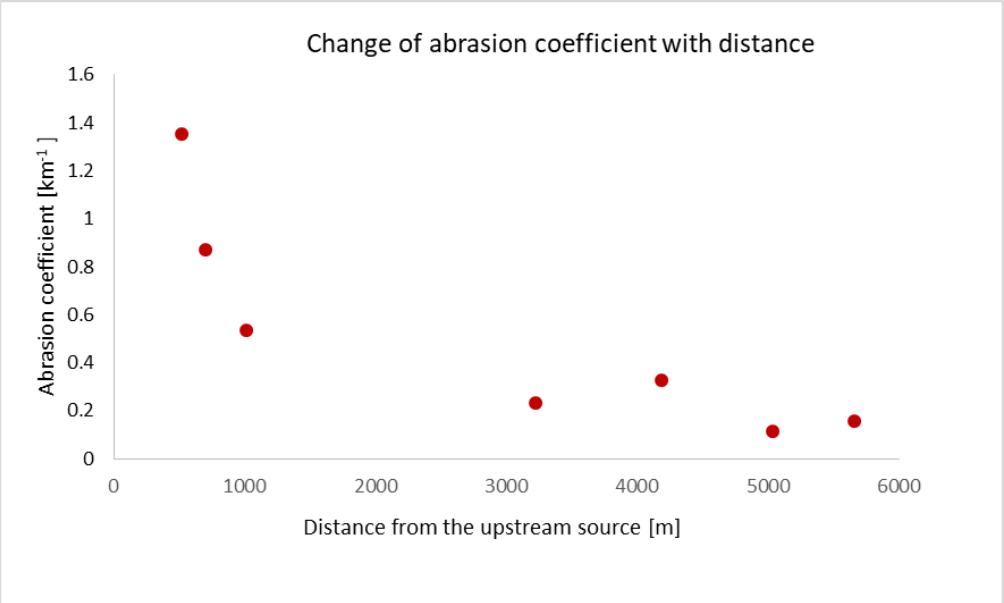

Figure 5-10 Downstream change in the abrasion coefficient

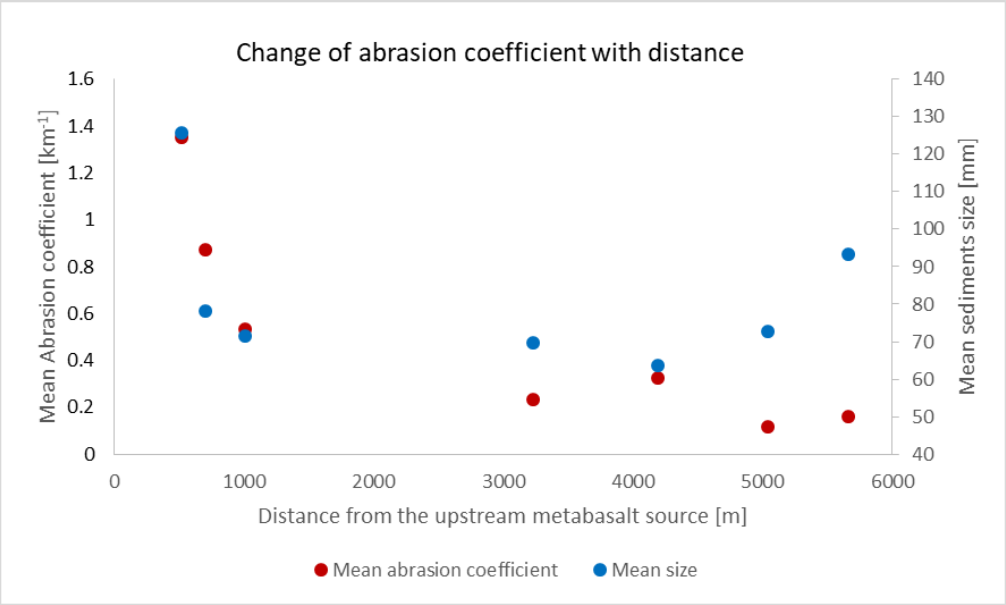

Figure 5-11 Evolution with distance of abrasion coefficient and mean size at each sampling location

5.5. Sensitivity analysis

The abrasion coefficient represents the speed with which metabasalts particles abrade as they are transported by the river flow. The values of this parameters relate the mean value of circularity and the measured distance from the source point. The next sections are focused on assessing how much the internal variability of the circularity values, due to the natural dispersion of the sample data, can affect the measure of the distance travelled. Moreover, the effects of the accuracy of the methods used for evaluating the particles' circularity values and the accuracy of the abrasion coefficient estimation on the estimation of the distance travelled by particles are analyzed.

5.5.1. Effects of the internal variability of the circularity data on the estimation of the possible distances travelled by particles

Within a sample, each particle has specific morphometric characteristics which can differ from the mean value of the sample.

The internal variability of the sample can be represented by the standard deviation. Table 5-8 summarizes the statistics of the samples relative to each location.

Location Station	Std C [-]	mean C [-]
14	0.071	0.858
15	0.086	0.853
17	0.084	0.848
21	0.071	0.860
22	0.077	0.876
23	0.074	0.851
24	0.070	0.866

Table 5-8 Mean and standard deviation and relative standard deviation of circularity measures for each sampling location

In the samples some of the values do not belong to the range $C_{MAX} - C_{SOURCE}$ where C_{SOURCE} is equal to 0.76 and C_{MAX} is 0.88. For this reason, for these particles it is not possible to estimate the mass loss $\Delta\mu$ using the curve derived from Domokos dataset and adapted to the inferred mean circularity values (Equations 4-4 and 4-5). As a consequence also the associated distance travelled cannot be directly computed. It has been assumed that for the particles with a circularity lower than C_{SOURCE} the relative mass loss is null as the travel distance, while, for values of circularity greater than C_{MAX}, the relative mass loss is equal to one and the correspondent distance is infinite. To associate a finite distance value to the most circular particles, it has been defined a maximum value of distance L_{MAX}, which corresponds to the maximum value of circularity belonging to the mentioned range. In this way, all the particles with circularity values higher than C_{MAX} result to have travelled a distance equal to L_{MAX}.

Under these assumptions and considering the abrasion coefficients estimated in Section 5.4, the application of the Equation 4-8 to all the sampled circularity values leads to a distribution of the travel distances characterized by a great dispersion as reported in Table 5-9. The standard deviation of the distance results to be of the same order of the mean value.

Location Station	Std L [m]	mean L [m]
14	586.066	522.455
15	686.883	703.175
17	1363.487	1009.955
21	3288.509	3227.715
22	2202.166	4189.775
23	5695.214	5039.057
24	4110.315	5663.835

Table 5-9 Mean and standard deviation of distance measures for each sampling location

To compare and underline the effects of the sequence of increasing functions on the internal variability of the sample it has been evaluated, for both the distance and circularity, the relative standard deviation:

$$Relative\ standard\ deviation = \frac{\sigma}{X} \qquad 5\text{-}2$$

Where X is the mean of the quantity analyzed and σ its standard deviation.

The normalized standard deviations estimated for the distance travelled result to be at least 6 times greater than the internal variability measured for circularity Table 5-10.

Location Name	Relative std L [-]	Relative std C [-]	ratio [-]
14	1.122	0.083	1.4E+01
15	0.977	0.101	9.6E+00
17	1.350	0.099	1.4E+01
21	1.019	0.083	1.2E+01
22	0.526	0.088	6.0E+00
23	1.130	0.087	1.3E+01
24	0.726	0.081	9.0E+00

Table 5-10 A-dimensional standard deviation of circularity C and travel distance L

According to the lithologic map, two diffuse sources of metabasalts are located within the Sarzana River Basin.

This means that all the points, belonging to these areas produce metabasalts fragments that travel into the Sarzana River following different paths, each one characterized by a flow distance from the outlet of the basin evaluated, as explained in Section 5.3, along the direction of the steepest descendent Figure 5-12. The analysis of a diffusive source allows to consider a distribution of the possible paths that particles can travel. The sub-basins underlying the most upstream sections, located along the Sarzana River, intersect the upstream metabasalt source. This means that the region partially contributes to the production of the sediments. The contributing area has been obtained by clipping, through the ArcGIS tool, the metabasalt source area with the sub-basin underlying the considered sampling location. Going downstream, the area that drains in the river section and, consequently, the number of possible paths increase. At the outlet, both the upstream and downstream sources drain, reaching the maximum number of possible along which sediments can travel. The variability of the path lengths distribution is expected to decrease approaching the outlet of the basin. Nevertheless, the presence of the downstream source makes the distribution more spread Table 5-11.

To evaluate the distance distribution, all points, belonging to the diffuse source, have been assumed to contribute with the same probability to the production of the fragments that flow into the Sarzana River.

The distribution of the path lengths, deduced from circularity and the distribution obtained considering a diffusive metabasalt source, have been comparerd under the hypotheses that all points belonging to the diffusive source contribute with the same probability to the sediment flow and applying the abrasion coefficient inferred in Section 5.4 from the mean data of circularity. Unlike what expected, the distributions do not overlap (Figure 5-13).

Furthermore, some of the distances evaluated from the circularity values are greater than the value of distance evaluated for the most far point belonging to the diffusive source, or lower than the distance measured for the closest source-point. This result is physically meaningless, since it suggests the presence of a source located more upstream, or *vice-versa* more downstream, than the one detected by the lithologic map of the basin Figure 5-13.

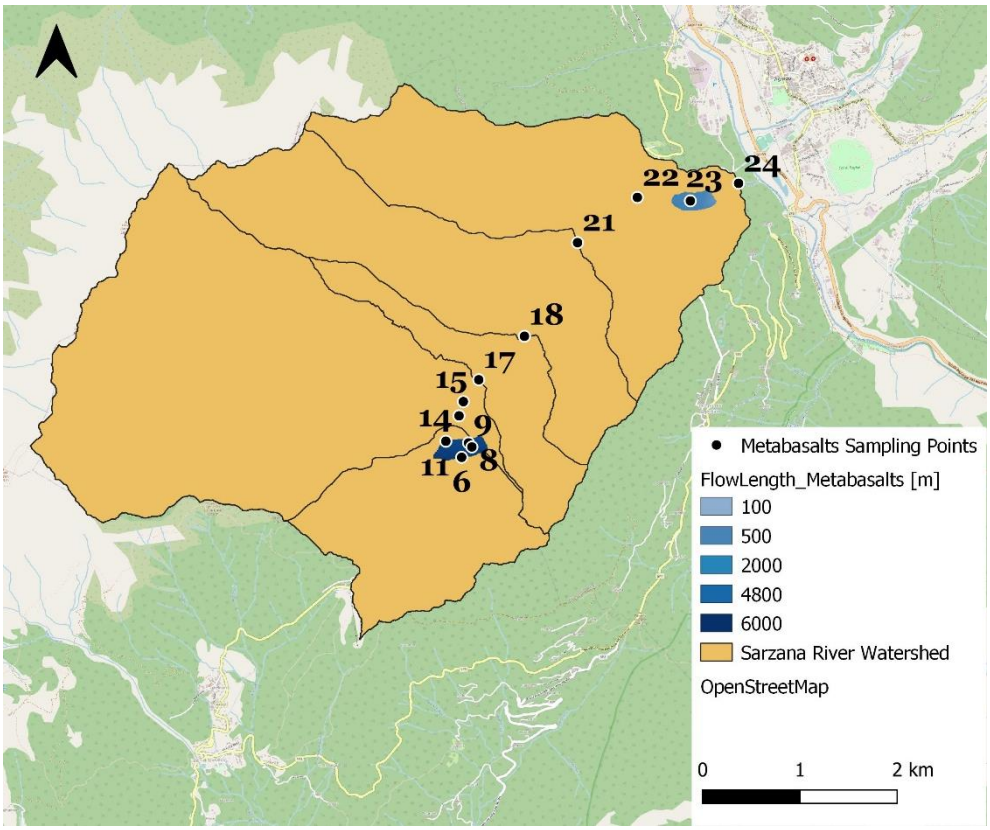

Figure 5-12 Sarzana River watershed and map of the distances associated to the points belonging to the metabasalt source.

Location Name	Mean L [m]	Std L [m]	Normalized std L [-]
14	481.49	80.11	0.166
15	652.90	81.66	0.125
17	959.68	81.66	0.085
21	3177.44	81.66	0.026
22	4139.50	81.66	0.026
23	3932.91	2008.24	0.485
24	3399.71	2488.29	0.633

Table 5-11 Statistics of the distribution of the possible path lengths deduced considering the diffusive source

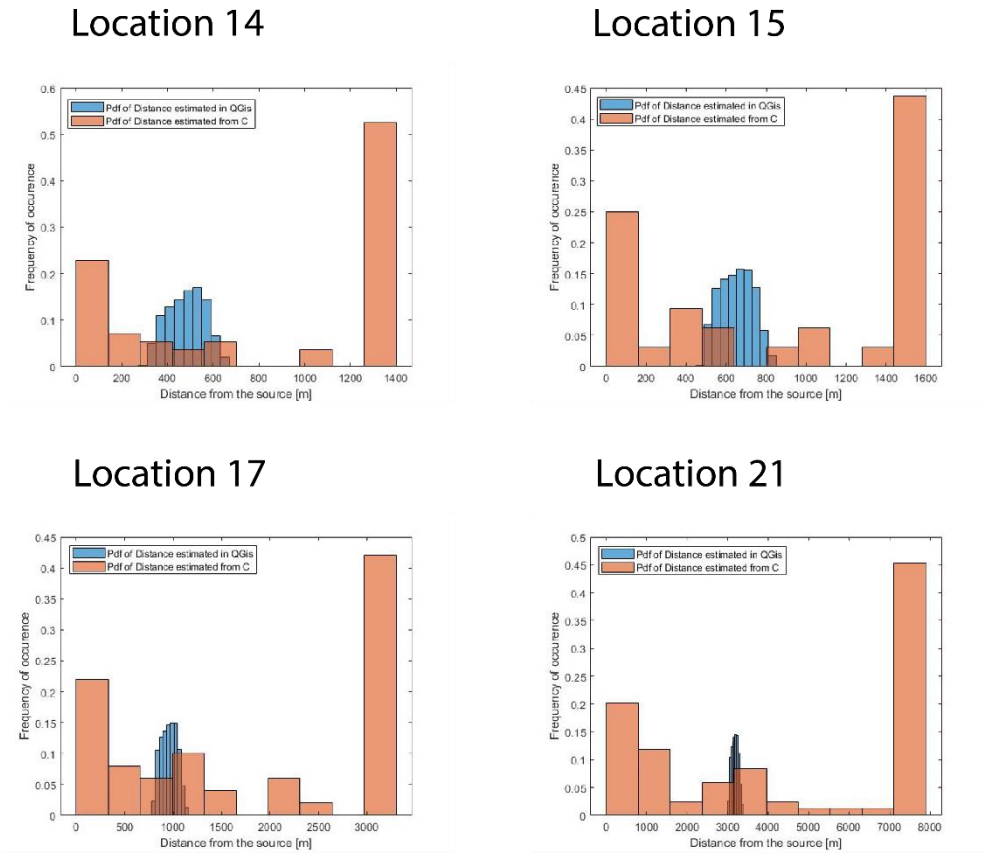

Figure 5-13 Distribution of the path lengths deduced from circularity values (red bars) and Distribution of the possible path lengths obtained considering the diffusive source of metabasalt (blue bars).

5.5.2. Effects of the uncertainties in the measures of circularity and abrasion coefficient on the distance inference

The accuracy of the inferred travel distances is a function of the accuracy of the measuring methods with which circularity and abrasion coefficient are estimated. As explained in Section 4.2, the accuracy of the travel distances can be evaluated by analyzing the first derivative of L with respect to C and k (Equations 4-9 and 4-10).

The accuracy of the models used here for the estimation of circularity was found to be approximately 1% (Section 5.2). This value was then used to assess its influence on the accuracy of the distance's estimation.

At the same time, since the ground-truth value of the abrasion coefficient is a priori unknown, it is not possible to evaluate the accuracy of the values inferred in Section 5.4.

It was then assumed that the accuracy associated with the estimated values of k, could be an order of magnitude lower than the value assumed by the parameter itself, i.e. ~10^{-2}. It has been first studied the influence of circularity, maintaining the abrasion coefficient set at 0.1 km^{-1} and varying circularity between C_0 and C_{MAX} reported in Table 4-1. Figure 5-14 represents the travel distance and its first derivative, both expressed as a function of circularity, its accuracy and the relative mass loss μ(C) Equations 4-8 and 4-9. Both the quantities increase with increasing values of circularity. Furthermore, the tendency of $\frac{dL}{dC}$ to infinite for higher values of circularity means that even small increment of Circularity leads to a great variation of distance ΔL (Equation 5-3).

$$\Delta L = \frac{dL}{dC}\Delta C \qquad 5\text{-}3$$

If we consider ΔC as the uncertainty of the measured C, the correspondent ΔL represents the accuracy in the distance inference and it depends on the values assumed by C as showed in Figure 5-15.

The absolute error of L is not significant since it depends on which value of distance is associated. For this reason, the relative error has been evaluated (Equation 5-4).

$$e = \frac{\Delta L}{L} \qquad 5\text{-}4$$

Figure 5-16 shows the evolution of the relative error *e* with the travel distance. It can be concluded that the relative error varies non-linearly with the distance and it does not increase monotonically.

Furthermore, the relative error rapidly decrease from infinite to the minimum value in less than three kilometres, underlying a great variability of this quantity for short distances. On the contrary, for longer distances, the curve increases again tending to an oblique asymptote.

The accuracy cannot reach values lower than 20%, meaning that whatever is the value assumed by the travel distance, the accuracy is however relatively low. Therefore, even a quite accurate estimation of circularity leads to a less accurate inference of distance travelled by particles.

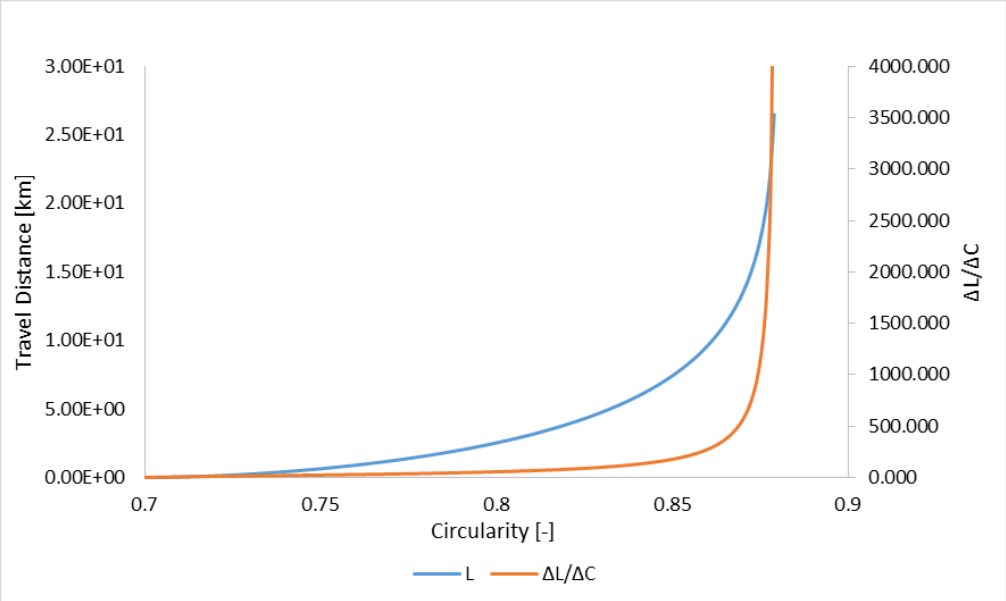

Figure 5-14 Evolutions of travel distance and its first derivative with circularity for k=0.1 [km^{-1}]

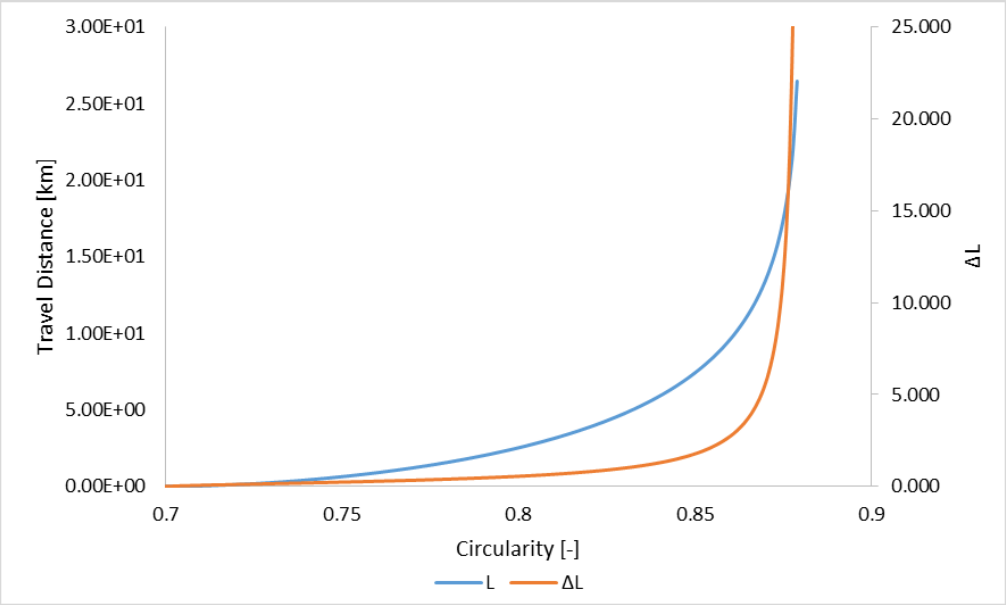

Figure 5-15 Evolutions of travel distance and its increment, given $\Delta C \sim 10^{-2}$ and for k=0.1 [km^{-1}]

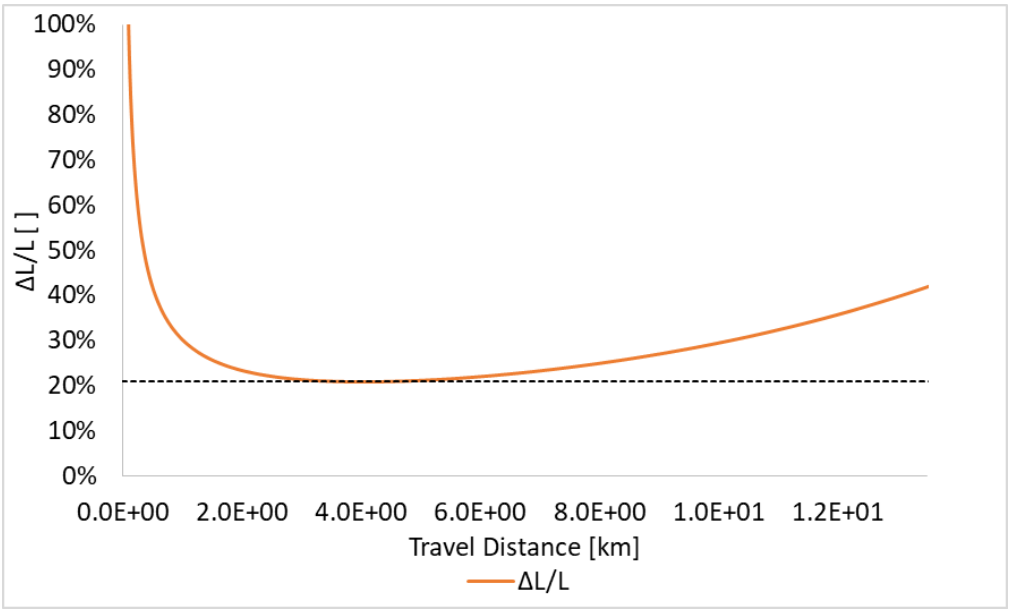

Figure 5-16 Change of the relative error as a function of L

The same analysis has been conducted considering different values of abrasion coefficient. In particular, k has been set equal to 0.1, 0.2, 0.9 and 1.

These values were chosen to assess how much the abrasion coefficient affects the estimated distance and to check whether the same increment of k, applied to both small and large values of k (Δk=0.1 [km^{-1}]), may affect the relative error differently.

Figure 5-17 shows the curves that represent the evolution tracked for each abrasion coefficient. As expected from Equation 4-8, the smaller the value of k, the greater the travel distance for each value of circularity. Furthermore, the influence of the abrasion coefficient on the estimation of distance decreases with the increase of the value assumed by the parameter. Indeed, one can see that the curves relative to the pair of the highest values of k are almost overlapped, while the curves obtained from the couple of the lowest values of k are more separated, even if the increment of k is the same for both the pairs of values.

The analysis of the relative error, expressed as a function of the accuracy of C, reveals that the greater the abrasion coefficient, the greater the variability of the relative error.

Seeing the yellow curve in Figure 5-18, correspondent to the highest value of abrasion coefficient, the error seems to vary faster than the other curves.

This means that a little increment of travel distance leads to a great variation of the accuracy of the inference.

Furthermore, independently on the value assumed by the abrasion coefficient, the relative error cannot decrease below the 20%, underlying that whatever is the abrasion coefficient, the inference of the travel distance results quite inaccurate.

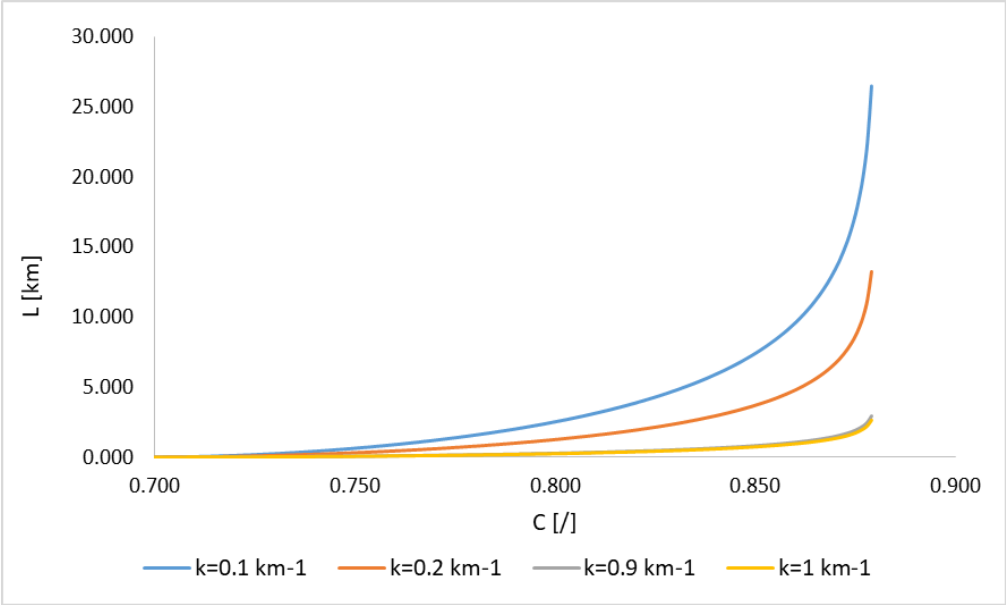

Figure 5-17 Evolution of travel distance with circularity for different values of abrasion coefficient.

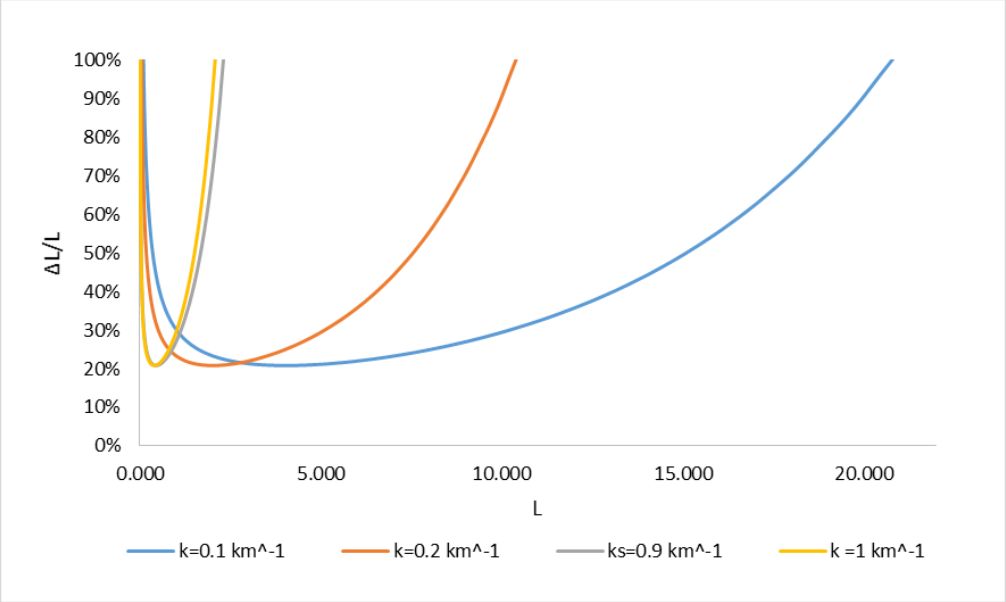

Figure 5-18 Evolution of the relative error with distance for ΔC=0.01 [-]

What has been just said about the influence of *k* on the inference of distance and its relative error, can be confirmed analyzing the effects of *k* when the circularity is fixed at 0.75 [-]. The abrasion coefficient is considered variable between 0.01 and 2 to include all the values of *k* estimated in Section 5.4.

The variability of L with respect of k is evaluated by studying the first derivative, a negative function, as shown in Equation 4-10, which decreases with increasing values of k. The orange curve in Figure 5-19 confirms what already underlined, i.e. the variability of the travel distance decreases with increasing values of the coefficient k (see yellow and gray curves in Figure 5-17), and it tends to be constant. Furthermore, as expected and demonstrated in Figure 5-17, the travel distance decreases with the increase of k, underlying that for a constant value of circularity and, therefore, of relative mass, shorter distances leads to higher abrasion rate.

The non-dimensional variability of the distance travelled, for a given uncertainty in abrasion coefficient estimation, varies linearly with the estimated distance travelled (Figure 5-21).

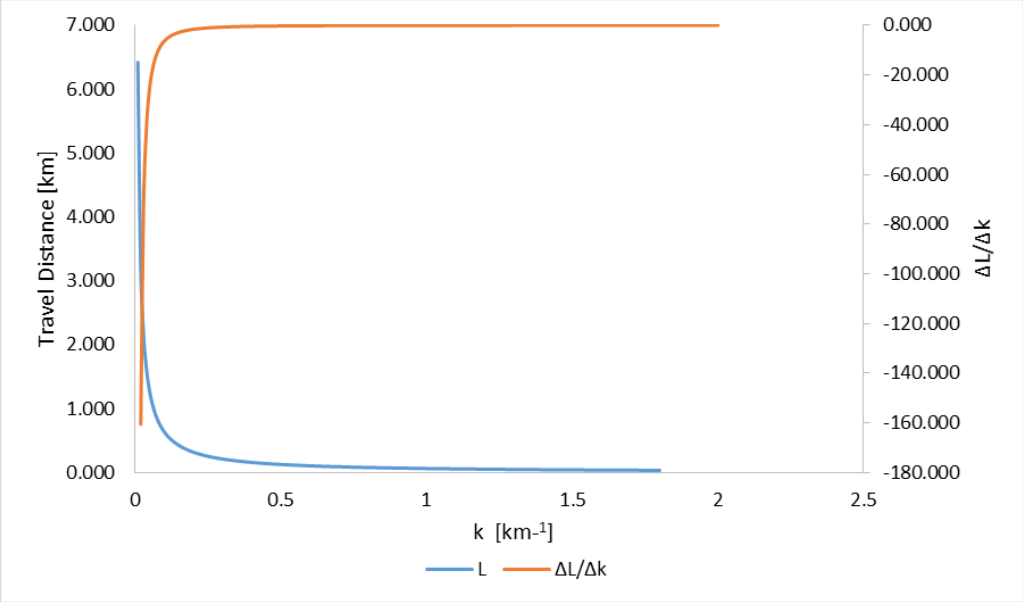

Figure 5-19 Evolution of travel distance and the first derivative for C=0.75 [-]

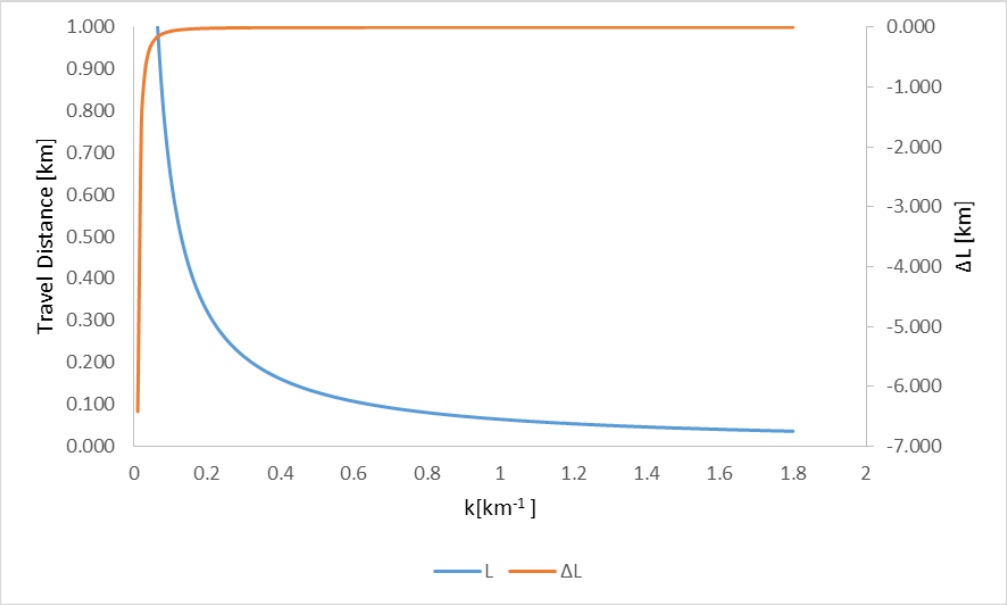

Figure 5-20 Evolution of travel distance L and its increment ΔL for C=0.75 [-] and Δk=0.01

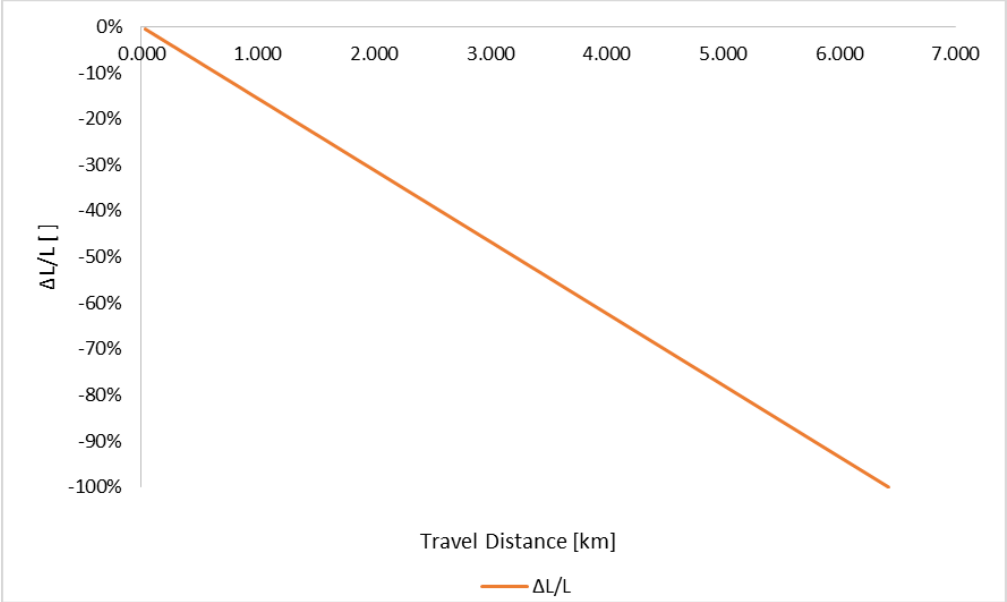

Figure 5-21 Evolution of the relative error with distance for a Δk=0.01 [km-1] and C=0.75 [-]

Considering the mean values of circularity C and the abrasion coefficient k (Table 5-12), estimated for the sections along the Sarzana River, the uncertainty associated with the distances has been estimated as a function of the accuracy of each of the parameters on which the distance depends.

To verify whether the mean circularity and the mean abrasion coefficient were correlated Table 5-12, the covariance has been estimated (Equation 5-5).

Since the covariance is reasonably equal to zero, the two parameters can be assumed as independent.

$$cov(C,k) = \frac{(C - C_{mean})(k - k_{mean})}{n} \qquad 5\text{-}5$$

Location Name	mean C [-]	mean k [km^{-1}]	Covariance [-]
Source	0.760	-	
14	0.858	1.352	
15	0.853	0.872	
17	0.848	0.535	-0.0006
21	0.860	0.234	
22	0.876	0.327	
23	0.851	0.117	
24	0.866	0.159	

Table 5-12 Mean circularity and abrasion coefficient data and the covariance of the two variables

The combined effect of the uncertainties of *C* and *k* on the accuracy of L has been evaluated studying the propagation of the errors. Being the two parameters independent, the combined effects of their accuracies on the distance inference can be assessed applying the formula in Equation 5-6.

$$\Delta L = \sqrt{\left(\frac{dL}{dC}\Delta C\right)^2 + \left(\frac{dL}{dk}\Delta k\right)^2} \qquad 5\text{-}6$$

The sensitivity analisis just shown have been executed considering, for the estimation of mu and therefore of L, the non-shifted interpolating curve μ(C) (Equation 4-4).

The following analyses, since they refer to the values of C and k measured from the field data, refer to the interpolating curve Δμ(C) (Equation 4-5). Therefore, L, $\frac{dL}{dC}$ and $\frac{dL}{dk}$ have been estimated through the Equations 4-8, 4-9 and 4-10 considering Δμ values instead of μ.

The circularity has been found to have a greater influence on the accuracy value in the estimation of L than the abrasion coefficient Table 5-13.

It is not possible to conclude that the greater is the distance the lower is the uncertainty (ΔL/L), unlike what one could expect.

Again, the relative error cannot decrease below the 20% independently on which combination of C and k is considered.

loc_name	C [ ]	k [1/km]	L [km]	ΔL(C) [km]	ΔL(k) [km]	ΔL [km]	ΔL/L [-]
14	0.858	1.352	0.523	0.148	-0.004	0.148	28%
15	0.853	0.872	0.703	0.188	-0.008	0.188	27%
17	0.848	0.535	1.011	0.261	-0.019	0.261	26%
21	0.860	0.234	3.230	0.953	-0.138	0.961	30%
22	0.876	0.327	4.201	2.528	-0.129	2.531	60%
23	0.851	0.117	5.045	1.341	-0.433	1.396	27%
24	0.866	0.159	5.672	1.915	-0.356	1.942	34%

Table 5-13 Absolute and relative error in the estimation of L. The relative contributions of both circularity and abrasion coefficient in the total error ΔL are reported in ΔL (C) and ΔL (k) columns.

Figure 5-22 shows the evolution of the abrasion coefficient with the distance travelled. If one considers the distance of each sampling location, estimated from the source as a function of mean circularity and abrasion coefficient, it is possible to infer the accuracy, represented as the error bars in Figure 5-22. Under the assumption that the variability of L with C, ΔL (C), is the most relevant, one can state that the variability increases with the distance, as confirmed by the orange curve in Figure 5-15 that represents an increasing trend of ΔL with C. This occurs for two effects: the increase in circularity and the decrease in abrasion coefficient, which are respectively directly and inversely related to ΔL (C).

After the fifth kilometer, the absolute error decreases due to the presence of the downstream source that reduces the mean circularity and, therefore, ΔL. Nevertheless, at point 23 of Table 5-13 occurs something different from what just explained.

Comparing ΔL at point 23 with ΔL at point 21 Table 5-13, the value estimated for the first point results higher than the second, even if the circularity, relative to the section 23, is lower. Nevertheless, the reduction of k results to be more relevant than the decrease in C.

This causes an increase of the absolute error associated to the estimated travel distance ΔL, instead of a decrease in ΔL that would be attended when the decrease in circularity is more relevant (Figure 5-15).

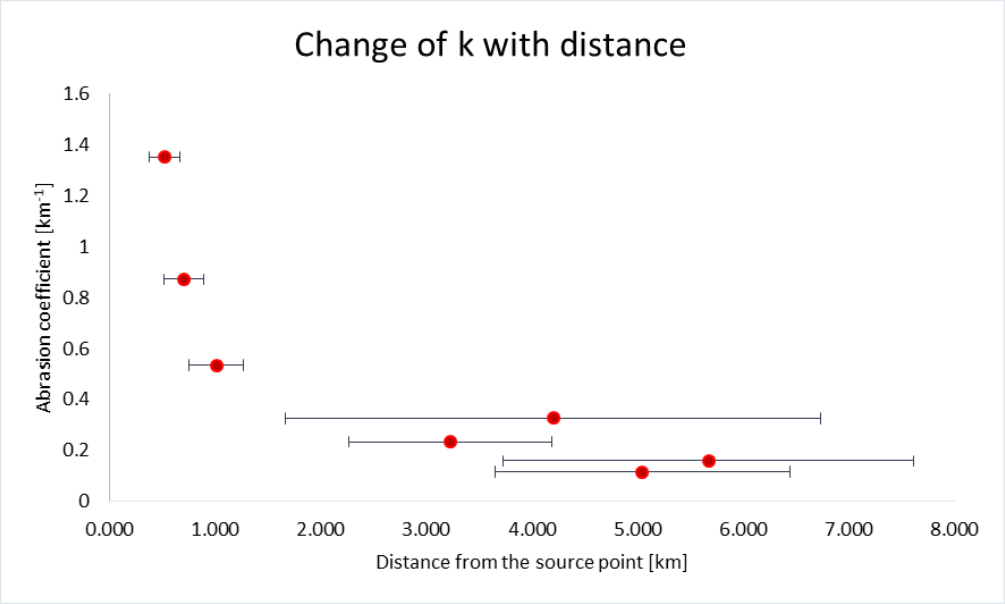

Figure 5-22 Representation of the evolution of the abrasion coefficient. The error bars represent the accuracy of the distance

6. Conclusion

This research aimed to assess whether the morphometric characteristics could be considered as a possible key property for the inference of the distance travelled by particles that can be integrated in the existent connectivity models allowing the calibration.

This work is divided in two phases. The first regards the implementation of a model for automatic edge detection and shape characterization of the particles. The existent computational geometry models cannot guarantee an accurate shape characterization of the particles if a first phase of manual segmentation, through the Adobe Photoshop software, is not applied. For this reason, the implemented model result to be a great improvement in the edge detection, since it allows an accuracy in the estimation of morphometric characteristics of the pebbles reasonably equal to 10^{-2} and it proved to be efficient for around 80% of the sample.

The second phase of this research focuses on the possibility to describe the particles' degradation finding a function that relates the particles' morphometric characteristics and the distance travelled. In particular, recent studies suggest a universal relation between the circularity and the relative mass loss (Novák-Szabó *et al.*, 2018). Since collecting data about mass loss in the field is not feasible, it is necessary to assess whether it is possible to find a relation between the relative mass loss and the distance travelled by particles. Sternberg, in 1875, found that the particles' mass decreases exponentially with the distance travelled. Under these assumptions, the abrasion coefficients that relate the mean circularity values, estimated from the field data, and the measured distances were evaluated.

From the analysis of the data it has been seen that the coefficient of abrasion decreases going downstream. In addition, it has been shown that, as already pointed out in the literature, the abrasion rate of the particles is mainly a function of the size and morphometric characteristics of the stones.

Finally, through a sensitivity analysis, it has been found that the accuracy in the estimation of distance, expressed as a function of circularity and abrasion coefficient is strongly dependent on the accuracy of the measuring method of circularity. Furthermore, the relative error of the measured distance increases non-linearly with both the distance and the circularity when the abrasion coefficient is constant. The increase of the abrasion coefficient, reduces the variation of the distance travelled with respect to the variation of the circularity.

Another effect is the increase of the internal variability of the sample produced by the sequence of two increasing functions applied for relating the morphometric characteristics to the distance travelled. A normalized standard deviation approximately ~ 10^{-2} for circularity sample results in a normalized standard deviation for inferred travel distance ~10^{-1}/1.

It is possible conclude that it is not possible to find a generalized function that relates morphometric characteristics to the travel distance and that can be used for the calibration of the connectivity models, since it depends on the internal conditions of the basin and on the lithology. Furthermore, the application of a sequence of two increasing function revealed an increase in the internal variability of the sample that leads even to physically incorrect estimations of travel distances. Finally, it has been found that the accuracy values of the distance inferred, cannot decrease under 20% considering the accuracy values of circularity and abrasion coefficient reasonably ~10^{-2}.

In the future, the influence of the particles' size on the estimated abrasion coefficient values could be studied by dividing the particles' sample into two size classes (small and large grains) and evaluating the coefficient for each class. Moreover, the study of the influence of the morphological characteristics of the riverbed, that affect mainly the size-selective transport, could be another possible step forward for the research.

The fieldwork experience showed that, unlike what suggested by the lithologic map, most of the metabasalts sources are punctual outcrops of the rock placed within small creeks draining the hillslopes.

To obtain a more accurate distribution of path-lengths, a possible future step forward, in this analysis, could be to consider only the paths that connect the downstream sink points with the upstream punctual outcrops and to associate, to each distance, a connectivity index. This value would represent the probability that a pebble deposited downstream comes from one of the identified outcrops. Nevertheless, this is still a preliminary idea that must should be studied in the future.

7. Acknowledgement

Thanks are due to Professor Francesco Ballio and Professor Mário Franca, who have allowed the realization of this work and have been always of great help and, above all, reference points.

I thank Alessandro Cattapan, for his availability. He supported and helped me many times and he has been of fundamental importance in the realization of this project.

Special thanks are also due to my family, and to my parents in particular, who have always been at my side although not physically. Thank you for having believed in me even when I doubted myself.

I thank my friends from Fano who, despite being almost always far away in these years, have never let me miss their affection for a moment, sending me a picture, a video of their evenings together, or video-calling me.

Thanks are due to my university friends, who have made the endless days of lessons and study more fun, including coffee breaks, lunches in the garden and card games.

I thank my roommates, who have been a second family. We argued, laughed, sang, joked, cried, and supported each other.

Without you this goal would not have been possible, so thank you.

Bibliography

Ancey, C. (2020a) 'Bedload transport: a walk between randomness and determinism. Part 1. The state of the art', *Journal of Hydraulic Research*. Taylor and Francis Ltd., 58(1), pp. 1–17. doi: 10.1080/00221686.2019.1702594.

Ancey, C. (2020b) 'Bedload transport: a walk between randomness and determinism. Part 2. Challenges and prospects', *Journal of Hydraulic Research*. Taylor and Francis Ltd., 58(1), pp. 18–33. doi: 10.1080/00221686.2019.1702595.

Anderson, E. P. *et al.* (2018) 'Fragmentation of Andes-to-Amazon connectivity by hydropower dams', *Science Advances*. doi: 10.1126/sciadv.aao1642.

Bradley, W. C. (1970) 'Effect of weathering on abrasion of granitic gravel, Colorado River (Texas)', *Bulletin of the Geological Society of America*. doi: 10.1130/0016-7606(1970)81[61:EOWOAO]2.0.CO;2.

Brewer, P. A. and Lewin, J. (1993) 'In-transport modification of alluvial sediment: field evidence and laboratory experiments', *Alluvial sedimentation*.

Canny, J. (1986) 'A Computational Approach to Edge Detection', *IEEE Transactions on Pattern Analysis and Machine Intelligence*. doi: 10.1109/TPAMI.1986.4767851.

Cassel, M. *et al.* (2018) 'Evaluating a 2D image-based computerized approach for measuring riverine pebble roundness', *Geomorphology*. Elsevier B.V., 311, pp. 143–157. doi: 10.1016/j.geomorph.2018.03.020.

Czuba, J. A. and Foufoula-Georgiou, E. (2014) 'A network-based framework for identifying potential synchronizations and amplifications of sediment delivery in river basins', *Water Resources Research*. Blackwell Publishing Ltd, 50(5), pp. 3826–3851. doi: 10.1002/2013WR014227.

Czuba, J. A. and Foufoula-Georgiou, E. (2015) 'Dynamic connectivity in a fluvial network for identifying hotspots of geomorphic change', *Water Resources Research*. Blackwell Publishing Ltd, 51(3), pp. 1401–1421. doi: 10.1002/2014WR016139.

Davis, C. M. and Fox, J. F. (2009) 'Sediment Fingerprinting: Review of the Method and Future Improvements for Allocating Nonpoint Source Pollution',

Journal of Environmental Engineering, 135(7), pp. 490–504. doi: 10.1061/(ASCE)0733-9372(2009)135:7(490).

Engel, K. (2006) 'Sobel operator', in *Real-time volume graphics*.

Frings, R. M. (2008) 'Downstream fining in large sand-bed rivers', *Earth-Science Reviews*. doi: 10.1016/j.earscirev.2007.10.001.

Green, B. (2009) 'Canny edge detection', *Retrieved: March*. doi: 10.1109/TPAMI.2005.173.

Gupta, S. and Ghosh Mazumdar, S. (2013) 'Sobel edge detection algorithm', *International Journal of Computer Science and Management Research*.

Guzmán, G. *et al.* (2013) 'Sediment tracers in water erosion studies: Current approaches and challenges', *Journal of Soils and Sediments*, 13(4), pp. 816–833. doi: 10.1007/s11368-013-0659-5.

Kodama, Y. (1994) 'Experimental study of abrasion and its role in producing downstream fining in gravel-bed rivers', *Journal of Sedimentary Research A: Sedimentary Petrology & Processes*. doi: 10.2110/jsr.64.76.

Krumbein, W. C. (1941) 'The Effects of Abrasion on the Size, Shape and Roundness of Rock Fragments', *The Journal of Geology*. doi: 10.1086/624985.

Krynine, P. D. (1960) 'On the antiquity of "sedimentation" and hydrology (with some moral conclusions)', *Bulletin of the Geological Society of America*. doi: 10.1130/0016-7606(1960)71[1721:OTAOSA]2.0.CO;2.

Kuenen, P. H. (1956) 'Experimental Abrasion of Pebbles: 2. Rolling by Current', *The Journal of Geology*. doi: 10.1086/626370.

Kummu, M. *et al.* (2010) 'Basin-wide sediment trapping efficiency of emerging reservoirs along the Mekong', *Geomorphology*. doi: 10.1016/j.geomorph.2010.03.018.

Lewin, J. and Brewer, P. A. (2002) 'Laboratory simulation of clast abrasion', *Earth Surface Processes and Landforms*. doi: 10.1002/esp.306.

Maini, R. and Aggarwal, H. (2009) 'Study and comparison of various image edge detection techniques', *International Journal of Image Processing*.

Manh, N. Van *et al.* (2015) 'Future sediment dynamics in the Mekong Delta floodplains: Impacts of hydropower development, climate change and sea level

rise', *Global and Planetary Change*. doi: 10.1016/j.gloplacha.2015.01.001.

Mikos, M. (1995) 'Fluvial abrasion: converting size reduction coefficients into weight reduction rates', *Journal of Sedimentary Research A: Sedimentary Petrology & Processes*. doi: 10.1306/d42680fe-2b26-11d7-8648000102c1865d.

Miller, K. L. *et al.* (2014) 'Quantifying the significance of abrasion and selective transport for downstream fluvial grain size evolution', *Journal of Geophysical Research F: Earth Surface*. Blackwell Publishing Ltd, 119(11), pp. 2412–2429. doi: 10.1002/2014JF003156.

'Mobile Topographer' (no date). Available at: https://play.google.com/store/apps/details?id=gr.stasta.mobiletopographer&hl=en&gl=US.

Morris, P. H. and Williams, D. J. (1999) 'Worldwide correlations for subaerial aqueous flows with exponential longitudinal profiles', *Earth Surface Processes and Landforms*. doi: 10.1002/(SICI)1096-9837(199909)24:10<867::AID-ESP16>3.0.CO;2-L.

Nations, U. (2030) *70/1. Transforming our world: the 2030 Agenda for Sustainable Development Transforming our world: the 2030 Agenda for Sustainable Development Preamble*.

Novák-Szabó, T. *et al.* (2018) 'Universal characteristics of particle shape evolution by bed-load chipping', *Science Advances*. American Association for the Advancement of Science, 4(3), p. eaao4946. doi: 10.1126/sciadv.aao4946.

P., J. A. (1879) 'Études Synthétiques de Géologie Expérimentale', *Nature*. doi: 10.1038/020501b0.

Pizzuto, J. E. (1995) 'Downstream Fining in a Network of Gravel-Bedded Rivers', *Water Resources Research*. doi: 10.1029/94WR02532.

Rice, S. (1994) 'Towards a model of changes in bed material texture at the drainage basin scale', *Process models and theoretical geomorphology*.

Roussillon, T. *et al.* (2009) 'Automatic computation of pebble roundness using digital imagery and discrete geometry', *Computers and Geosciences*. Elsevier Ltd, 35(10), pp. 1992–2000. doi: 10.1016/j.cageo.2009.01.013.

Salman, N. (2006) 'Image Segmentation Based on Watershed and Edge Detection Techniques', *The International Arab Journal of Information*

Technology.

Schmitt, R. J. P., Bizzi, S. and Castelletti, A. (2016) 'Tracking multiple sediment cascades at the river network scale identifies controls and emerging patterns of sediment connectivity', *Water Resources Research*. Blackwell Publishing Ltd, 52(5), pp. 3941–3965. doi: 10.1002/2015WR018097.

Schumm, S. A. and Stevens, M. A. (1973) 'Abrasion in place: A mechanism for rounding and size reduction of coarse sediments in rivers', *Geology*. doi: 10.1130/0091-7613(1973)1<37:AIPAMF>2.0.CO;2.

Szabó, T. *et al.* (2015) 'Reconstructing the transport history of pebbles on Mars', *Nature Communications*. doi: 10.1038/ncomms9366.

Vauclin, S. *et al.* (2019) 'Effects of river infrastructures on the floodplain sedimentary environment in the Rhône River', *Journal of Soils and Sediments*. doi: 10.1007/s11368-019-02449-6.

'Veneto Region Geoportal https://idt2.regione.veneto.it/' (no date). Available at: https://idt2.regione.veneto.it/.

Wadell, H. (1932) 'Volume, Shape, and Roundness of Rock Particles', *The Journal of Geology*. doi: 10.1086/623964.

Zheng, J. and Hryciw, R. D. (2015) 'Traditional soil particle sphericity, roundness and surface roughness by computational geometry', *Geotechnique*. Thomas Telford Services Ltd, 65(6), pp. 494–506. doi: 10.1680/geot.14.P.192.