UNIL | Université de Lausanne
Faculté des géosciences
et de l'environnement

Master of Science in Environmental Geoscience

Estimating the discharge of large rivers using remotely sensed information

Bungener Rachel

Sous la direction du Prof. Grégoire Mariéthoz
Codirection du Dr. David Sebag

Source :http://www.cefirse.ird.fr

Maîtrise universitaire ès sciences en géosciences de l'environnement | June - 2018

Secrétariat du Master MSc GSE | www.unil.ch/masterenvi

Contents

Abstract

Water is the most important resource on earth as it is necessary to every living system. A lot of populated areas are near rivers and human activities are strongly linked to rivers. River discharge is a key value for water management and water monitoring. However, their values are not known in many parts of the world due to rare, non-existent or proprietary river gauge measurements (Gleason and Smith, 2014). Remotely sensed information offer an important alternative to obtaining river discharges, as they are accessible and provide a better temporal and spatial overview. Different methods exist but they often depend on in-situ measurements and apply on very large rivers. In this study, a new approach is developed. It combines knowledge from three different disciplines. An image analysis is done to estimate the river width. Then, relations from open-channel hydraulics are used with a probabilistic inversion. Probabilistic inversion is a method widely used for example in Geophysics. It has been proved that it is a very efficient way to analyze *highly nonlinear problems with complex a priori information and data with an arbitrary noise distribution* (Mosegaard and Tarantola, 1995). The method is tested on the Nyong River located in Cameroon which is a relatively small river comparing to usual case studies in the field of interest. Width is obtained from WorldView-3 and Landsat 7 images and the method is implemented in Matlab. The results obtained are very promising which prove that probabilistic inversion is a powerful tool to estimate river discharge using remotely sensed information. It should be the object of further research and interesting developments.

Acknowledgements

Comme toutes les personnes remerciées sont francophones, cette partie fait plus de sens dans cette langue.

Tout dabord, je voudrais remercier mon superviseur, le professeur Grégoire Mariéthoz, pour sa confiance, son intérêt, sa patience et son aide dans ce travail. Grâce à lui, j'ai eu l'opportunité de travailler sur un sujet passionnant et découvrir de nouvelles disciplines. Ses indications m'ont permis de régulièrement remettre en question mon raisonnement et le mûrir.

Je souhaiterai ensuite remercier mon co-directeur, Dr. David Sebag, pour son aide et sa connaissance du terrain. Ses informations et ses contacts ont été précieux pour parvenir à aborder le Nyong avec des éléments plus complets. Je tiens également à remercier les services du BVET dont j'ai utilisé les données pour ce travail.

Enfin, merci à Lionel Benoit qui m'a aidé à mieux comprendre ce mystère qu'est le Metropolis-Hastings. Merci pour ta patience, ton temps et ton enthousiasme!

Ce travail clôture non seulement mon master mais également dix ans de vie Lausannoise bien remplie! Durant ces années incroyables, jai eu la chance de rencontrer des gens merveilleux qui m'ont accompagné et je finis ainsi ces études avec le sourire et une curiosité intacte. Maira, Remy, Maité, Leila, Leandra, Sam, Yohan, pour avoir été un soutien infaillible, une épaule, une deuxième famille quoiqu'il arrive, pour m'avoir supporté dans tous les sens du terme, pour les rires et les pleurs, les angoisses et les bonheurs! Philipp, Stéphane, Olivier pour les longues pauses déjeuners, les passages éclairs programmés deux mois à l'avance, les discussions, débats, les doodles et les fous rires. Jérôme et Johann pour les mardis soirs, les lundis soirs et même des fois les samedis matins ou les dimanches soirs, bref pour être les compagnons de semaines pleines de rires et d'aventures. Blaise, Daniel et Gauthier pour l'accueil plus que chaleureux dans vos locaux et sur les terrasses! Les sudistes, Nour, Tressy et Robin qui m'ont permis d'avoir des petites bulles d'air pleines de chaleur. La classe de master, petite mais soudée, prête à traverser cette étape academique avec enthousiasme et curiosité! Les actuaires en herbes, Etienne, Vladimir, Guillaume, Alexis et Romain qui m'ont accompagné dans la dernière ligne droite, je peux ainsi quitter la vie étudiante fraîche et sereine!

Pour me permettre d'assumer cette étape, j'ai eu la chance extraordinaire de tomber sur des employeurs qui m'ont toujours soutenu dans mes démarches et donné une flexibilité rare afin de s'adapter à mon rythme et mes études. D'abord les Alumni, collègues de la première heure, qui ont suivi tous les rebondissements de ce parcours de près. Annelies, merci pour ta confiance, tes conseils, ton soutien et bien plus encore! Emmanuelle pour tous ces bons moments partagés et tes encouragements! Fiorella pour les fous rires et les vignerons, Arnaud pour les parties de jeux les dimanches, Emilie, François, Leila pour les rires, votre confiance et le coaching pour la suite, je suis un peu plus armée grâce à vous!

Claude, qui m'a toujours fait confiance et avec qui j'ai découvert des domaines auxquels je n'aurai pas eu accès autrement. Et pour toutes les discussions sur la vie, j'ai beaucoup appris en ta compagnie.

Et les bars, qui m'ont permis de changer d'air, de me sentir à la maison à la cité et de faire des tas de rencontres. Hesat, grâce à ta confiance je me suis formée, merci de m'avoir donné une chance. Mehdi et Tomaso pour avoir partagé un bout de votre si belle aventure avec moi! Antoine, Raph, Virginie, Jessie, Bobby, Cassandre, Douniya, Gaelle et Joanna, avec vous les nuits sont plus folles !

Pour finir, loin des yeux mais dans le coeur, Ana, Jean, Marc et Zoé, vous êtes mon ADN et il n'y a pas de mots pour décrire tous les sentiments que j'ai pour vous, merci, je n'en serais pas là aujourd'hui sans vous.

Chapter 1 Introduction

Water is the most important resource on earth as it is necessary to every living system. Rivers are one source of freshwater. A lot of populated areas are near rivers and human activities are strongly linked to rivers.

One of the most important hydraulic observation is the temporal and spatial variation in **river discharge** (Alsdorf et al., 2003). Indeed, river discharge is a key value in numerous application.
First, it could be used to study the watershed function. This aspect is necessary for example when predicting inundations patterns. Discharge values are also often used in numerical models to estimate extreme values. It is also useful for fisheries and wildlife habitat monitoring (Smith, 1997). In addition, knowing how the watershed works is a useful tool to study how a pollutant could be transported inside a watershed.
Second, river discharge is used to implement and monitor various human activities such as irrigation, navigation or hydro-energy production. It is a useful value for evaluating environmental strategies (USGS, 2017) or development and rehabilitation project around rivers. It is also a good indicator when studying the impact of urbanization on a river.
Finally, on a larger scale, river discharge is a useful data that could help evaluate climate variations (Cazenave et al., 2016), (Negrel et al., 2011).

Nowadays, river discharge is mainly measured with in-situ gauging stations or with estimation using rainfall-runoff models. But those methods are not optimal. Indeed for the rainfall-runoff models is a simplification on how the watershed and its different compartments work and are linked. Moreover, a lot of observation stations have been interrupted or discontinued (Sichangi et al., 2016), (Alsdorf et al., 2003), (Cazenave et al., 2016), (Pan and Nichols, 2013). **Remotely sensed information** constitute therefore an important alternative to obtain river discharges. It offers many advantages.
First, they offer a better temporal and spatial overview.It provides information over large areas (Bjerklie et al., 2005), with data that are contiguous and political boundary-free (Vörösmarty et al., 2005). Data are over an area and not single points and they are standardized. This is particularly useful for example for time series analysis. Information is measured since the 1990', so it could be used to study temporal evolution. Moreover, there are more than one information on an image, water as well as urban or agricultural objects could be studied from a single image.
Second, it offers a greater accessibility. On one hand, for the financial aspect, as im-

ages are free or can be bought individually for a particular time and place (compare to install and manage many gauging stations). On the other hand, for the physical aspect, as sometimes for geographic or politic reasons, it is difficult to access an area of interest.
Finally, the number of tools to obtain and analyze this kind of information increase each year. Satellite and acquisition devices like radar are more and more powerful and precise. Software and platforms like Google Earth Engine or Matlab's and Python's specific toolbox offer a lot of options to process the data as well as sophisticated algorithm (Tang et al., 2009). In the future years, it should easier to produce useful analytics from this source of information.

Estimating hydraulic parameters and river discharge using remotely sensed information is a subject of interest for a long time.
In 1997, a review of existing methods is presented by Laurence Smith (Smith, 1997). It explains how remotely sensed information is useful for the estimation of inundated areas and flood extent. It also presents some studies that use altimetry to estimate river stage and discharge. In 2005, a study (Bjerklie et al., 2005) try to estimate discharge using exclusively remotely sensed data. In 2007, swath-altimetry is explored to estimate river discharge and depth (Andreadis et al., 2007). A project published in 2012 (Pan and Nichols, 2013), presents a method using the cross-sectional inundation area-river stage relationship (IARSR). In 2011, a paper (Negrel et al., 2011) explores how river flow equations could be modified in order to obtain simpler relations between hydraulic parameters and use only some of them to compute discharge. A very promising method is presented by Gleason and Smith in 2013 (Gleason and Smith, 2014), using at-many-stations hydraulic geometry. Finally, the work published in 2016 (Sichangi et al., 2016) gives good results. In this paper, a state-of-the-art of the field of study is presented in a Table that could be found in Appendix A. They also propose that studies realized can be divided into four categories:

- Methods using altimetry data to obtain water stage and in-situ measured discharge, and then using rating curve to estimate river discharge.
- Methods using satellite to obtain water surface area and in-situ measured discharge, and then using water area-discharge rating curve.
- Methods using hydraulic equations and estimation of their associated parameters by remotely sensed information.
- Methods using width and law of the at-many-stations hydraulic geometry.

In the present master thesis, we have used a mixed approach that combines satellite images analysis, hydraulic equations and probabilistic inversion. Probabilistic inversion is a method widely used for example in Geophysics. It has been proved that it is very efficient to analyze *highly nonlinear problems with complex a priori information and data with an arbitrary noise distribution* (Mosegaard and Tarantola, 1995). So there are two main objectives in this work. First, to create this method and explore its capacity. Second, all these studies presented before have been realized on very large rivers (Amazon, Mississippi, etc), and we would like to explore method on a smaller river which is why the Nyong river in Cameroon is the study case.

The method explored in this project is described in Figure 1.1. It regroups techniques from three disciplines. All aspects will be explained in details in next chapters. First (1), an image analysis (Chapter 3), is done in order to obtain widths along the river. The analysis was done using Matlab and Google Earth Engine on two kinds of images: WorldView-3 (high resolution) and Landsat 7 (lower resolution). Then,(2), knowledge from hydraulics (Chapter4) were used. The link between the width and the discharge is found using Manning's Equation. The definition of hydraulics parameters relatives to our case study constitute the a priori information that will be then used in the inversion. The parameters found were also used to create synthetic data (3) in order to test and assess the quality of the probabilistic inversion. The probabilistic inversion (4), (see Chapter 5), consists of a Metropolis-Hastings algorithm, which allows us to obtain information about a system (discharge) using observational data (width computed at step (1)), prior information (bounded hydraulic parameters specific to the case study) and theoretical relationships (Manning's equation). At the end of the method, we obtain an estimation of the river discharge (5). Details of the case study are presented in Chapter 2.

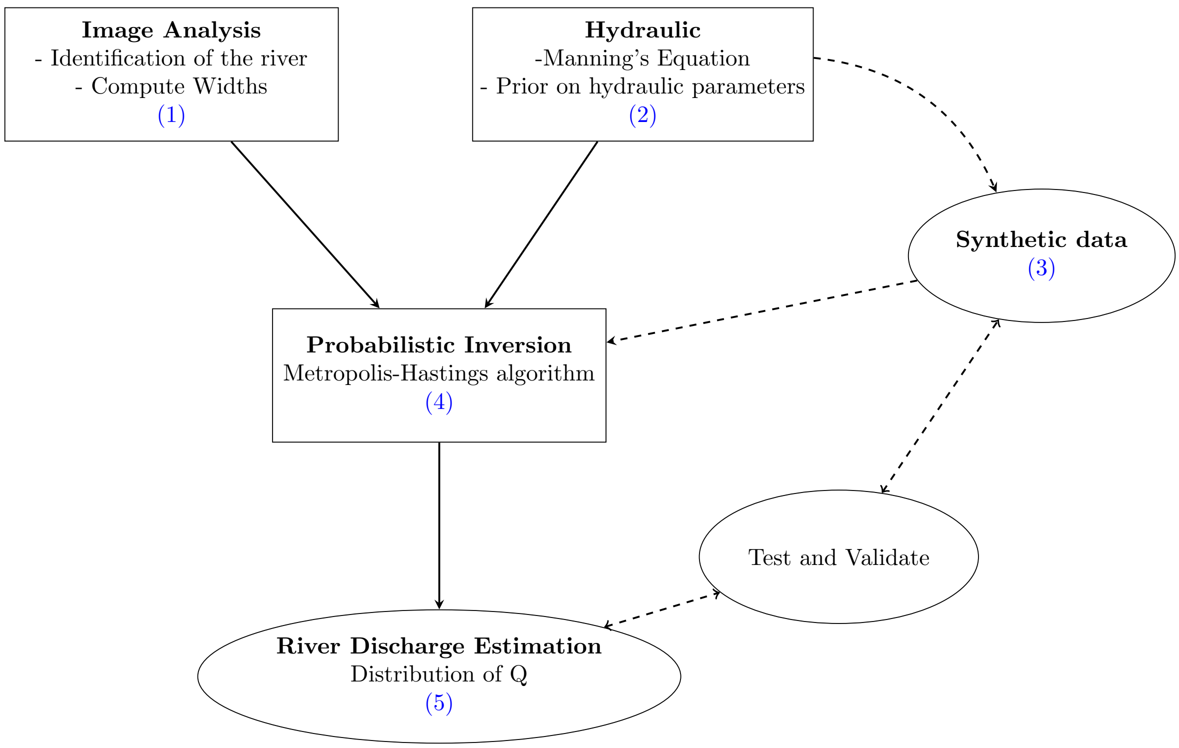


Figure 1.1: Scheme of the Method Explored in the project

Chapter 2 Case Study: The Nyong river

2.1 Generalities

Our case study is on the Nyong river, located in Cameroon, as shown in Figure 2.1.

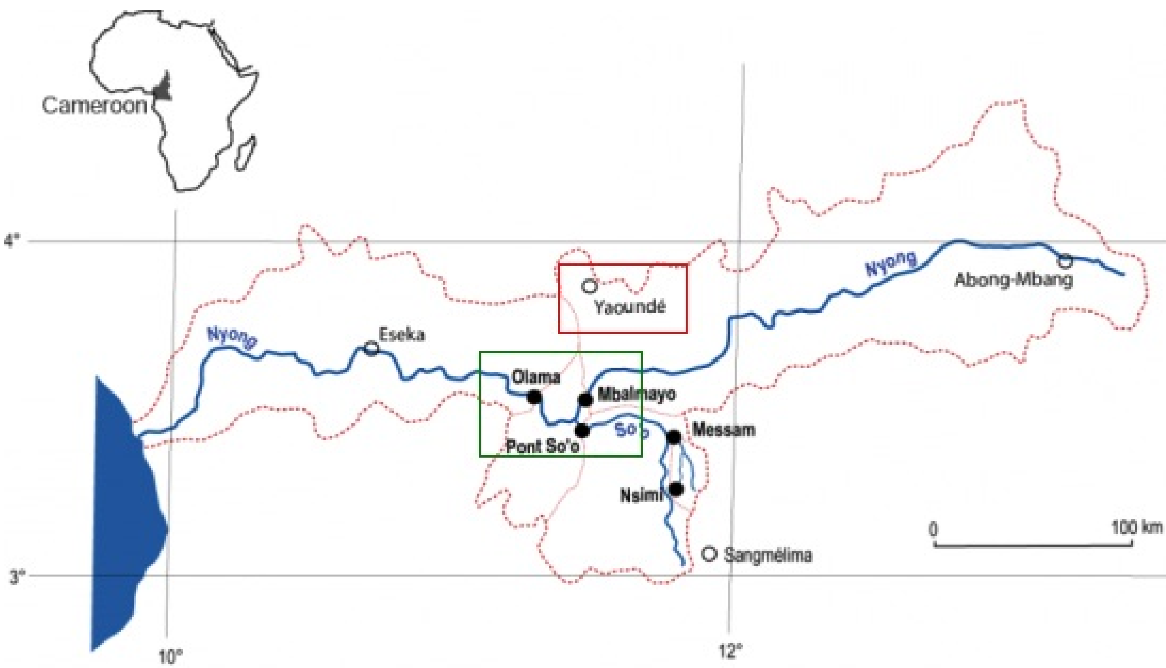


Figure 2.1: Nyong River (http://bvet.obs-mip.fr/fr)

The Nyong is the second river of the country in length (690km) (Olivry, 1986) but it has a catchment area of 27 800 km^2 which is relatively small. It originates at the east of Abong Mbang in the rainforest. It could be divided into three different morphological regions (Viers et al., 2000). First, a swamp zone of 2-3km width (Viers et al., 2000) and a lot of vegetation in the river bed. Around 250km after the Nyong source, the swamp zone decreases and is replaced by a more defined channel with a width around 100m. Then, approximately 230km further, the Nyong River leaves the South Cameroonian Plateau (losing 470m altitude) what leads to a succession of rapids and waterfalls. Finally, it reaches the coastal plain and its mouth is in Petit Batanga in the Guinea Gulf.

The Nyong river presents various advantages regarding our study.

First, a lot of studies, as explained in Introduction, are focused on very large rivers. It is interesting to investigate the method on smaller rivers. Indeed, large rivers are often monitored with on-site stations for many years. Smaller rivers represent a great challenge and are consistent with the aim of our study. The method will provide locals with an easy tool that give an estimation of discharge, useful for example in water management. The community concerned is often smaller than the one on large rivers, as are their financial resources.

Second, the Nyong river is monitored since the fifties. This is necessary because, as said in the introduction, the method created has to be consistent with real settings and we need information to define the prior in the inversion part.

Since the beginning of the fifties, hydrologists from ORSTOM (today IRD - Institut de Recherche pour le Développement) have monitored the river in 8 stations (Olivry, 1986). The first set of data are from *Monographie du Nyong et des fleuves cotiers*(Olivry, 1979) where discharge and water height from 1947 to 1964 are available.

In 1993, the OS-BVET (Observatory for Research in Environment-Experimental Tropical Watersheds) have begun to study the Nyong River basin. This program reunites different French institutions and organizations. *It is a monitoring tool that aims to increase our knowledge regarding the continental water and biogeochemical cycles and the dynamics of weathering processes in tropical environments. It is also dedicated to the study of anthropogenic impacts on the natural environment. These goals are achieved by the combined use of hydrological, geophysical, mineralogical, geochemical methods and modeling*(INSU, 2017). This program has collected a lot of hydrological and geochemical data (INSU 2017). In our case, we use water discharge measurements. All the data used will be explained later in this chapter.

2.2 Area of interest

In order to create and test the methodology, we have to choose an area along the Nyong river. Several elements have to be taken into account. First, the Nyong has to be visible on satellite images and distinct from the rest of environment. Then, it should present variations of width along its path. Moreover, measured data of the area should be available. The last constraint is imposed by Digital Globe, which is a provider of one set of images, and concern the size of the area. It has to be 100 km^2 that could fit in a kind of rectangular shape. The area selected is between Mbalmayo and Olama stations. It is situated in the green rectangle in Figure 2.1 and a zoom is presented in Figure 2.2. It presents variations in width and shape (with meanders). The two measurement stations exist since the fifties and all the data are available. Moreover, the So'o affluent is also visible.

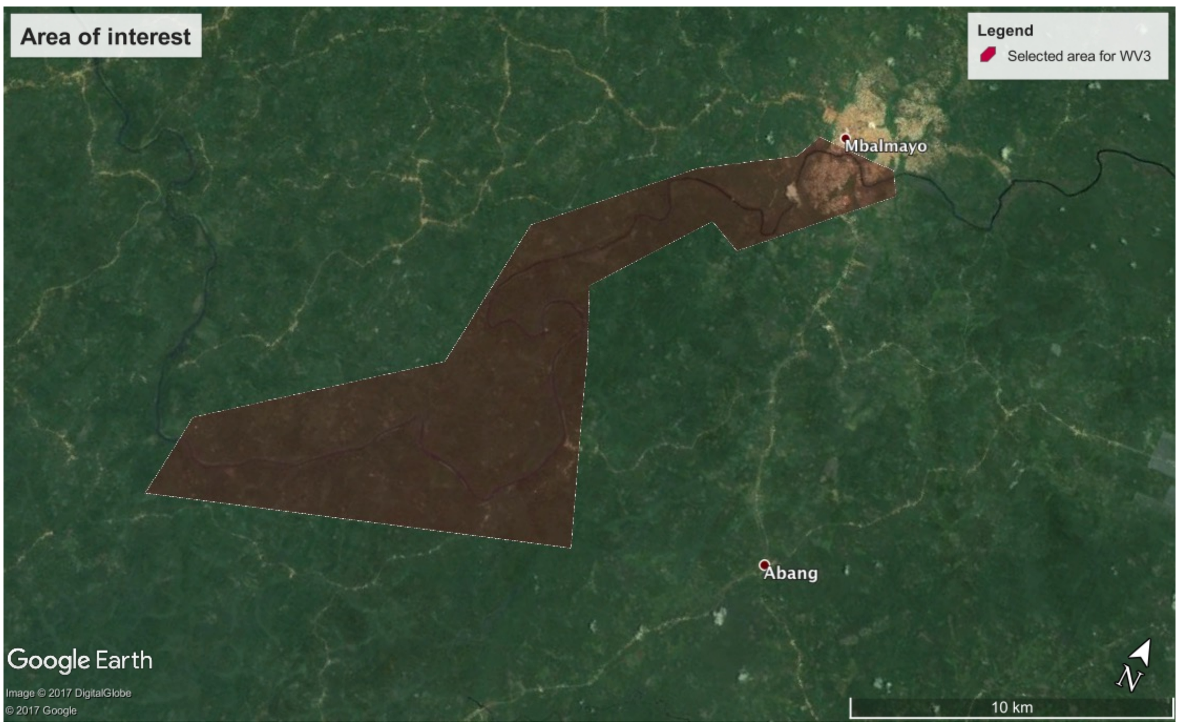


Figure 2.2: Area of interest (Google Earth Pro)

Several parameters that will be useful for the project are known in this area. We have water discharge, slope, a relation between the two stations during the low-water period and stage-discharge curves. For example, on Figure 2.3, the stage-discharge curve at station Olama using measurements between 1964 and 1976 is presented. A lot of other parameters such as geochemical data are also available but we will not use them in our study.

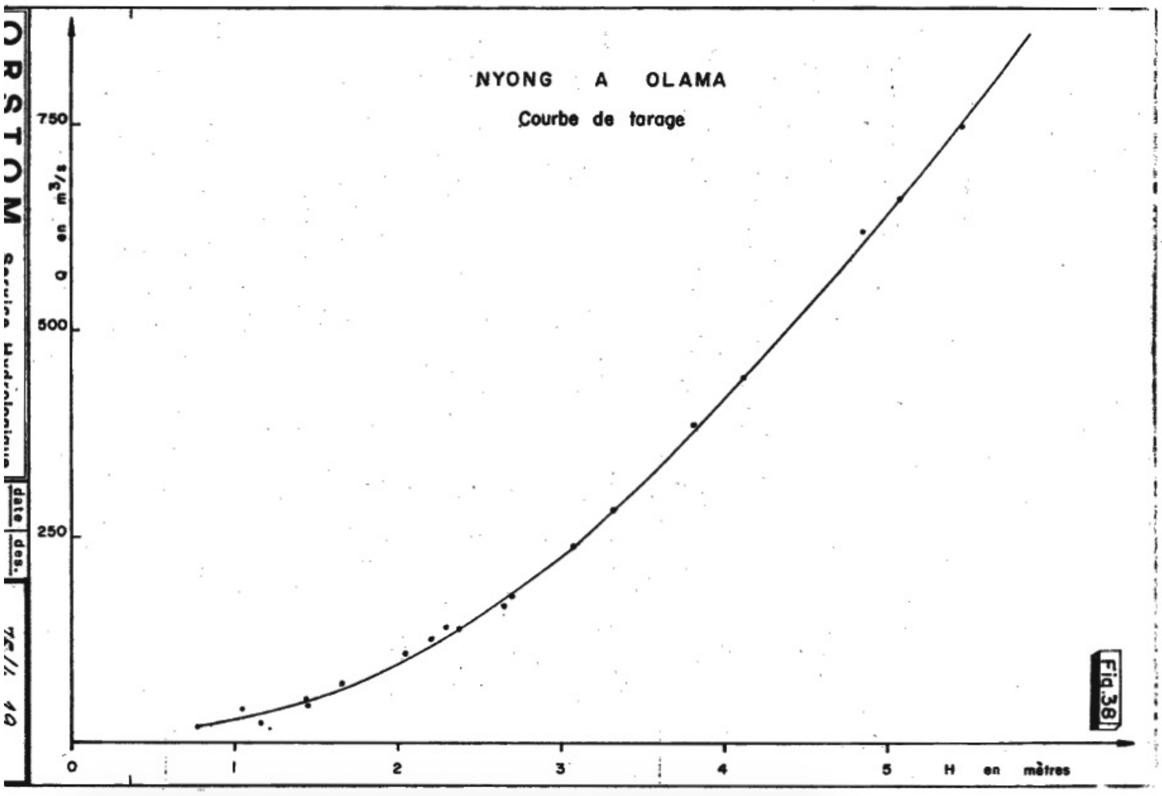


Figure 2.3: Stage-discharge curve in Olama Station (Olivry, 1979)

2.3 Climate and Hydrology

As explained in the introduction, in order to have a good and consistent model, it is important to rely on real data. It will help us to select our images and afterward define the prior.

The climate in the studied area is equatorial with four seasons. It includes two wet seasons, the small one between March and May and the big one between September and November. The big dry season is between December and February, and the small one goes from June to August (Olivry, 1986). Using the data provided by BVET program, the alternation of seasons is visible through variations of river discharge as shown in Figure 2.4 for the Mbalmayo station and in Figure 2.5 for the Olama station.

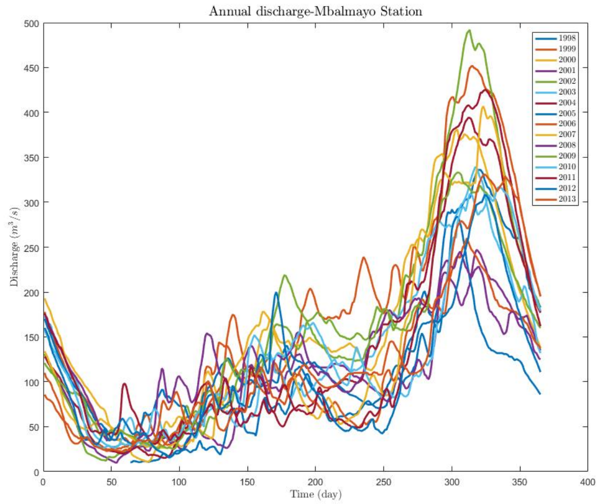


Figure 2.4: Discharge at the Mbalmayo station between 1998 and 2013

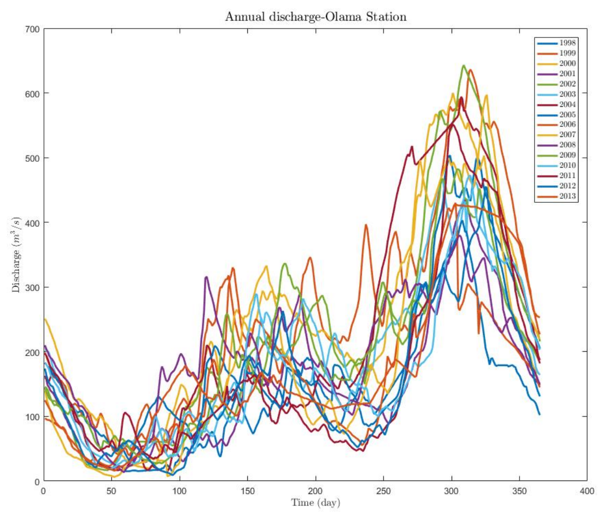


Figure 2.5: Discharge at the Olama station between 1998 and 2013

River discharges at Olama station are generally greater than in Mbalmayo due to the So'o tributary. An example of the relation between the two stations is shown in Figure 2.6. The WorldView set of images was acquired during February 2017. The river discharge measurement for this period are presented in Figure 2.7.

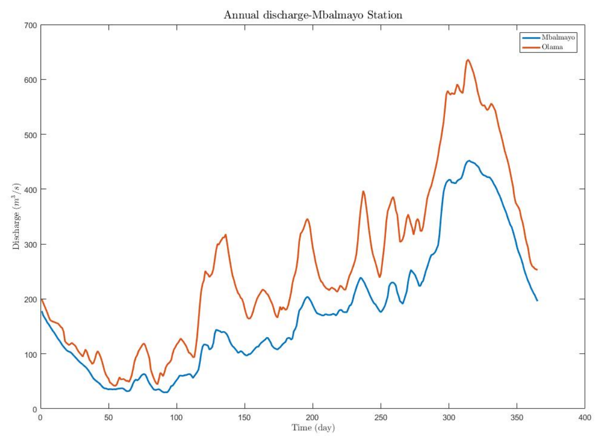


Figure 2.6: Discharge at Mbalmayo and Olama station for year 1999

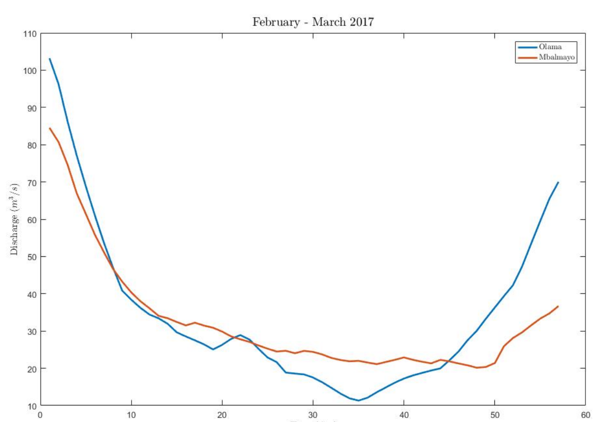


Figure 2.7: Discharge at Mbalmayo and Olama station for 2017 dry season

Chapter 3 Images analysis

In this chapter, the dataset and methodology used to compute the width are detailed. The idea is to obtain a width using images containing a river independently of the satellite used or river's morphology. A satellite image is a collection of bands that contain different spectral information. The two kinds of products used in this project offer Multispectral (MS) imagery and Panchromatic (PAN) imagery. This means that for each image, several spectral bands are available. Depending on the subject of the research, different bands or combination of bands are used. For example, a well-known combination of Near Infrared (NIR) and Red (R) bands, is the NDVI (Normalized Difference Vegetation Index) which highlight the presence of vegetation in images. For our project, we have used the NIR Band which correspond to Band 4 in most products. The specific wavelength range (0.77 - 0.90 micrometers) emphasizes boundaries between vegetation and water (USGS, 2017).

3.1 Dataset

In this project, two kinds of images were used. One is a set of nine images acquired with the satellite WorldView-3 (WV3) and the others are Landsat 7 products found on Google Earth Engine.

3.1.1 WorldView-3

WorldView-3 is a satellite owned by Digital Globe company. It provides panchromatic imagery of 0.31cm resolution and multi-spectral (MS) imagery at 1.24m resolution. The data were captured on the 18th of February 2017, at 09:56. The set of images used is presented in Figure 3.1.

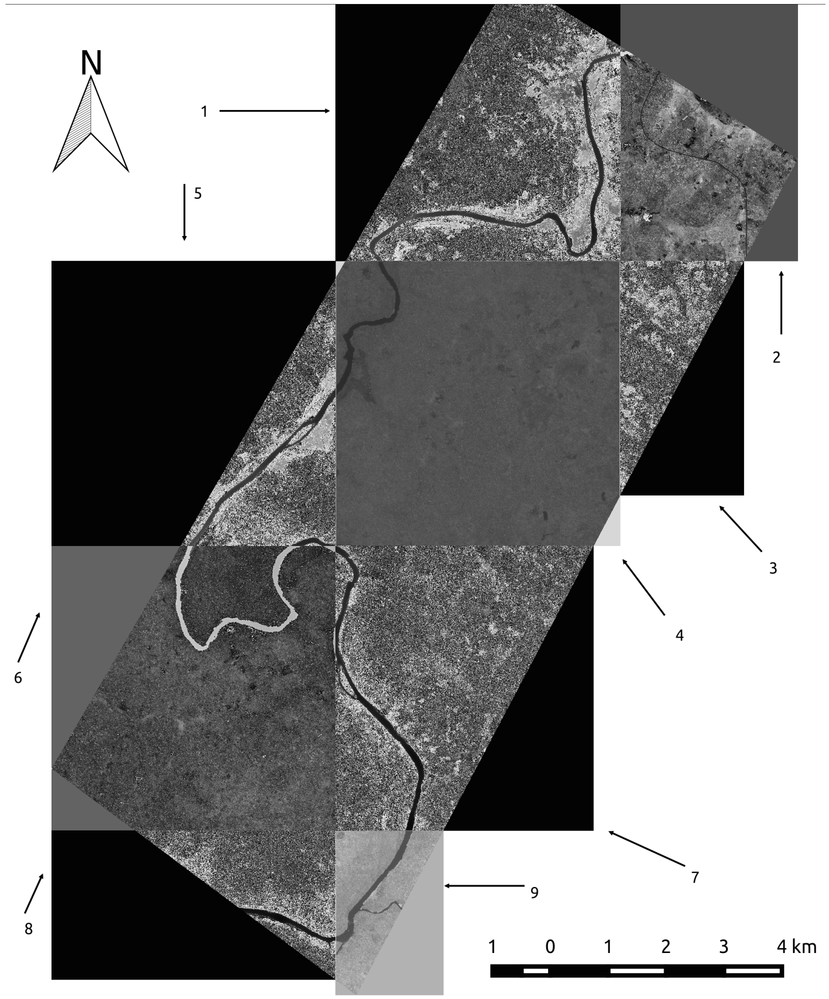


Figure 3.1: Set of WV-3 images, NIR Band

3.1.2 Landsat

Google Earth Engine was used to obtain the Landsat images. It has to be consistent with the WV-3 images in order to compare the results afterward. As the region is very cloudy, it is difficult to find a raw product that works for the analysis. Google Earth Engine provides processing methods that help to obtain a useful image. In this paper, the Landsat TOA (top-of-atmosphere) reflectance percentile composites were used. *This method selects a subset of scenes at each location, converts to TOA reflectance, applies the simple cloud score and takes the median of the least cloudy pixels*(Google, 2017). Pixels closest to the 50th percentile were chosen to minimize clouds and shadows. For a more detailed analysis, the choice of the upper and lower percentiles can be estimated by taking cloud frequency and topographic conditions into account (Donchyts et al., 2016). The technique creates an image composed of

Landsat 7 images that cover a defined period of time. The period of interest, in this project, is the low-water season for two reasons. First, the river bed is better defined and also the WV-3 images were acquired during this season. As the resolution of Landsat images is lower, we wanted to have particularly dry years. The driest years were identified using the data described in Chapter 2. To obtain the TOA percentile composites, three years in a row have been selected and a time range has to be defined. The period goes from 01/02/2000 to 04/10/2003 between Day 30 (January 30th) to Day 100 (April 10th) of each year. We can see on Figure 3.2 that it corresponds to the driest years (lowest peaks) that are in a row. During those days, discharge is below or equal to $75m^3/s$ at Mbalmayo as you can see in Figure 3.3. To explore this issue of low resolution we have also thought of pansharpening but due to time constraints it was only tested on WV-3 images as explained in Appendix E.

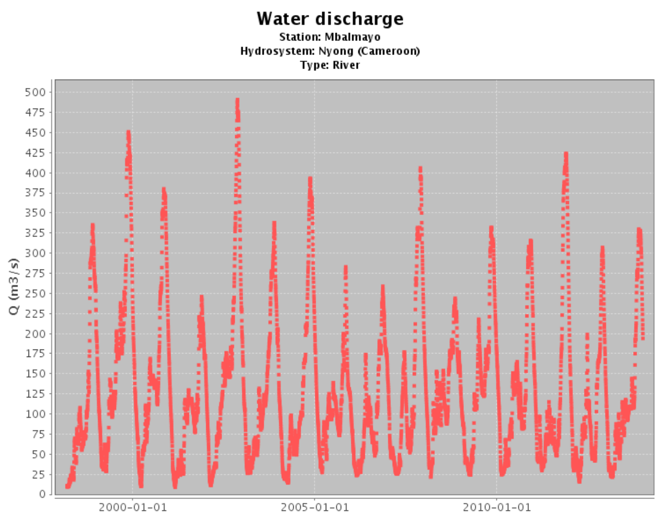


Figure 3.2: Discharge between 1998 and 2013 at the Mbalmyo station (source:bvet)

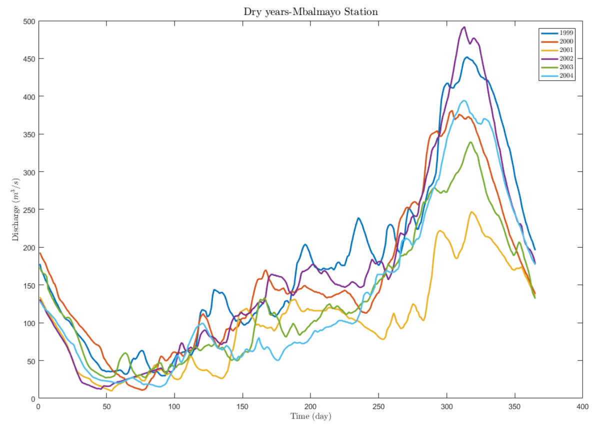


Figure 3.3: Discharge from dry years in Mbalmayo station

Only the NIR band of the composite image was exported and in order to explore various options, two export resolutions were chosen. The first one is 30m which is the real resolution of Landsat 7 images and the second one is 10m. Google Earth Engine use *the scale specified by the output to determinate the appropriate level of the image pyramid to use as input* (Google, 2017). The detail of methods used for smoothing and sub-sampling is not detailed on the API so the distortion induced by those techniques are not known. But as the aim of this work is to estimate river discharge using only satellite images, it is a good lead to follow and see if the difference of resolution induces high variation in the calculated width or not and if those variations impact the estimation of discharge. The images obtained are presented in Figure 3.4.

3.2 Methodology

To compute the width, some steps are required. First, the river has to be isolated from the other component of images using image processing tools and after that, a method has been created to compute the width. All those steps are explained below in details.

3.2.1 Isolation of the river

The river have to be highlighted and isolated from other components (water, trees, lands) on the image. We have used four steps and implemented them in Matlab to achieve this part. First, we have **selected and extracted the NIR band** which

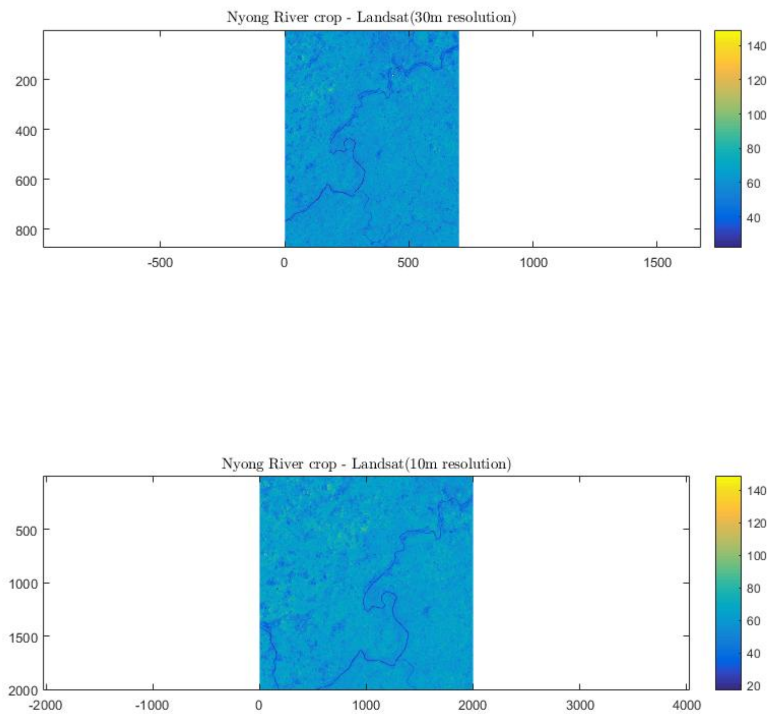


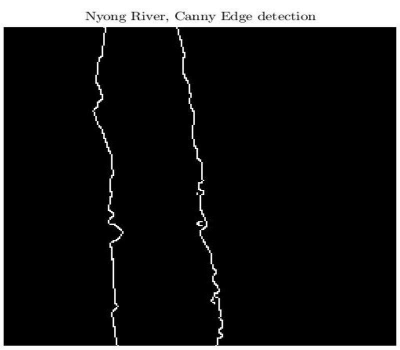
Figure 3.4: Nyong River, Landsat TOA percentile for dry years, NIR Band

is useful for delineating water bodies (Lillesand et al., 2015). For the Landsat 7 product, it was done just by selecting the chosen band on Google Earth Engine. For the WV-3 product, we imported images on Matlab and use the function imread and then select the Band 4. At the end of this step, images are defined by a matrix where each cell is a pixel with a location address and an intensity value. An example in shown in Figure 3.6 (a). For each image, size of the matrix is huge which means high computing time for the rest of the analysis. As the river is not everywhere on the dataset and the aim is not to compute all widths in the case study's area, we have selected only two areas with a cropping. They correspond to parts of Area 1 and Area 8 (identified in Figure 3.1), which are relevant for two reasons. First, they are near gauging stations and also they have structures easily identified that we could use as a landmark to compare WV-3 and Landsat 7 sets.

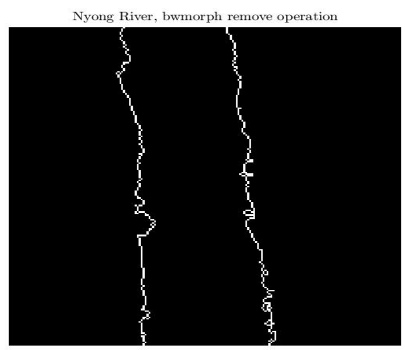
Then, the **water must be selected from other components** of the image. The principle behind this step is that water has a different spectral signature than tree or land, therefore its intensity value is different. We fix a threshold using the visual inspection of the image that takes only the water. For instance, in the Figure 3.4 the river is clearly identified by values below 40 (deep blue). The threshold needs to be found for each kind of image. The visual inspection is sufficient to complete this step. It is an important step because if the threshold is not good, the computing time for the next steps explodes. Exploratory data analysis using data presentation as in Figure 3.4 is a good way to reduce the range of possible values. Moreover, some methods to automatize the threshold selection process exist (Gonzalez et al., 2009) but they were not used in this project due to time constraints as they are not easy to implement. An example of the result obtained is shown in 3.6, (b).

As the region contains a lot of trees and there are wetlands, some residues are still present as in the Figure 3.6 (b). To overcome this issue, a **connected component analysis** is done. This technique identifies objects that are connected and the specify one can be selected (Gonzalez et al., 2009). As the river after the threshold treatment will be the biggest of all water objects, its search can be automatized and the river as a unique object is the output of this step. In Figure 3.6 (c), we clearly see that the river is the only element left in the image.

The last step is the **definition of the river banks**. Our first idea was to use an edge detection analysis. It is an *approach for detecting meaningful discontinuities in intensity values* (Gonzalez et al., 2009). In Matlab, several methods can be used, in this case, the Canny Edge detector was at first chosen. But a simpler operation was finally used, using the function bwmorph, we have removed pixels inside our object. We have chosen this method instead of Canny Edge Detection for two reasons. As shown in Figure 3.5, the Canny Edge detection (a) gives smoother contours than the bwmorph function (b). Moreover, it creates small subparts that could lead to errors when defining the river banks. The output gives the two river banks as shown in Figure 3.6 (d).

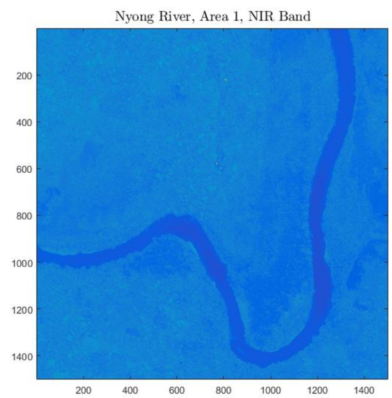


(a)

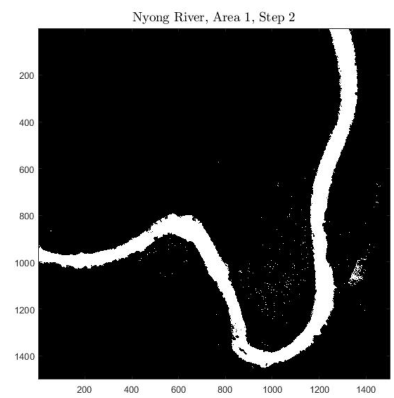


(b)

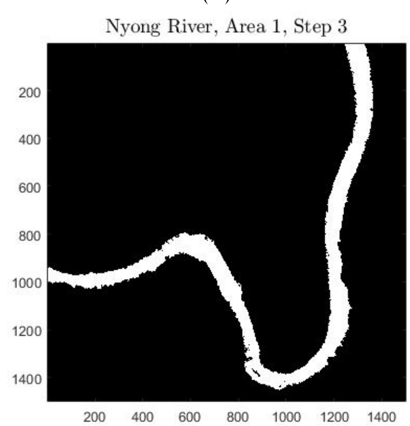
Figure 3.5: Comparison between Canny Edge Detection and Bwmorph operation



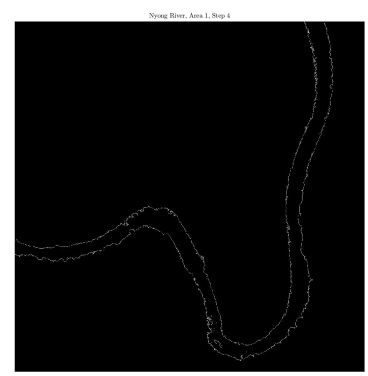
(a)



(b)



(c)



(d)

Figure 3.6: Isolation of the river, Area 1

3.2.2 Computation of the width

Finally, we have to compute the width. As the aim of the method is to work with any kind of river and configuration, some considerations must be taken into account. The orientation of the river must not influence the results. In order to have the real width (perpendicular distance between two banks), the centroid was first identified. The distance of each bank to all other points of the grid is computed and then the centroid is the point where the absolute difference between the two distance matrices is equal to 0. The results are shown in Figure 3.7. To obtain the width, the distance between the bank and the centroid is multiplied by 2 and then by the resolution of the image. A result on a portion of the river is shown in Figure 3.8.

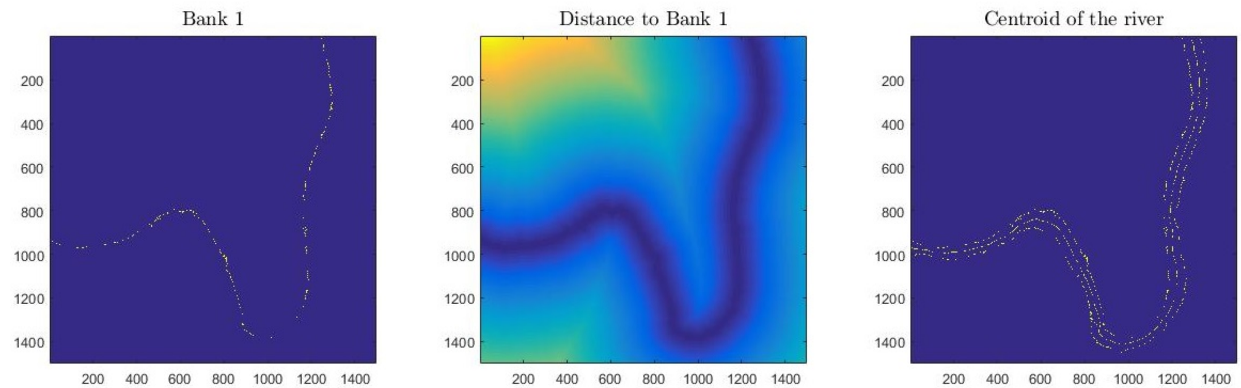


Figure 3.7: Centroid determination

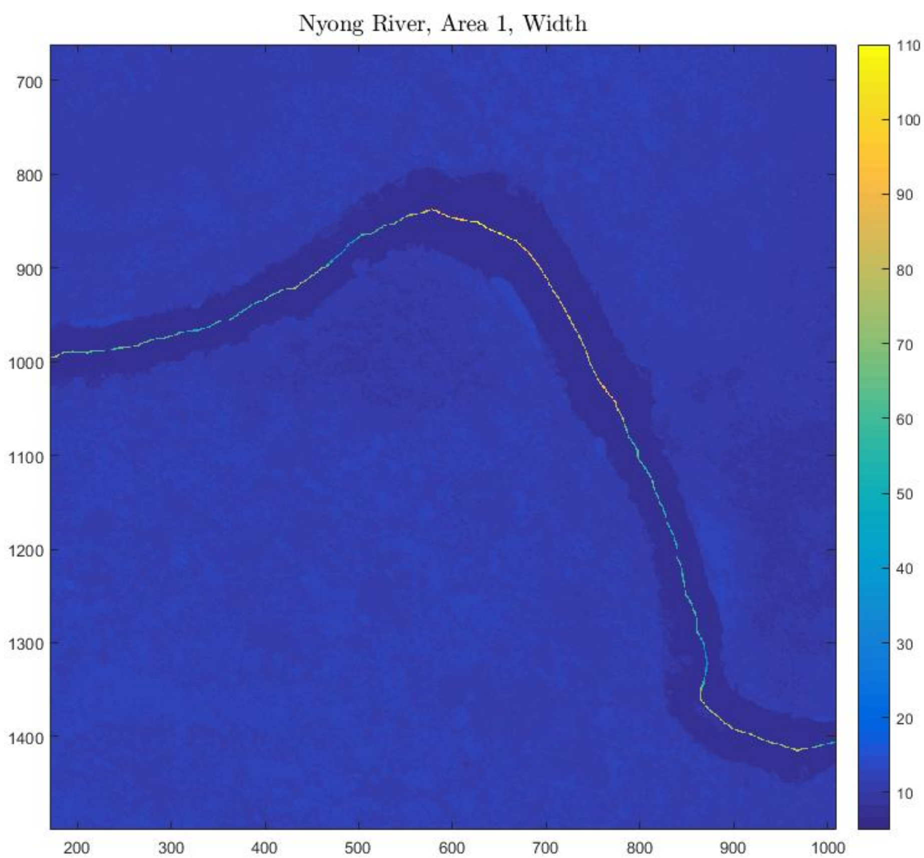


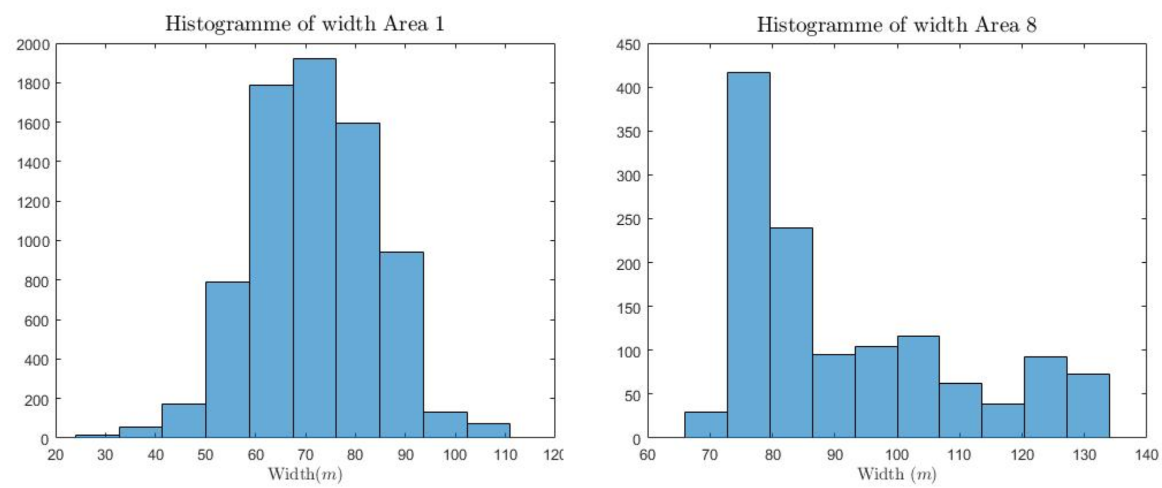
Figure 3.8: Width on World View 3 image

3.3 Results

In this section, we discuss the general results and the difference between images from different resolutions.

3.3.1 General results

As we want to assess the validity of the method, we will consider results only on the WV-3 images in a first step. As shown in Figure 3.8, the width is computed at each point of the river and variations are well represented. The superposition with the real image lets us check if qualitative variation is correct. A measure on ArcGIs have also been made and the width value in five different points is as obtained by the algorithm. Widths computed for Area 1 (a) and Area 8 (b) are presented in Figure 3.9. The range of values for Mbalmayo area goes from 30 to 110m and for Olama area from 65-135m. The distribution and extreme values obtained is consistent with the literature. In (Olivry, 1986), the width is of order 100m. Moreover, the fact that the width at Olama is greater than in Mbalmayo is also coherent as the So'o tributary join the Nyong river just a little upstream.



(a) (b)

Figure 3.9: Histograms of computed width for Area 1 and Area 8

3.3.2 Comparison between WorldView-3 and Landsat images

The same segment of the river has been selected in order to have a coherent comparison. We have chosen a geometric form as a reference. The positive aspect is that the method works, meaning the code for computing the river width run and gives results also for Landsat images. The problem appears with the 30m export of Landsat images. Indeed, the resolution is to low to identify distinct river banks in step 4 and so the centroid cannot be found. In Figures 3.10 and 3.11, we see that the method gives good qualitative results for the 10m export. In the 30m export, we clearly see that there is closing pattern issue. Moreover, even with 10m export, the section of the river must be carefully selected due to a risk of a closing pattern. If in some place, the river is too thin, it closes the river and Matlab interprets it as a single bank. In Figures 3.12 and 3.13, we can see that variations are well identified and width values are almost the same or at least of the same order (also verified in Figure 3.14). The image processing and method to compute the width are working with images with lower resolution but to be efficient for the river discharge estimation, it is necessary to use the 10m export tool from Google Earth Engine.

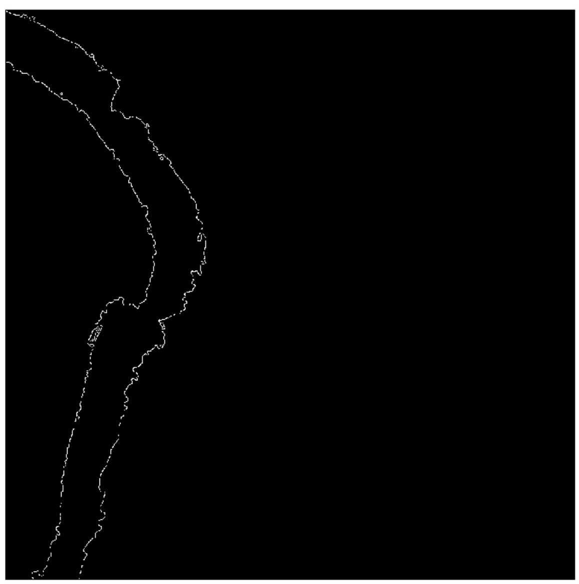


Figure 3.10: Bwmorph operation on WorldView-3 image

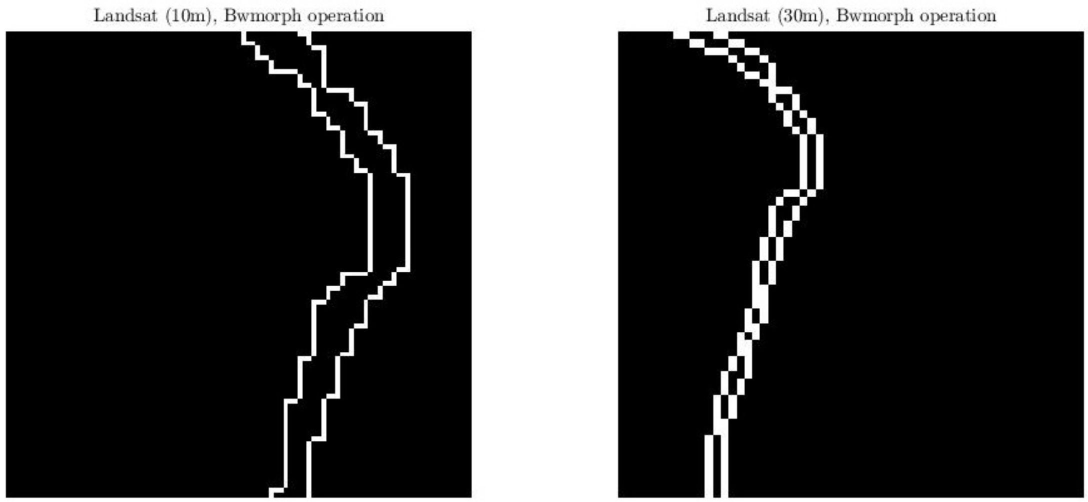


Figure 3.11: Bwmorph operation on Landsat images

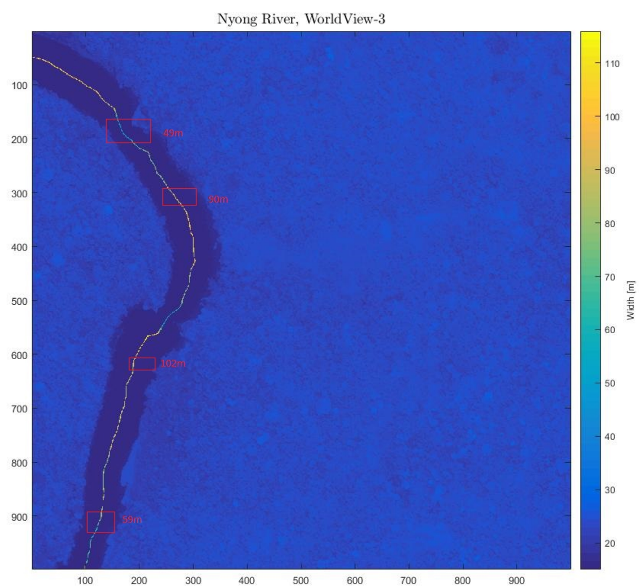


Figure 3.12: Computed widths on WV-3 image

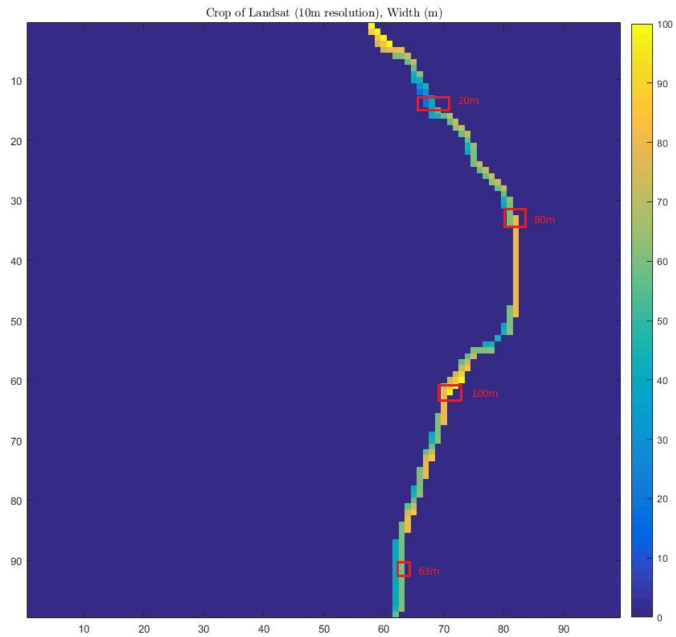
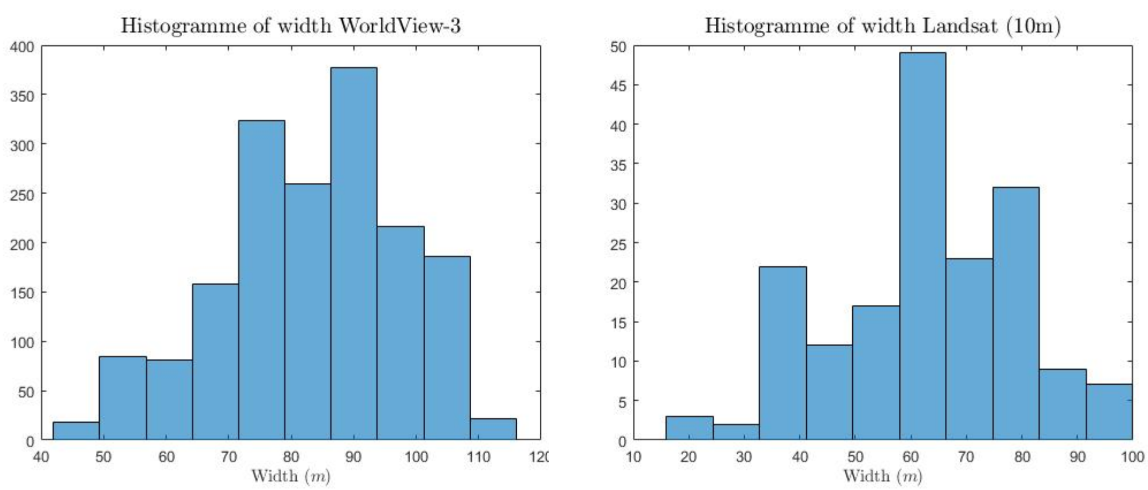


Figure 3.13: Computed widths on Landsat 10m image



(a) (b)

Figure 3.14: Histograms of computed width for WV-3 and Landsat(10m) images

Chapter 4 Hydraulic component

The hydraulic principle used in this project is the Manning Equation (4.1), which is *the most widely used of all uniform-flow formulas for open-channel flow computations* (Chow, 1959). This means that the assumption that the flow is uniform has been made. As we want a relation between discharge and width, a modified Manning's Equation has been used and it will be presented below. Manning's Equation (modified or not) links river discharge to several parameters that will be described in this chapter. Moreover, to perform the Metropolis-Hastings, a good prior is needed. The range of possible values for each parameter will be discussed, and we will present the synthetic river created. It will be used afterwards to assess the quality of results produced by inversion.

$$Q = A * R_h^{2/3} * \sqrt{J} * K \tag{4.1}$$

With Q the discharge (m^3/s), A the section (m^2), $R_h = A/P$ the hydraulic radius (m) with P the wetted perimeter (m), J the slope ($unit$) and K the Manning's roughness.

4.1 Modified Manning's Equation

The width does not appear directly in the Manning's Equation. To do the inversion we need to find a way to express parameters of the equation as a function of the width. We have used the relation 4.2 and it can be then applied in 4.3 and 4.4 to obtain $R_h = A/P$. All parameters of the following equations are explained in Section 4.2.

$$W = B_W + 2 * z * h \tag{4.2}$$

$$A = (B_w + z * h) * h \tag{4.3}$$

$$P = B_w + 2 * h * \sqrt{1 + z^2} \tag{4.4}$$

4.2 Hydraulic parameters

In this project, parameters are not estimated using remotely sensed information except the width. It is due to time issues but a lot of studies exist and are promising and could be used to complete the work done here. In the following section, each parameter will be presented and its values relative in the context of the case study will be discussed. It is an important step of the method because it will define the prior that will be used in the Metropolis-Hastings.

Slope: The channel bottom slope is given by $J = tan\theta$ with θ the angle of inclination (Mays, 2010). It is often given in % or in ‰. For the Nyong river between Mbalmayo and Olama, the slope can be found in the literature (Olivry, 1986) and (INSU, 2017). It is equals to $J = 0.16$‰.

Channel Geometry: The channel geometry is very important as it influences the section and the hydraulic radius. In this project, we assume that variation of width will influence the discharge so we have to define geometry where the width is changing. The main channel geometries in hydraulics are the circle, triangle, rectangle and, trapezoid. In this project, we only consider the trapezoid and the triangular cases like in Figure 4.1. In order to limit the computational cost, we have chosen to vary the angle of the riverbank and the width at the bottom. It let us go from a triangle ($B_w = 0$) to a trapezoid. B_w is set to never be superior to W. The river banks slope goes from $z = 0.5$ to $z = 6$ based on Appendix B of the book of Chow (Chow, 1959). If $z = 0$, the geometry is rectangular.

(a) (b)

Figure 4.1: Trapezoidal (a) and triangular (b) geometries

Width: We used the width computed on our dataset of images. Results have already been presented in Chapter 3. Between the two areas of interest, it gives values between $40m$ and $120m$.

Water height: To define the limits of water height we have used old records presented in (Olivry, 1986) and you can find an example of the data for Olama station in Figure C.1 in Appendix. For the Mbalmayo station, we had a little more data. In the work of Jean-Claude Olivry, heights are related to discharges. The values of discharges are consistent with the one we have for the period of interest, that it is why we assume the value of height can be used. We have set values between $0.04m$ and $2.6m$. These values seem small but they are what is found in the literature for the low-water period.

Manning's Roughness: The Manning's roughness is a key factor in the Manning's equation and there is no way to obtain the direct value. It needs to be determined using different factors that influence it. Those factors are related to material of the bed, irregularity, vegetation, variation of the cross-section, obstructions, and meandering. Chow proposed a method in (Chow, 1959), that allow to estimate a value for each factor and then compute the roughness coefficient n with an equation. The table is presented in Figure C.1. Based on the information we found, we have chosen the following values: $n_0 = 0.025$, $n_1 = 0.020$, $0.005 < n_2 < 0.015$,$n_3 = 0.015$, $0.010 < n_4 < 0.025$, $m_5 = 1.150$ which gives us values of n between 0.035 and 0.1. This is consistent with the literature regarding $K = 1.49/n$ in $m^{1/3}s^1$ (Mays, 2010). Indeed for natural streams with vegetation $K = 10$, for natural streams with meanders $20 < K < 30$ and for uniform section and earth channel $40 < K < 60$. With n between 0.035 and 0.1 it gives K between $14.9m^{1/3}s^1$ and $42.6m^{1/3}s^1$.

All those parameters are useful for the prior in the Metropolis-Hastings. The values used are summed up in the Table 4.1

Parameter	Values range
Slope (J[‰])	0.00016
Water height ($h[m]$)	0.04-2.6
Bank slope (z)	0.5-6
Bottom width ($B_w[m]$)	0-120
Roughness ($K[m^{1/3}s^1]$)	14.9-42.6

Table 4.1: Values used for the prior

4.3 Synthetic river

In order to test and validate the inversion method, synthetic data have been created. It is important because as it is a new methodology we have no idea of its quality. With synthetic data, we know for sure what is the real solution. So we can compare our results to the value we should have obtained and then assess the quality of our methodology.

To generate those data we have used the hydraulic parameters described in the previous section with their chosen boundaries. We have then computed the width using Equation 4.2 and discharge using Manning's equation for every combination possible. To obtain all combinations we have fixed 4 out of 5 parameters and taken all possible values of the fifth. We have done it for every parameter. The parameters are the slope (fixed), the water height, the riverbank slope, the bottom width and the roughness. Their boundaries and their number of values are presented in Table 4.2. In the code, n is converted in K (from 14.9 to 42.57).

The result is a matrix of 211 302 rows with 9 columns (parameters plus W, A, P, R_h and Q). Out of those possibilities, we have removed all inconsistent values, meaning ones where the width was off boundaries (40 and $120m$), as well as the discharge was off boundaries (10 and $140m^3/s$). It gives 93 800 combinations. The main interest to have all these combinations is that it is easy to do subsets depending on what characteristics we need without more computation.

Parameter	Values range	Step between values	Number of values
Slope (J)	0.00016	-	1
Water height (h)	0.04-2.6	0.02	129
Bank slope (z)	0.5-6	0.5/1	9
Bottom width (B_w)	0-120	10	13
Roughness (n)	0.035-0.1	0.005	14

Table 4.2: Values used for synthetic river

Chapter 5 Probabilistic Inversion

In order to compute the discharge, the Metropolis-Hastings algorithm is used. It is a well known probabilistic inversion method. *Inverse problem theory is the mathematical theory describing how information about a parameterized physical system can be derived from observational data, theoretical relationships between model parameters and data, and prior information*(Mosegaard and Tarantola, 1995). In our case, observational data are the computed widths, the theoretical relationship is the Manning's Equation and prior information are the bounded definition of each parameter.

It exists two kinds of inversion methods. Deterministic ones, which usually provides a single estimate, one of many different acceptable inversion solutions. Probabilistic inversion methods, on the other hand, provides (almost) all solutions with their associated statistics. Those methods are very useful for highly non-linear problems (Zahner et al., 2015).

In this section, we will present the Metropolis-Hastings algorithm, then we will discuss its structure and how we have implemented it in the context of the project and finally, we will discuss the obtained results.

5.1 Metropolis-Hastings

Metropolis-Hastings algorithm is a Markov Chain Monte Carlo (MCMC) method. MCMC methods are used for sampling from a probability distribution. A Monte Carlo is a random walk in the model space, which means that the method explores solutions in all the model space avoiding to be trapped in local likelihood maxima (Mosegaard and Tarantola, 1995). The Markov Chain are *sequences of events that are probabilistically related to one another. Each event comes from a set of outcomes, and each outcome determines which outcome occurs next, according to a fixed set of probabilities* (Shaver, 2017). In other words, the Metropolis-Hastings will do a random walk (Monte Carlo) in the model space and to go from a step to the next one, the outcomes are evaluated and added or not to the chain via a Markov chain.

The Figure 5.1 tries to sum up the process. Moreover, to have a concrete view, the structure is detailed in Section 5.2.

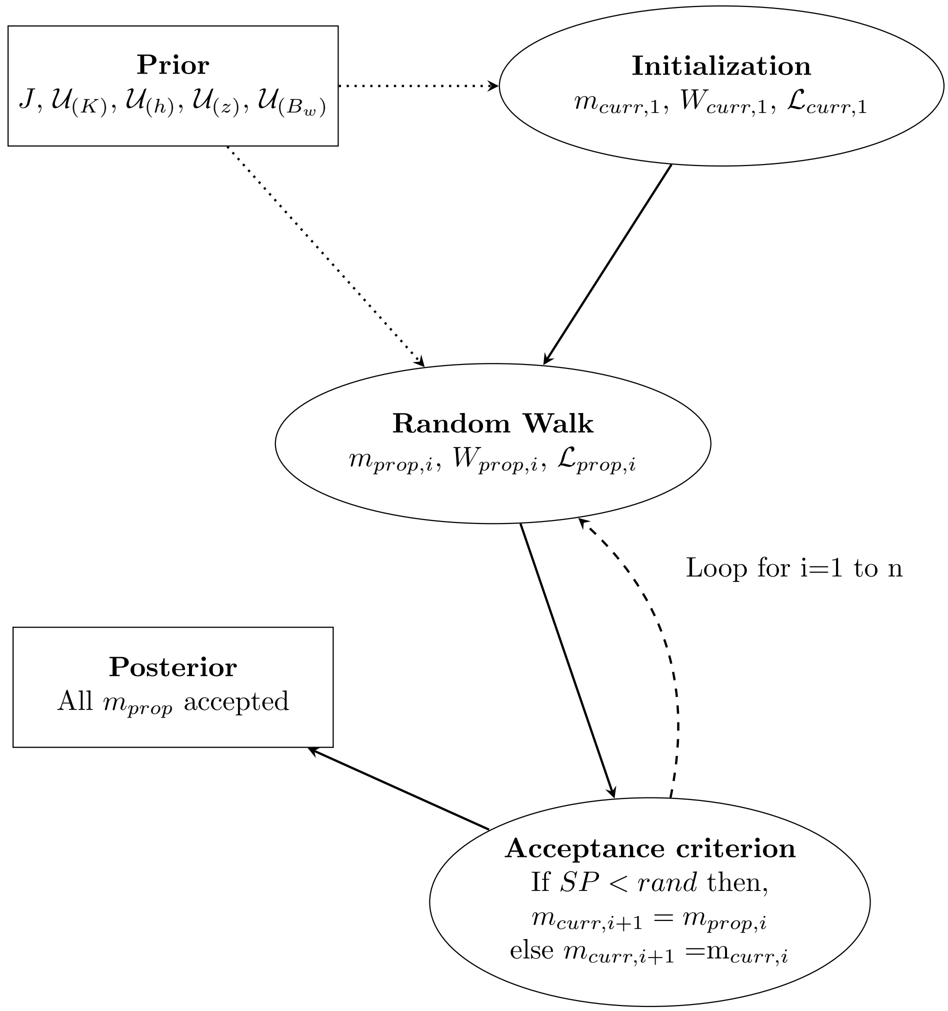


Figure 5.1: Scheme of the Metropolis Hastings Algorithm

5.2 Structure

The method could be seen as a succession of three major steps. First, the initialization of the parameters before entering the loop. It could be seen as the starter point of the walk. Then, entering the loop, the random walk is started, which corresponds to a small step from the initial stage. Finally, the acceptance where a choice is made: if the proposal becomes the current starting point for the next step of the walk or if the current position is kept as a starting point. All accepted proposals are stored and then used to compute the river discharge.

The method has been implemented in Matlab.

Initialization

In this part, the prior is set and the initial current stage is created. First, all parameters, except the slope (J) which is fixed, are defined. A uniform distribution has been chosen with the lower and upper parameters defined in the synthetic case. Then, the first current stage is created in order to initialize the MCMC. A random selection within the distributions is done. So, at this stage J, K_{curr}, Bw_{curr}, z_{curr} and h_{curr} are obtained. The combination of them is called m_{curr}. Using those parameters and the transformed Manning's Equation, we also define W_{curr}, A_{curr}, P_{curr}, Rh_{curr}. With all these elements, the current likelihood $\mathcal{L}_{curr}$ between W_{curr} and W_{obs} is computed using the Equation 5.1. In this project $\sigma = 1.24$ which is the size of a pixel. Different formulas exists for the likelihood, this one was chosen because it corresponds to a Gaussian likelihood.

$$\mathcal{L}_{curr} = exp(-0.5 * (W_{curr} - W_{obs})^2/(\sigma^2)) \tag{5.1}$$

Random walk

When initialization is done, the random walk begins. For each parameter, a new proposal value is defined, it corresponds to $m_{prop} = m_{curr} + \theta * f * rand$. The size of the step between m_{prop} and m_{curr} is defined by $\theta * f * rand$. Where θ is a value that is adapted for each parameter to have a proportional step with regards to its distribution. f is just a scaling factor to easily test the impact of variation in step size. The random number warranties variations between each iteration. The proposed values at each iteration have to be between the boundaries set in the prior and this condition is verified for each loop. The chosen θ values for parameters are presented in Table 5.1.

Parameters	θ value
Roughness	0.005
Bed width	5
Bank slopes	2
Water height	0.01

Table 5.1: Theta values

Then, as in initialization step, W_{prop} is computed and compared to W_{obs} to compute the proposal likelihood $\mathcal{L}_{prop}$ using the Equation 5.1.

Acceptance
The last part consists of an acceptance step. A selection parameter (SP) is set as the minimum between 1 and the ratio $\alpha = \mathcal{L}_{prop}/\mathcal{L}_{curr}$. Then, the proposal is accepted if SP is smaller than a random number between 0 and 1. If the proposal is accepted, in the next iteration $m_{curr} = m_{prop}$, else the same m_{curr} is used and another small random variation from this state is done.

This structure will be applied to particular settings because the method is based under the assumption that there is variation in width but the discharge is constant if the river section studied is not too long. Those conditions are what constrain the inversion. Different widths will be tested and a condition on their related discharge will be defined. An example of three sections is presented in the next section but the idea is to test many more sections at the same time to improve the results.

5.3 Results

In this part, the obtained results will be presented and methodological issues, as well as limits will be discussed in Chapter 6.
In order to give an overview of the results and not have too many figures, results are often presented three by three. So the size of the graphics is often small which is enough to have an overview and analyze them. But to have a more precise view, some are also presented in Appendix G.

5.3.1 Synthetic case

From the synthetic data, in order to explore a case similar to a real setting, we have made a subset that respects some conditions. First, we have selected consistent W values with regards to Mbalmayo station. Then, we wanted a constant discharge around the value measured at the station the day of the image acquisition. Finally, we have chosen a constant K. It gives us 19 combinations.
Among those 19, the three sections presented in Table 5.2 were used for the inversion. The likelihood is computed using the Equation 5.2 with the relations 5.3 and 5.4 .

$$\mathcal{L} = l_1 * l_2 * l_3 * q_1 * q_2 \tag{5.2}$$
$$l_i = exp(-0.5 * (W_i - W_{obs(i)})^2/(\sigma^2)) \tag{5.3}$$
$$q_i = exp(-0.5 * (Q_i - Q_{i+1})^2/(\sigma^2)) \tag{5.4}$$

# index	z	K	B_w	h	W	Q
1	0.5	17.53	50	1.92	51.92	32.2
2	0.5	17.53	60	1.72	61.72	32.2
3	0.5	17.53	70	1.56	71.56	32.2

Table 5.2: Synthetic data used for the test

5.3.1.1 Variation in step size

First, we explored the impact of modifying the step size. We have tried three different values for f: $f = 1$, $f = 5$ and $f = 10$. For each value we have run the inversion three times because it gives different distributions. The number of iterations is set to one million. In Figures 5.2, 5.3 and 5.4, the distribution is more centered on the real value (vertical red line) for $f = 5$.

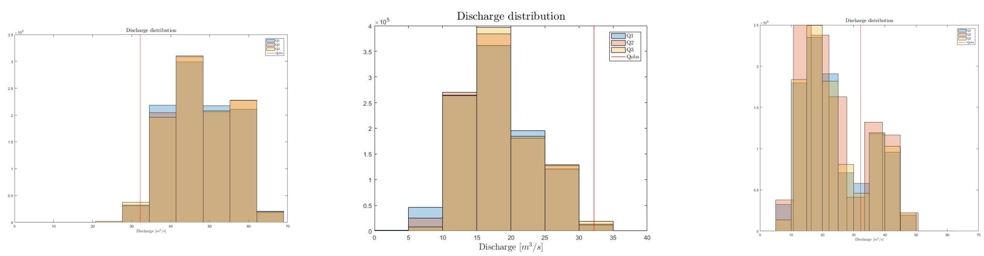


Figure 5.2: Distribution of discharge for $f = 1$

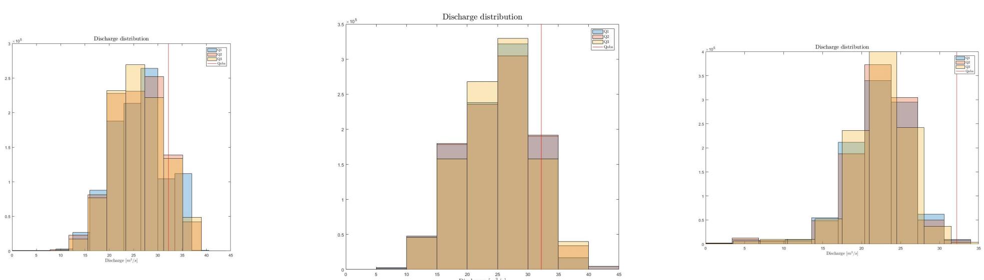


Figure 5.3: Distribution of discharge for $f = 5$

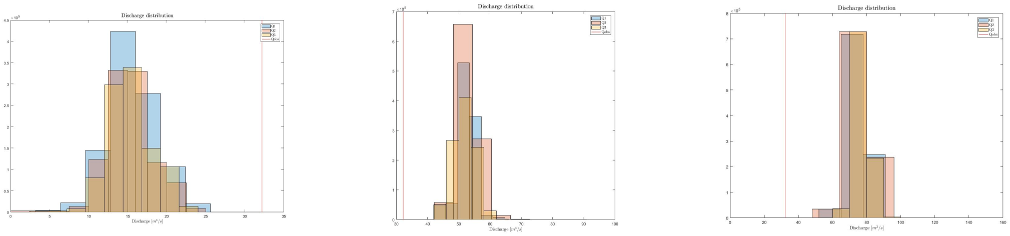
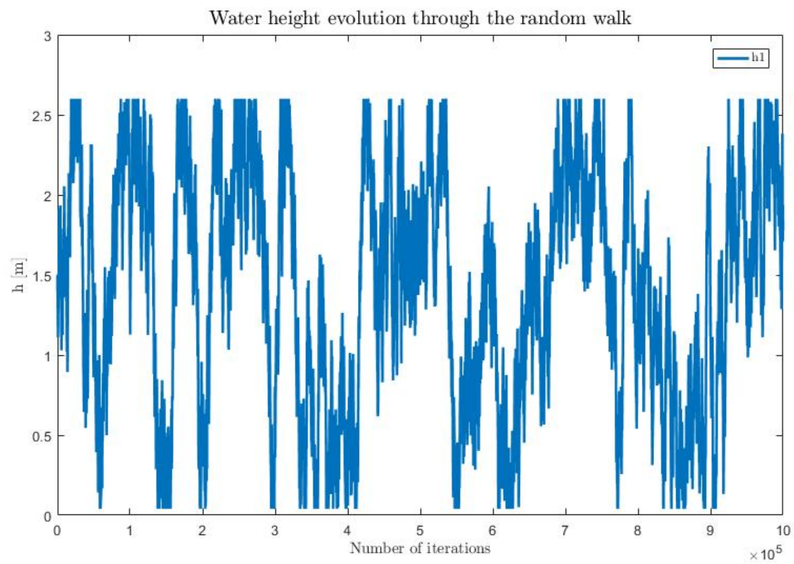
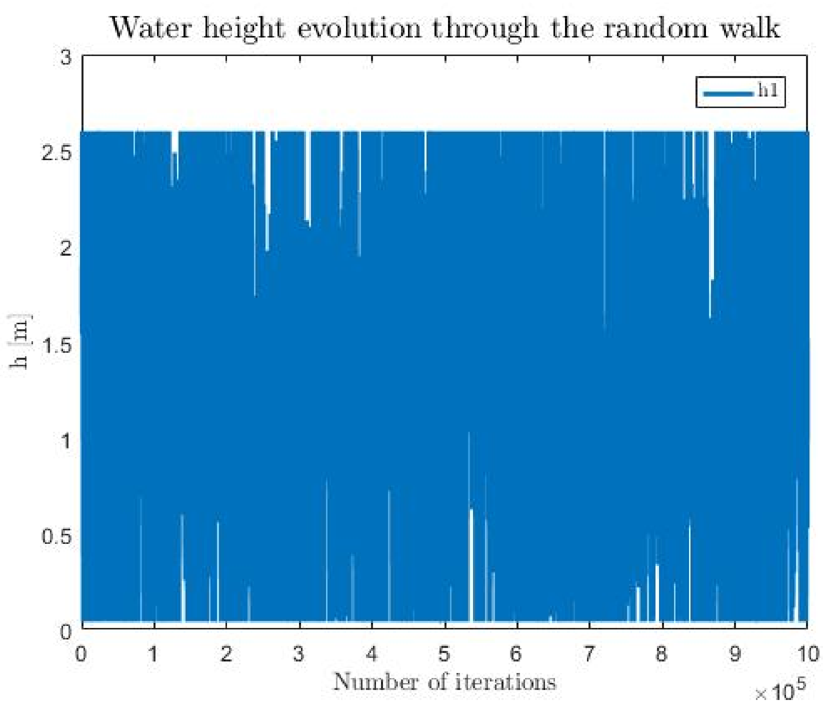


Figure 5.4: Distribution of discharge for $f = 10$

We have also observed how the algorithm explores the prior depending on the step size. For that, we have set $\alpha = 1$, meaning the Markov chain is not conditioned by observation. On Figure 5.5, the impact of the step size is clearly visible. With $f = 5$, almost all the possible values for h are tested whereas, for $f = 1$, the algorithm tries only some values. Those results confirm the choice of $f = 5$. For other parameters, the change does not impact as much the exploration. This is due to the θ values that are different for each parameters (0.01 for h, 5 for B_w).



(a)



(b)

Figure 5.5: Evolution of h_1 through the random walk for $f = 1$ (a) and $f = 5$ (b)

5.3.1.2 Number of iterations

As shown in Figure 5.3, results are different each time we run the code. This means that we need to test more proposals. Two ways are possible, either we multiply the number of Markov chains or we do more iterations. This two methods will lead to similar results as explained in (Tierney, 1994). We have chosen to change from 1 million to 50 million iterations. We see that the discharge distribution tends to be more similar with more iterations. But they are still not perfectly similar. The more we had iterations, the more it is expensive in computing time. Creating multiple Markov chain will help with this issue but due to time constraints, we have not implemented it.

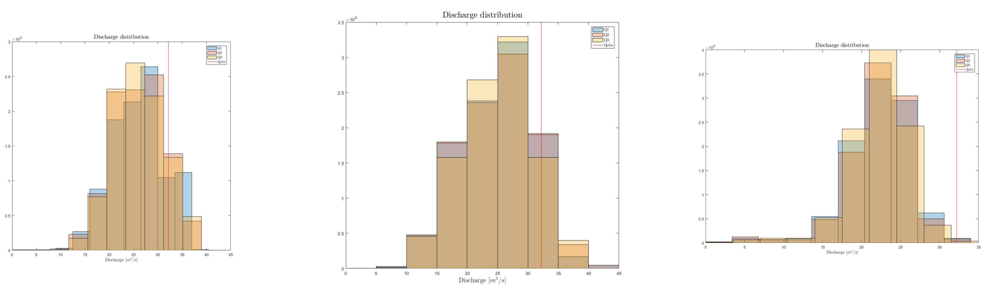


Figure 5.6: Distribution of discharge for $f = 5$,$n = 1 * 10^6$

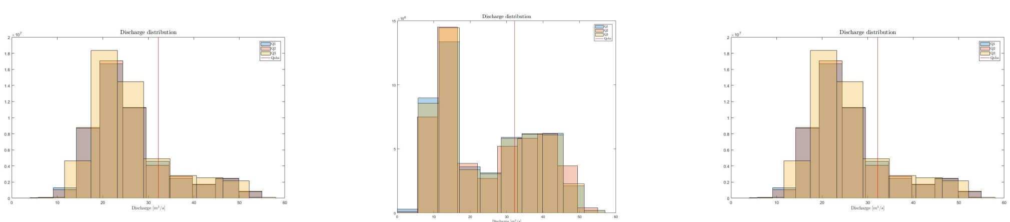


Figure 5.7: Distribution of discharge for $f = 5$, $n = 5 * 10^7$

5.3.1.3 Evolution of the parameters

The results presented in previous sections are satisfying with regards to the discharge. They are often around the real value and they are not really inconsistent values (for example none is above $100m^3/s$). In Figure 5.8, we see that widths tends to their real value. But, in Figure 5.9, we see that the other parameters are not close to their expected value. This means that the overall results are satisfying but do not necessarily take into account consistent hydraulics possibilities and do not correspond the synthetic data. We have also done a bivariate analysis of the parameters. It is presented in Figure 5.10. The bottom width and the water height seems to be dependent variables. For the other variables, it is difficult to infer relationships.

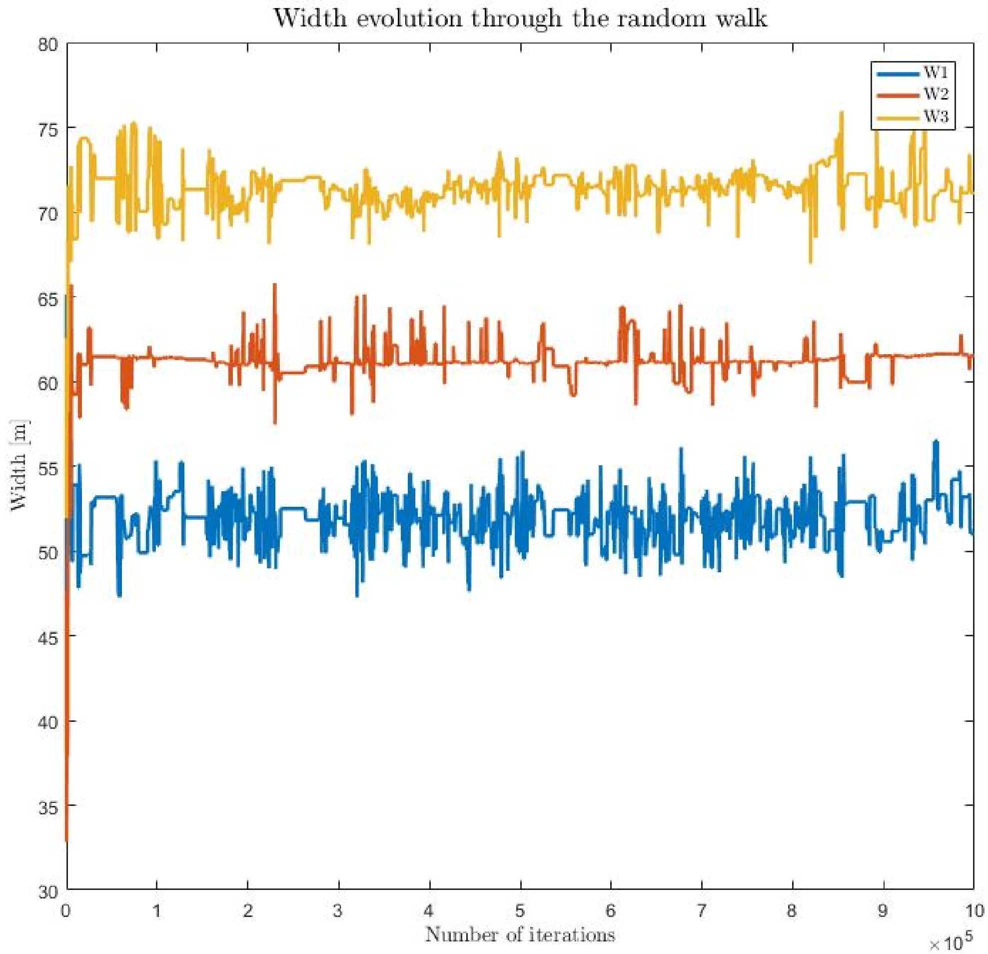


Figure 5.8: Evolution of W_1, W_2, W_3 for $f = 5$, $n = 1 * 10^6$

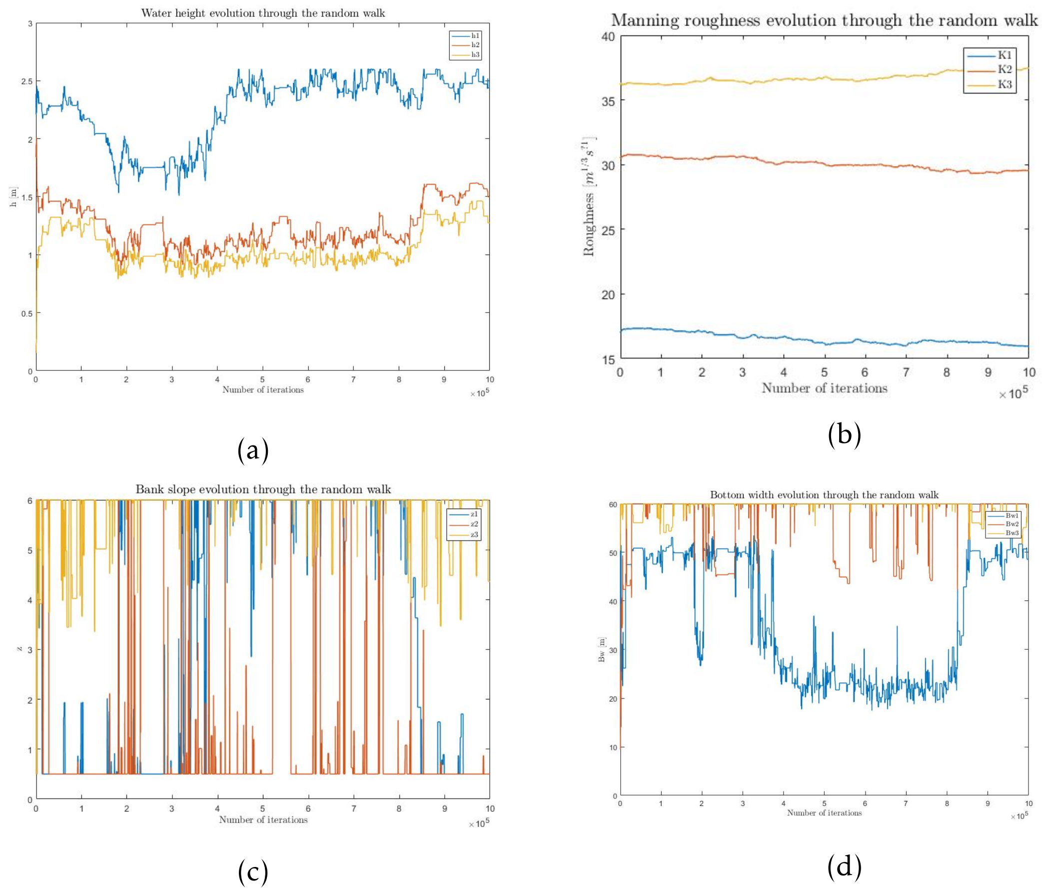


Figure 5.9: Evolution of h (a), K (b), z (c), B_w (d) for $f = 5$, $n = 1 * 10^6$

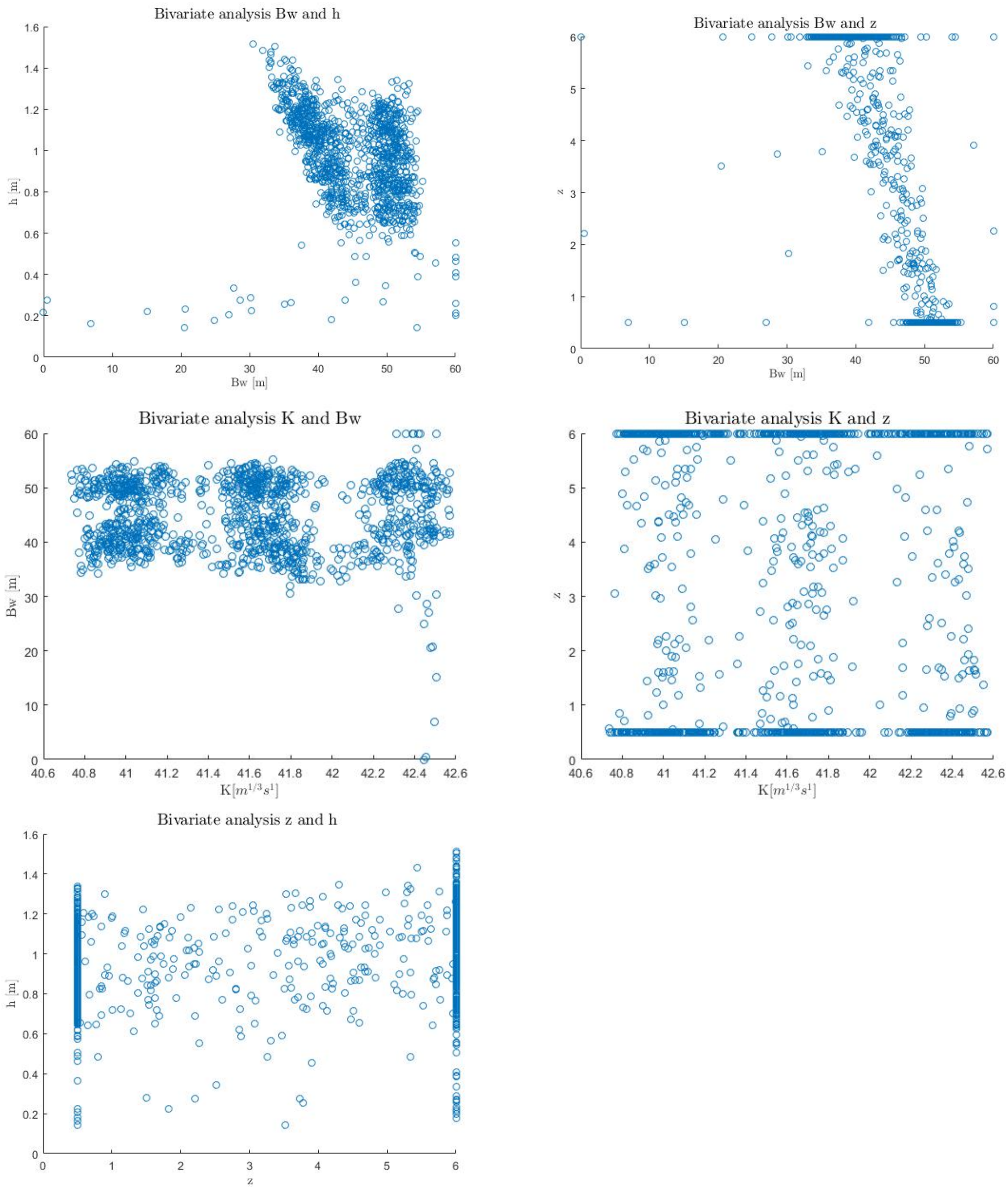


Figure 5.10: Bivariate analysis with $f = 5$, $n = 1 * 10^6$

5.3.2 Real data

In this section, we will present results obtained for real data. For the two stations, three widths have been chosen. they are from the WorldView-3 images. The values are the nearest of the station. The chosen values are presented in Table 5.3 and discharge values measured the day of the images acquisition are $32.2m^3/s$ at Mbalmayo and $27.5m^3/s$ at Olama. They are identified with red lines on Figure 5.11 and Figure 5.12. In those Figures, it seems that results are quite satisfying but as shown in Appendix in Figures G.4 and G.5, it is one chain among others that are worst. Even when we increase the number of iterations it does not give satisfying results. The variation between the different width is maybe too small to lead to a good result.

Mbalmayo widths	Olama widths
81.84	79.36
74.4	74.4
77.5	80.13

Table 5.3: Synthetic data used for the test

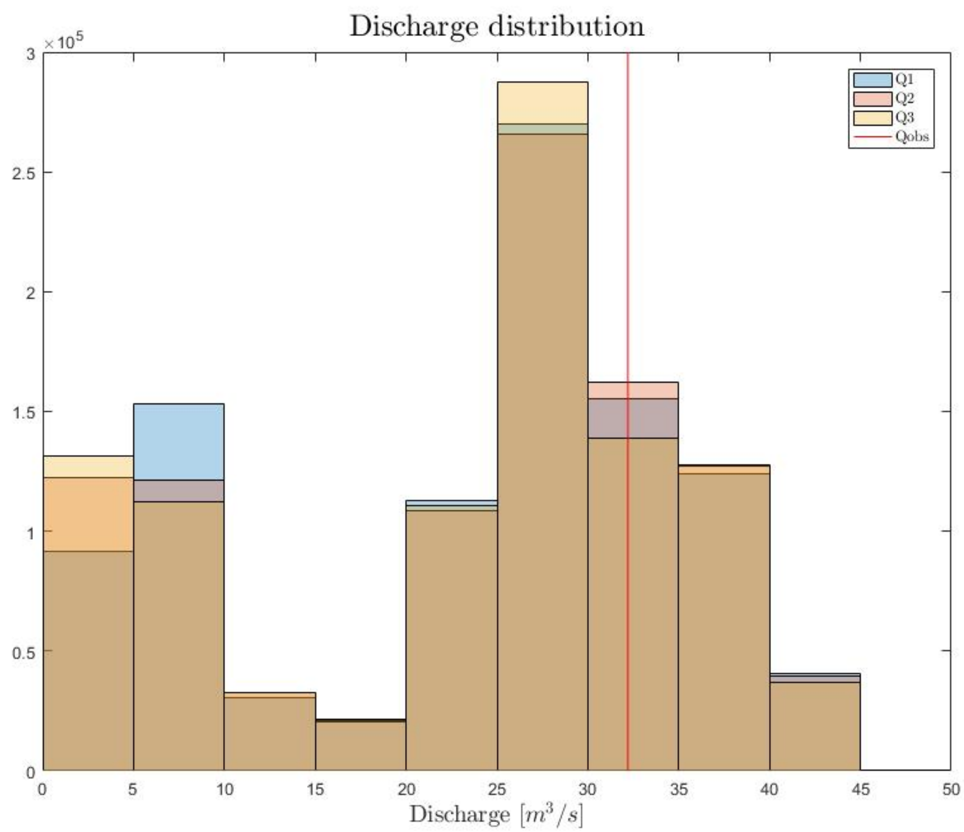


Figure 5.11: Distribution of discharge at Mbalmayo Station

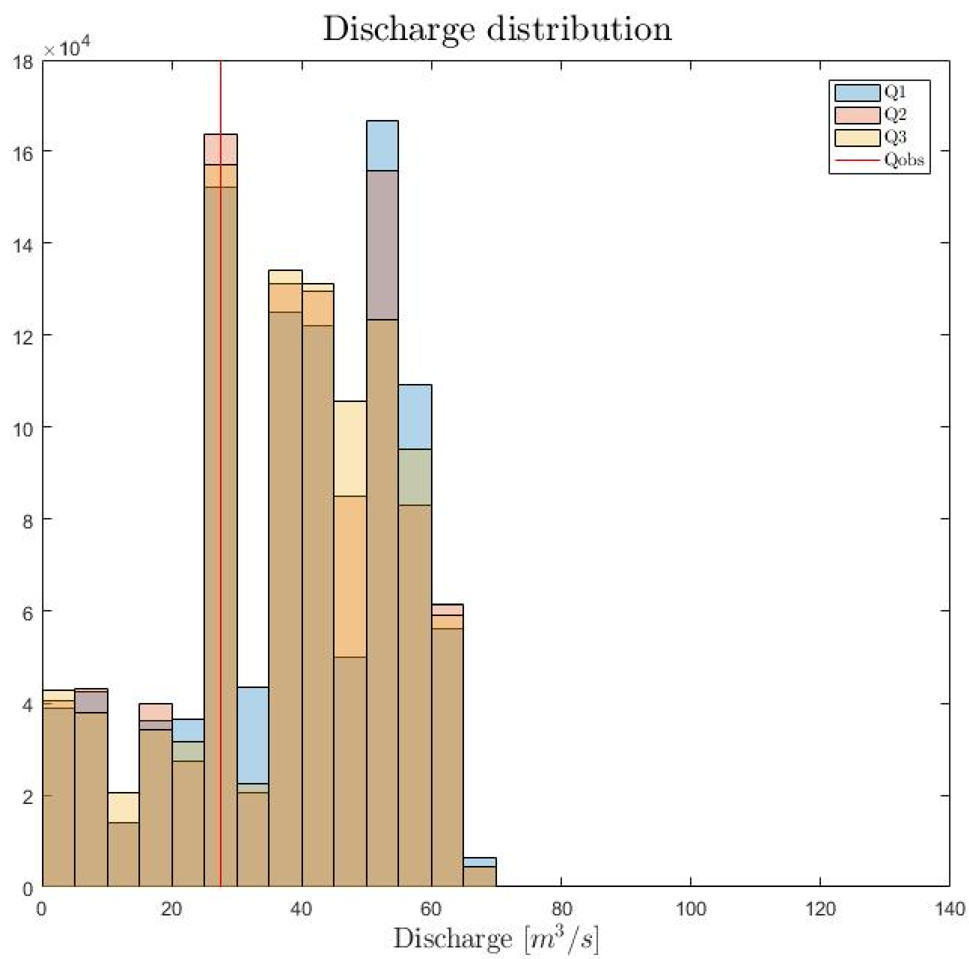


Figure 5.12: Distribution of discharge at Olama Station

Conclusion

In this study, several points have been treated. First,using image processing tools and a creative computation idea, a method to obtain widths from satellite images has been developed. Its efficiency has been proved independently of the river orientation or the satellite used. Results were good for images with high resolution. The test on lower resolution images had given satisfying results. However, as the Nyong river is small, the impact of the differences found could be important on final results. Overall, this method could be a useful tool and not only for hydraulics application. Tests on rivers with different size and morphology will be needed to confirm its efficiency.

Second, using classical laws of open-channel hydraulics, a consistent prior has been created and a simple relation between width and discharge has been established. It seems consistent with the literature but it would have been interesting to investigate further the relation between the different parameters. The quality of the prior has an impact on the final results and also, better knowing the importance of each parameter and their relations could be useful to condition the Metropolis*Hastings.

Finally, the results obtained with a probabilistic inversion are very promising. Indeed, with a generalist approach with regards to hydraulics laws and inversion rules, obtained distribution of river discharge were consistent with the expected value. Without a lot of tests and calibration, the algorithm has given values that were not in outliers. However, the distribution obtained were different between each chain, more iterations will be needed to really confirm the potential of the method.

To sum up, the objectives were to develop a new approach and explore its capacities and also to see if it would work on small rivers. We can conclude that even if there is a lot of possible ameliorations, the uses of probabilistic inversion in order to estimate river discharge using remotely sensed information is a very promising field of study and opens new research perspectives that are discussed in the next chapter. And this method works on small rivers.

Chapter 6 Perspectives

6.1 Methodological issues

The method developed to compute the width demonstrated its efficiency as we have obtained good results for the WorldView-3 and Landsat 7 images. However, the method could be improved by creating another way for finding the centroid as it is a step with a high computing cost. This means it will be difficult to use for very large rivers. Moreover, with the Landsat 7 images, we have used a TOA percentile composite with standard parameters but a customization could impact the results and it will be interesting to explore this aspect. The same remark stands for the export resolution on Google Earth Engine.

Common rules of open-channel hydraulics provided relationships between the different parameters and helps in defining the prior. However, it exists different relations that may be more appropriate to our case study. Open-channel hydraulics was not the main purpose of this work, and within the time given, it was difficult to explore all existing equations. The Nyong river has a lot of aquatic vegetation and we have not take into account the sediment transport in our calculus. It could influence the discharge. Hydraulics is a very complex science with a lot of unknown. In this project, choices about the definition of parameters and the use of Manning's Equation instead of another could be discussed. This part should be carefully examined because the better is the prior and the relationships between its parameters, the better will be the results.

The creation of synthetic data was really helpful to assess the quality of our method and discuss the results obtained. However, when creating it, we have only looked at boundaries for each parameter, independently of the others. But, some combinations obtained are surely false and not exist in natural streams. Two kinds of approaches could be used regarding this issue. The first one could be to construct a river in laboratory and measures and controls every parameter in-situ. The other option will be to better look at the relation existing between the parameters and add conditions when generating the synthetic data.

Probabilistic inversion is a very efficient method to have information about non-linear problems. And in our case, it seems to give encouraging results. However, numerous aspect of our methods could be discussed. First, in the choice of parameters like the size of a step in the random walk or the equation for the likelihood.

We have chosen classic relations in order to have a first idea of the results. But, by using some rules and studies, the calibration of those parameters could improve the results. Moreover, we have used only one Markov chain, but it could have given better results if we had done it on several chains.

6.2 Applied perspectives

The Nyong river is *predominantly covered by tropical rainforest and is free of anthropogenic disturbance (industrial or agricultural)*(Viers et al., 2000). In this context, it could be very interesting to see if the urban expansion of Yaoudé had an impact on the discharge of the Nyong. Indeed, one of the main tributaries of the Nyong river is the Mefou river which drains the region of Yaoudé.

6.3 Limits, open questions and research perspectives

This project presents some limits and new questions have emerged. Thus, there are further interesting developments of this project that might be imagined.

First, the method to compute the width have worked on the different segment that we have tried but to really asses that it is an efficient method to compute the width of rivers, it should be tried on different streams. For example, *the fact that water is almost never clear in the real world results in changes in its spectral curve, making it difficult to use spectral indices with a single threshold to separate water pixels from non-water pixels*(Donchyts et al., 2016). Moreover, the computation of the width with the real resolution of Landsat images (30m) is impossible and with the 10m export, it works under specific conditions. To tackle this problematic, pansharpening on Landsat images could be a good idea to pursue. Indeed, it will give 15m resolution images which should be satisfying for the method and could avoid the change of resolution during the export (which is not controlled). It will also be interesting to work directly in the Google Earth Engine API to perform at least the first fourth steps of the analysis as it will apply to any area of interest. It will also be useful to customize for example the parameters of the TOA percentile composites or explore if other bands than the NIR could work.

Second, there were some limits regarding the hydraulic part. We had a lack of knowledge of the channel geometry and more generally of the relation between the different parameters and their influence on the result. A sensitivity analysis of the parameters could give insights into how they are linked and how they influence the result. It could be location specific or linked to the vegetation or sediment. Moreover, we have assumed that the discharge was constant through a certain segment of the river, but we have data only at the station and they could be different 200m up or down-stream.

Then, some limits have appears during the probabilistic inversion. The results obtained still have a lot of errors. One idea that could be tested is to add a condition in the Metropolis-Hastings using the stage-discharge curve that will maybe improve the general result. Another condition could be the relation between discharges from the two stations. Moreover, as the results are different between each chain, an idea would be to combine the method with knowledge from data mining. For example, a neural network method could teach the algorithm to restrain between different chains some parameters. It would certainly help for the overall result but also for the estimation of the different hydraulic parameters.

Another interesting area to complete this research will be to use results from other studies to estimate and bound the hydraulic parameters in order to create a method that really uses only remote sensed information. For example, using methods described in (Andreadis et al., 2007) or in (Bjerklie et al., 2005).

Probabilistic inversion has been proved to be a good strategy in geophysics and give good results in our method. Using simple concepts we have obtained promising results. It should be more and more used.

Finally, the Surface Water and Ocean Topography (SWOT) satellite mission will certainly create new possibilities regarding the subject of this project. The launch planned in 2020, and as described in (Biancamaria et al., 2016), discharge estimation directly from SWOT data will be possible. There will be some limits, for example, it will not necessarily apply small or braided rivers, but the data provided could at least help to improve the definition of the prior in the context of our project. Moreover proof-of-concept experiments are realized using synthetic generated SWOT measurements (Durand et al., 2008). The use of probabilistic inversion regarding hydraulic parameters estimation seems totally founded to complete those studies and better are the data for the prior, better are the results.

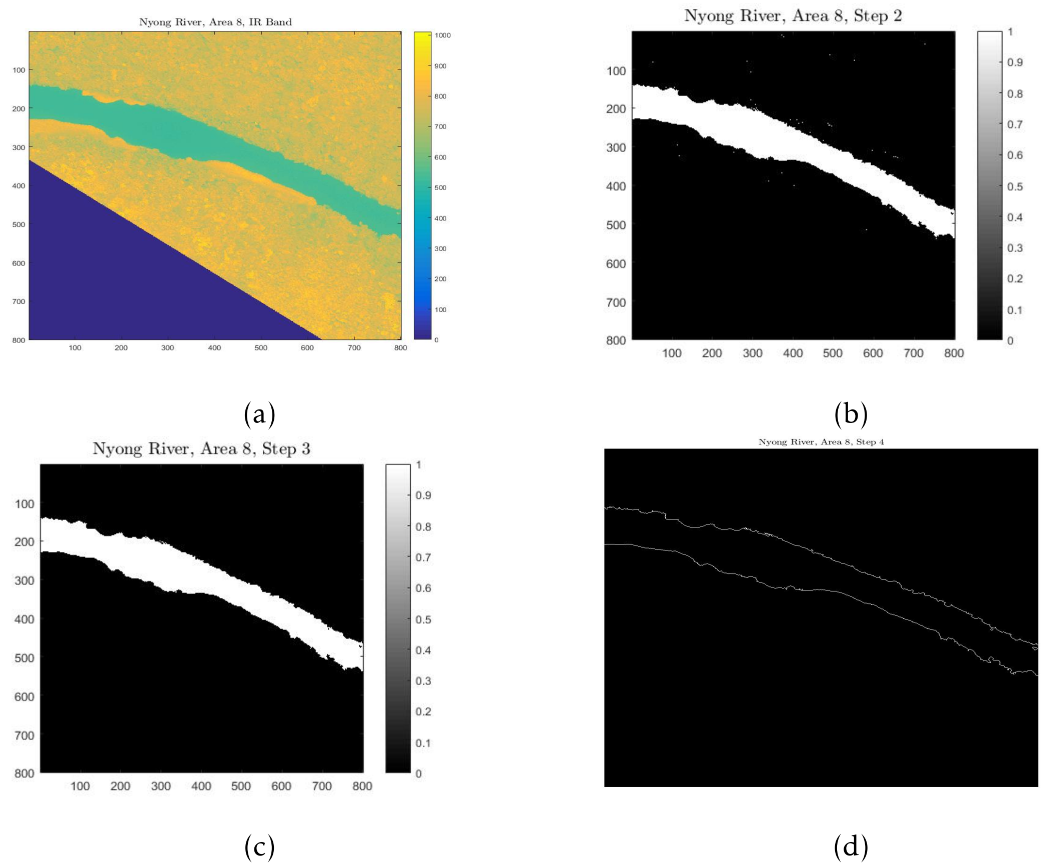
Appendix A State of the art by Sichangi and al.

Table 1
Overview of studies using remote sensing to estimate river discharge (the current paper is added for completeness). Not all studies have performed river discharge estimation by incorporating multiple remote sensing derived datasets. In addition, not all studies have compared the accuracies in considering the single verses the multiple derived remote sensing data. Not all models used in the studies are optimized for diverse river cross-sections. In this case, the three are flagged as not applicable (N/A) in the key results. In the 'Key results' column the abovementioned three components are identified by the code: (1) use of multiple remote sensing derived dataset, (2) comparison in discharge estimate using single verses multiple derived remote sensing data and (3) performance over diverse river cross-sections.

Study	Data used	Approach	Key results
Gleason and Smith (2014)	Landsat TM	Landsat TM are used to approximate at-many-stations (AMHG). AMHG used to retrieve instantaneous river discharge.	1. N/A 2. N/A 3. AMHG discharge retrievals are successful for most investigated river morphologies. However, poor performance is observed in a few rivers (Gleason et al., 2014).
Birkinshaw et al. (2014)	ERS 2, Envisat & Landsat	The satellite data are applied to the Bjerklie et al. (2003) equation.	1. Multiple parameters (width, slope, stage) 2. N/A 3. Nash–Sutcliffe efficiency values of 0.90 at Nakhon Phanom and 0.86 at Vientiane on the Mekong, and 0.86 at Kalpashevo on the Ob
Pavelsky (2014)	Rapid Eye & Landsat	Width based rating curves are used.	1. N/A 2. N/A 3. Optimized to use width and tested on river Tanana with a 6.7% relative error
Tarpanelli et al. (2013)	ERS 2 & ENVISAT	Rating Curve Model is applied to estimated discharge using altimetry data. Comparison of the method is carried out against the empirical equation proposed by Bjerklie et al. (2003).	1. N/A 2. N/A 3. Optimized to use stage and tested on Po River with a relative root mean square error of 30%
Negrel et al. (2011)	Acoustic Doppler Current Profiler (ADCP)	The hydraulic parameters values are formulated using least square method by minimization of the error criterion.	1. No remote sensing data was used to test this method. Instead datasets were taken at gauging stations in the Amazon basin. 2. N/A 3. Two sites Obidos and Manacupuru are considered with the former recording poor performance.
Birkinshaw et al. (2010)	ERS-2 and ENVISAT	Stage based rating curves are used	1. N/A 2. N/A 3. Optimized to use stage giving Nash–Sutcliffe values between 0.823 and 0.935 on River Mekong
Smith and Pavelsky (2008)	MODIS	Width based rating curves are used	1. N/A 2. N/A 3. Optimized to use river width giving a mean absolute error <25% on River Lena
LeFavour and Alsdorf (2005)	SRTM DEM, SAR, nautical charts	Manning's equation is used with SRTM DEM slope and with channel width measurements from the SAR, channel depths averaged from nautical charts, and reasonable estimates of Manning's *n*.	1. Remotely sensed parameters width and slope are used 2. N/A 3. Three sites are considered giving discharge values within 6.2% at Manacupuru, 7.6% at Itapeua, and 0.3% at Tupe of the in situ gage-based estimates.
Bjerklie et al. (2005)	Topographical maps aerial photos, SAR	Hydraulic relationship is used to estimate in-bank river discharge using remotely sensed data (Bjerklie et al., 2003).	1. Remote sensing derived slope, velocity and width are used 2. N/A 3. Attained mean discharge estimate accuracy within 10% over River Missouri.
Bjerklie et al. (2003)		Multiple regression analysis used to develop multi-variate river discharge estimating equations is developed.	1. Different combinations of remote sensing derived effective river width, stage and surface velocity of river estimates discharge with average uncertainty of <20%. 2. N/A 3. N/A
This study	MODIS & Multiple Satellite altimetry	River stage is incorporated with the effective river width. A comparison is done against a stage derived discharge estimation equation and the empirical equation proposed by (Bjerklie et al., 2003).	1. Satellite derived river stage and effective river width are used. 2. A comparison is drawn showing a better performance. 3. We formulate our model to capture variations in river cross sections. Good performance is exhibited in the tested cross sections.

Figure A.1: Overview of studies using remonte sensing to estimate river discharge (Sichangi et al., 2016)

Appendix B Image processing, Area 8



(a) (b) (c) (d)

Figure B.1: Isolation of the river, Area 8

Appendix C Water Height

Year	Month	Day	h	Q
1964	11	12	5,45	750
1964	11	21	5,07	655
1964	11	27	4,84	611
1964	12	9	4,13	447
1964	12	24	3,33	287
1964	12	29	3,07	241
1965	1	12	2,79	201
1965	1	29	2,3	142
1965	2	9	2,22	127
1965	2	24	2,38	142
1965	3	13	2,69	180
1965	4	20	2,64	1,69
1966	2	5	2,06	109
1966	3	1	1,06	35,8
1966	3	11	1,41	51,5
1966	3	21	136	40,4
1966	4	1	1,45	47,2
1967	2	25	169	73,9
1969	12	6	3,83	389
1971	2	15	1,44	49,1
1972	3	2	1,16	25
1973	3	13	0,78	21,4
1976	1	6	2,65	184,8
1976	2	23	1,82	85
1977	3	19	1,09	23,4

Figure C.1: Water height data for Olama station

Appendix D Roughness Coefficient

Values for the computation of the roughness coefficient (Chow, 1959)

Channel Conditions		Values	
Material Involved	Earth	n0	0.025
	Rock Cut		0.025
	Fine Gravel		0.024
	Coarse Gravel		0.027
Degree of irregularity	Smooth	n1	0.000
	Minor		0.005
	Moderate		0.010
	Severe		0.020
Variations of Channel Cross Section	Gradual	n2	0.000
	Alternating Occasionally		0.005
	Alternating Frequently		0.010-0.015
Relative Effect of Obstructions	Negligible	n3	0.000
	Minor		0.010-0.015
	Appreciable		0.020-0.030
	Severe		0.040-0.060
Vegetation	Low	n4	0.005-0.010
	Medium		0.010-0.025
	High		0.025-0.050
	Very High		0.050-0.100
Degree of Meandering	Minor	m5	1.000
	Appreciable		1.150
	Severe		1.300

$$n = (n_0 + n_1 + n_2 + n_3 + n_4)\, m_5$$

Figure D.1: Values and Equation for the computation of the roughness coefficient

Appendix E Pansharpening

Pansharpening is the process of merging a high-resolution panchromatic image with a lower multispectral image acquired over the same area in order to obtain a higher resolution color image. In the present work, it was first explored in order to improve the resolution of Landsat images.

Different methods exist that can be divided into *two categories: the component substitution (CS) family and multi-resolution analysis (MRA) family. The CS approaches focus on the substitution of a component that is obtained by a spectral transformation of the MS bands with the PAN image. The MRA-based techniques rely on the injection of the spatial details that are obtained through a multi-resolution decomposition of the PAN image into the up-sampled MS bands* (Li et al., 2017).

Various studies have been done in order to compare and test pansharpening techniques Snehmani et al. (2016),Vivone et al. (2015). To obtain the best result, a careful selection of the fusion method is required because each pansharpening method will induce distortions. Based on the three following paper Li et al. (2017),Snehmani et al. (2016), Vivone et al. (2015) the method that seems to be the better for our case is a technique from the MRA family that use the Generalized Laplacian Pyramids (GPL) with Context Based Decision (CBD) model as it gives the better performance, particularly for water bodies on WV-3 images. Unfortunately, it is a method that is complex to implement and as it is not the main objective of the project, we have decided to only test a CS method which is the Intensity Hue Saturation (IHS), using ArcGIS, as it was quite easy to run and to see if the differences were really important. The aim of this step is not to have a real quantitative measure but more a qualitative idea of the impact that pansharpening could have. It would have been also relevant to perform it on Landsat images as they are the one who needed the most to improve resolution. But, it was difficult to do because we used TOA percentile composites products and not raw data.

Intensity Hue Saturation method The IHS pan-sharpening method converts the multi spectral image from RGB to intensity, hue, and saturation. The low-resolution intensity is replaced with the high-resolution panchromatic image. We have used the default parameters of ArcGis and the result is presented in Figure E.1.

Limits In Figure E.2, the result of step 2 applied to the pan-sharpened image is presented. The problem is clearly visible, the river is now a combination of small distinct objects and when the connected component analysis is done, the bigger ob-

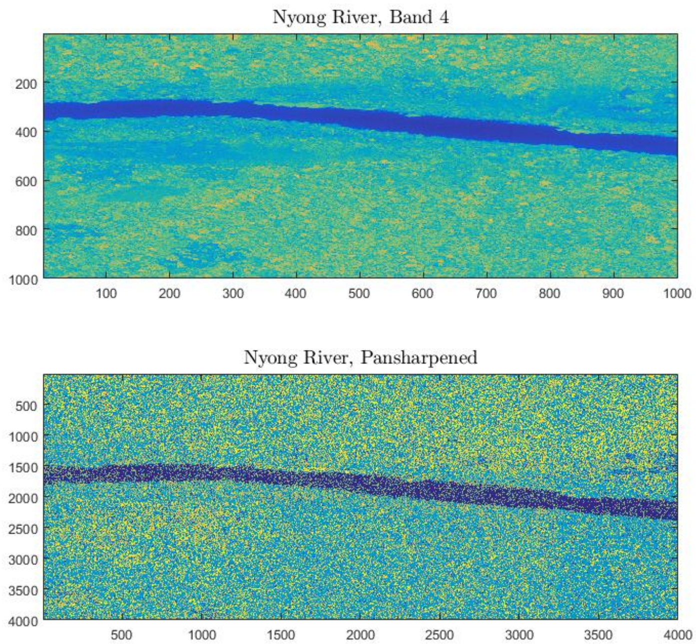
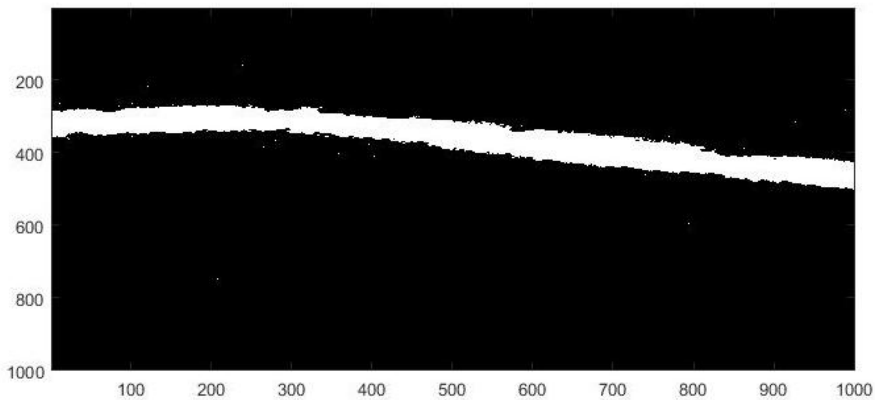
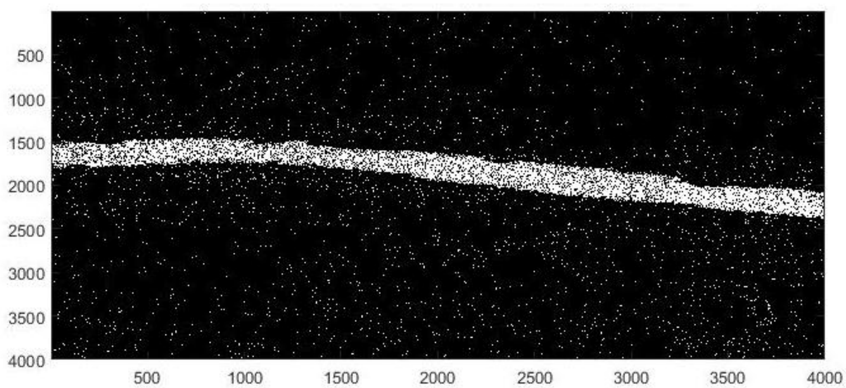


Figure E.1: Nyong River, World View 3, MS image and Pansharpened image (IHS)

ject is not anymore the river. So it is impossible to find the river banks and compute the width. It could be something specific to the Nyong river because it contains a lot of aquatic plants (Olivry, 1986). The resolution of WV-3 panchromatic is 0.31cm, which will be enough to detect a pile of aquatic grass. Pansharpening could be a useful tool but at the scale of a large river (¡100m), a resolution around a meter seems sufficient.



(a)



(b)

Figure E.2: Nyong River, Band 4 image (a) and Pan-sharpened image (b), step 2

Appendix F Coding structure

1. image_analysis.m (see Chapter 3):
    - import the file and select the NIR band;
    - Cropping;
    - Threshold selection
    - Connected Component analysis;
    - Bwmorph operation;
    - Distance to bank;
    - Distance to bank and centroid determination;
    - Width computation;
    - Plot the results.
2. analysis_bvet_data.m (see Chapter 2):
    - Put data in Year Month Day Discharge format;
    - Plot the discharge for each year;
    - Comparison between the two stations;
    - Plot the dry years.
3. Comparaison_bvet_2017.m (see Chapter 2):
    - Select data for the good period (February-March);
    - Find the years that are similar;
    - Plot the results.
4. synthetic_data.m (see Chapter 4):
    - Boundaries definition;
    - Generation of possible values of parameters;
    - Loop to create all combinations;
    - Computation of W, A, P, Rh and Q;
    - Creation of a subsection with Q and K constants and W around values measured at Mbalmayo station.

5. metropolis.m (See Chapter 5):

    - Definition of the prior;
    - Initialization, m_{curr} ;
    - Compute the width, the discharge and the likelihood using likelihood.m;
    - Definition of the size of the step.
    - Begin the loop
    - Definition of m_{prop};
    - Compute the width, the discharge and the likelihood using likelihood.m of the proposal;
    - Definition of the acceptance (”if” loop);
    - Update the current stage;
    - Store the result;
    - End of both loops.
    - Plot the results.

6. likelihood.m (See Chapter 5):

    - Definition of the input of the function;
    - Compute the log likelihood;
    - Compute the likelihood using exponential;

Appendix G Results

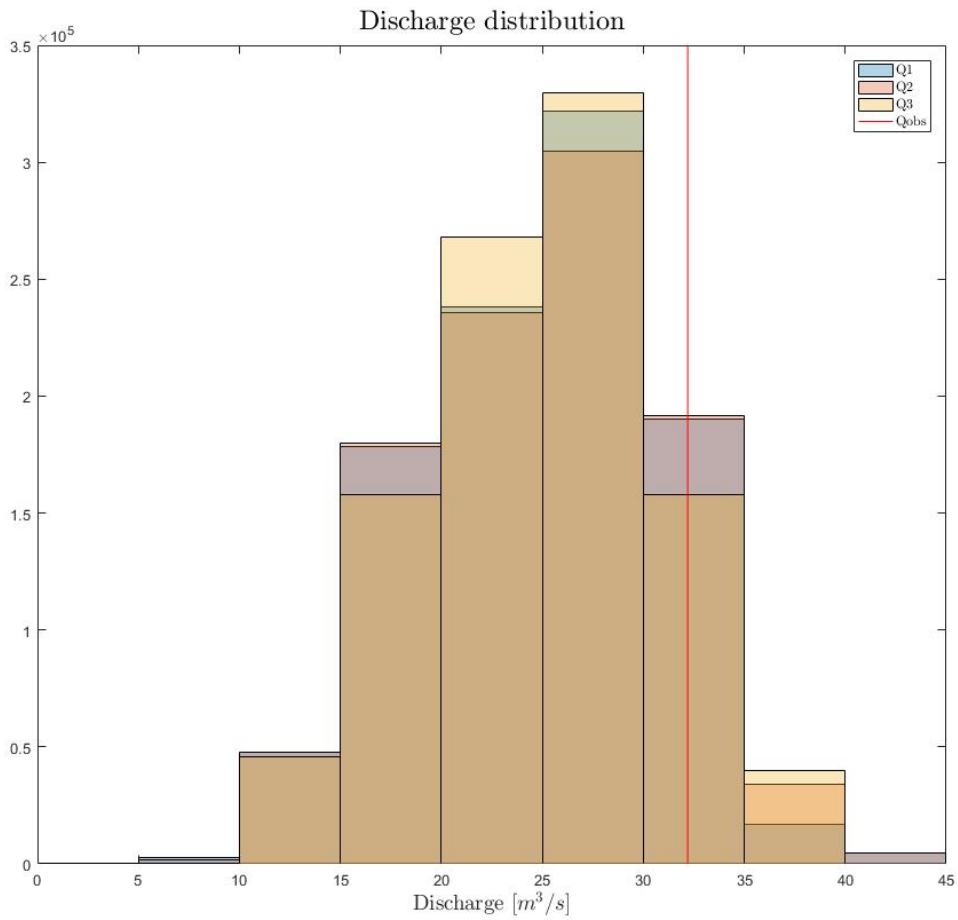


Figure G.1: Distribution of discharge for $f = 5$

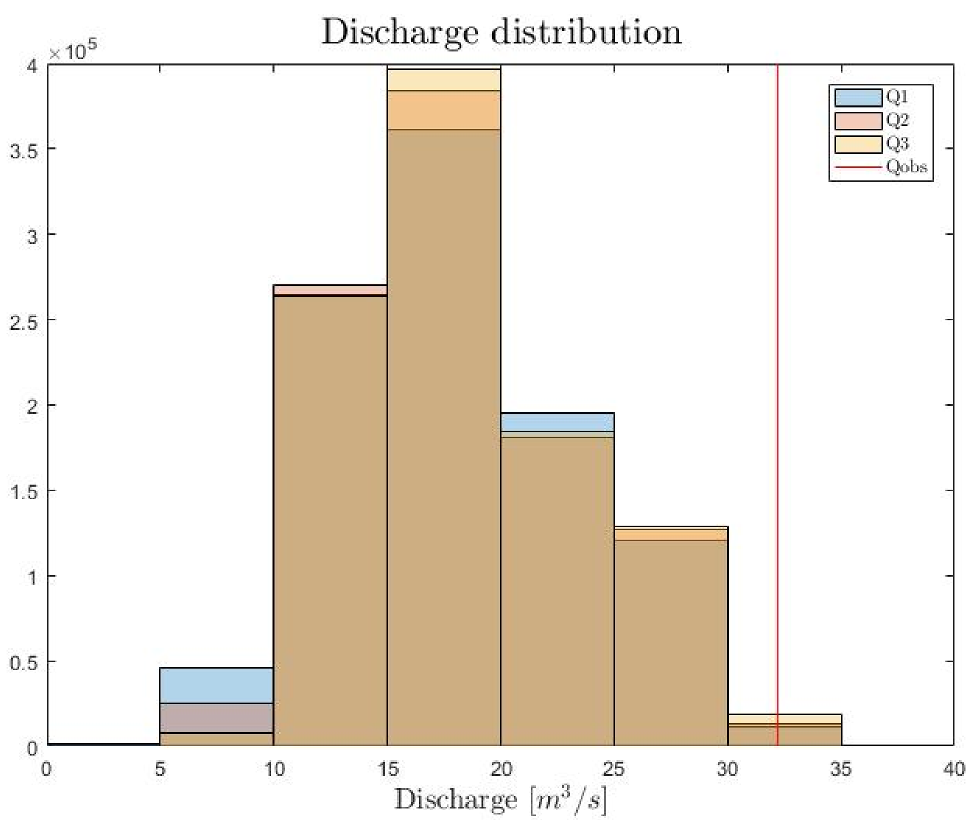
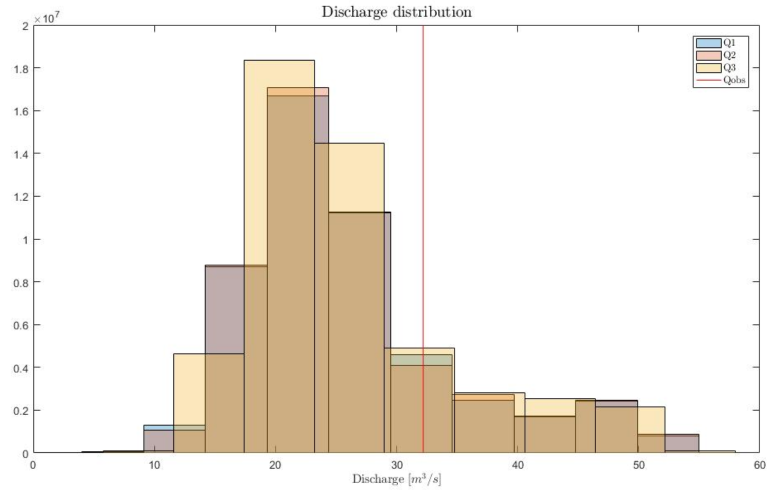
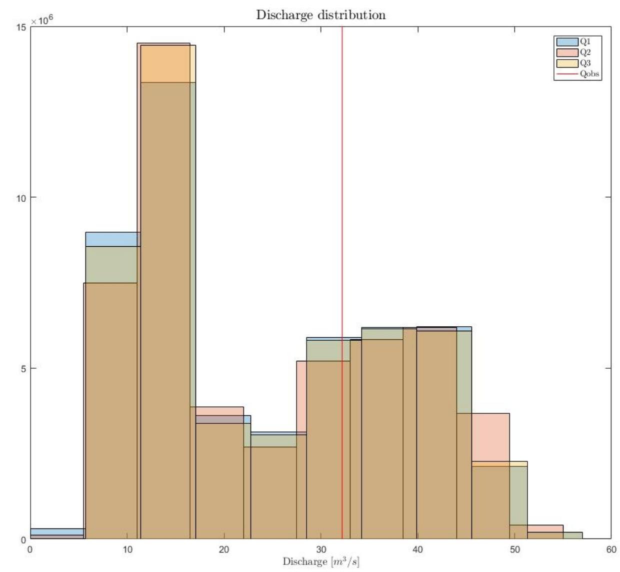


Figure G.2: Distribution of discharge for $f = 1$





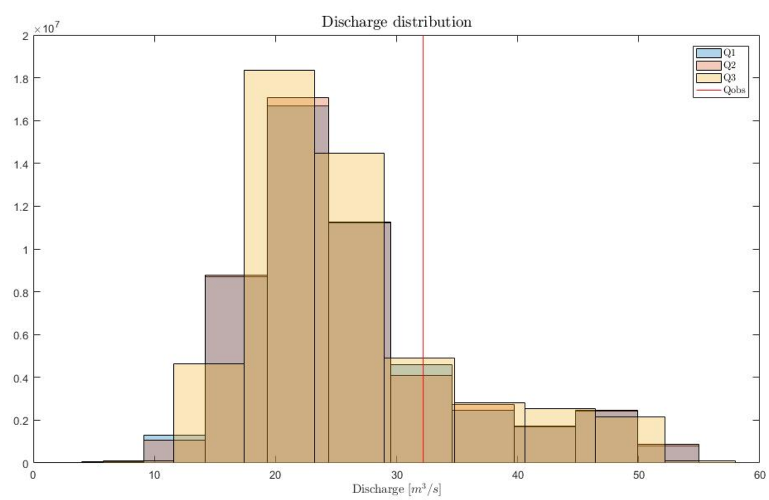


Figure G.3: Distribution of discharge for $f = 5$, $n = 5 * 10^7$

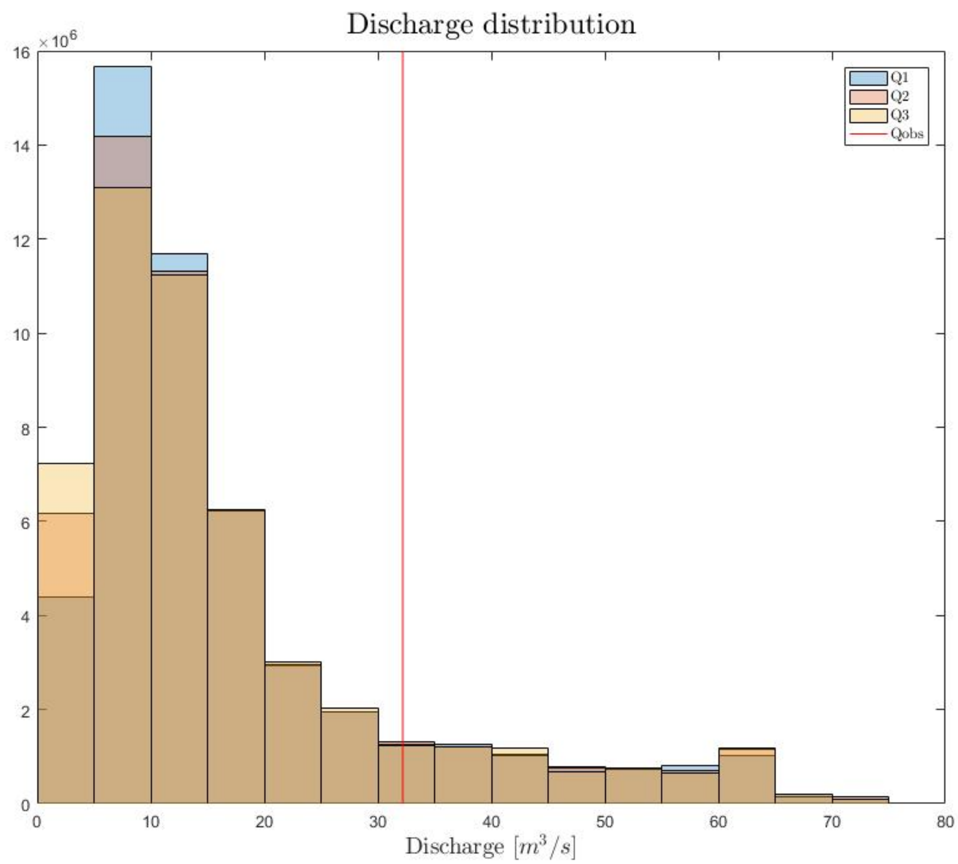


Figure G.4: Distribution of discharge at Mbalmayo Station

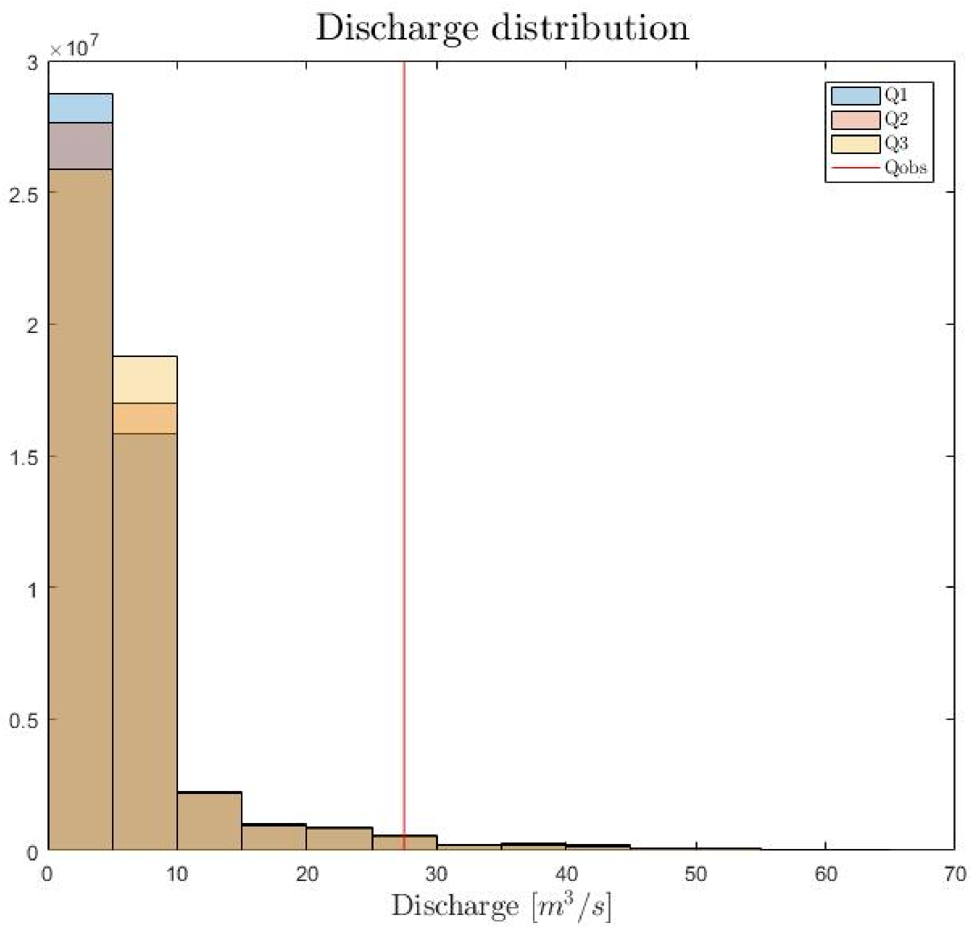


Figure G.5: Distribution of discharge at Olama Station

Bibliography

Doug Alsdorf, Dennis Lettenmaier, and Charles Vörösmarty. The need for global, satellite-based observations of terrestrial surface waters. *Eos, Transactions American Geophysical Union*, 84(29):269–276, 2003.

Konstantinos M Andreadis, Elizabeth A Clark, Dennis P Lettenmaier, and Douglas E Alsdorf. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, 34(10), 2007.

Sylvain Biancamaria, Dennis P Lettenmaier, and Tamlin M Pavelsky. The swot mission and its capabilities for land hydrology. *Surveys in Geophysics*, 37(2):307–337, 2016.

David M Bjerklie, Delwyn Moller, Laurence C Smith, and S Lawrence Dingman. Estimating discharge in rivers using remotely sensed hydraulic information. *Journal of Hydrology*, 309(1-4):191–209, 2005.

JN Callow and GS Boggs. Studying reach-scale spatial hydrology in ungauged catchments. *Journal of hydrology*, 496:31–46, 2013.

James B. Campbell and Randolph H. Wynne. *Introduction to remote sensing*. 5th Edition. The Guilford Press, 2011.

Anny Cazenave, Nicolas Champollion, Jérôme Benveniste, and Jianli Chen. Foreword: International space science institute (issi) workshop on remote sensing and water resources. *Surveys in Geophysics*, 37(2):191–194, 2016.

Gyanesh Chander, Brian L Markham, and Dennis L Helder. Summary of current radiometric calibration coefficients for landsat mss, tm, etm+, and eo-1 ali sensors. *Remote sensing of environment*, 113(5):893–903, 2009.

Van Te Chow. *Open-channel hydraulics*. 5th Edition. McGraw-Hill Book Company, 1959.

Petra Döll, Hervé Douville, Andreas Güntner, Hannes Müller Schmied, and Yoshihide Wada. Modelling freshwater resources at the global scale: Challenges and prospects. *Surveys in Geophysics*, 37(2):195–221, 2016.

Gennadii Donchyts, Jaap Schellekens, Hessel Winsemius, Elmar Eisemann, and Nick van de Giesen. A 30 m resolution surface water mask including estimation of positional and thematic differences using landsat 8, srtm and openstreetmap: a case study in the murray-darling basin, australia. *Remote Sensing*, 8(5):386, 2016.

Michael Durand, Konstantinos M Andreadis, Douglas E Alsdorf, Dennis P Lettenmaier, Delwyn Moller, and Matthew Wilson. Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophysical Research Letters*, 35(20), 2008.

Colin J Gleason and Laurence C Smith. Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. *Proceedings of the National Academy of Sciences*, 111(13):4788–4791, 2014.

Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins. *Digital Image Processing using Matlab*. 2nd Edition. McGraw-Hill, 2009.

Google. Google earth engine api, 2017. URL `https://developers.google.com/earth-engine/scale`.

N Hassan, JRA Hamid, NA Adnan, and M Jaafar. Delineation of wetland areas from high resolution worldview-2 data by object-based method. In *IOP Conference Series: Earth and Environmental Science*, volume 18, page 012017. IOP Publishing, 2014.

INSU. Bassins versants exprimentaux tropicaux, 2017. URL `http://bvet.obs-mip.fr/fr`.

Abdollah A Jarihani, John N Callow, Tim R McVicar, Thomas G Van Niel, and Joshua R Larsen. Satellite-derived digital elevation model (dem) selection, preparation and correction for hydrodynamic modelling in large, low-gradient and data-sparse catchments. *Journal of Hydrology*, 524:489–506, 2015.

Hui Li, Linhai Jing, and Yunwei Tang. Assessment of pansharpening methods applied to worldview-2 imagery fusion. *Sensors*, 17(1):89, 2017.

Thomas M. Lillesand, Ralph W. Kiefer, and Jonathan W. Chipman. *Remote sensing and image interpretation*. 7th Edition. Wiley, 2015.

Grégoire Mariethoz. Cours 12, pansharpening. Teledetection, Bachelor en Environnement et Geosciences, 2017.

Larry W Mays. *Water resources engineering*. John Wiley & Sons, 2010.

Klaus Mosegaard and Albert Tarantola. Monte carlo sampling of solutions to inverse problems. *Journal of Geophysical Research: Solid Earth*, 100(B7):12431–12447, 1995.

J Negrel, P Kosuth, and N Bercher. Estimating river discharge from earth observation measurement of river surface hydraulic variables. *Hydrology and Earth System Sciences*, 15(6):p–2049, 2011.

Jean-Claude Olivry. *Monographie du Nyong et des fleuves cotiers*. Tome 1. ONAREST, 1979.

Jean-Claude Olivry. *Fleuves et rivières du Cameroun*. MESRES-ORSTOM, 1986.

Feifei Pan and Jennifer Nichols. Remote sensing of river stage using the cross-sectional inundation area-river stage relationship (iarsr) constructed from digital elevation model data. *Hydrological Processes*, 27(25):3596–3606, 2013.

Ben Shaver. A zero-math introduction to markov chain monte carlo methods, 2017. URL `https://towardsdatascience.com`.

Arthur W Sichangi, Lei Wang, Kun Yang, Deliang Chen, Zhongjing Wang, Xiuping Li, Jing Zhou, Wenbin Liu, and David Kuria. Estimating continental river basin discharges using multiple remote sensing data sets. *Remote Sensing of Environment*, 179:36–53, 2016.

Laurence C Smith. Satellite remote sensing of river inundation area, stage, and discharge: A review. *Hydrological processes*, 11(10):1427–1439, 1997.

Akshay Gore Snehmani, Ashwagosh Ganju, Satish Kumar, and PK Srivastava. A comparative analysis of pansharpening techniques on quickbird and worldview-3 images. 2016.

Qiuhong Tang, Huilin Gao, Hui Lu, and Dennis P Lettenmaier. Remote sensing: hydrology. *Progress in Physical Geography*, 33(4):490–509, 2009.

Luke Tierney. Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728, 1994.

USGS. Landsat missions, 2017. URL `https://landsat.usgs.gov/`.

Jérôme Viers, Bernard Dupré, Jean-Jacques Braun, Samuel Deberdt, Bernard Angeletti, Jules Ndam Ngoupayou, and Annie Michard. Major and trace element abundances, and strontium isotopes in the nyong basin rivers (cameroon): constraints on chemical weathering processes and elements transport mechanisms in humid tropical environments. *Chemical Geology*, 169(1-2):211–241, 2000.

Gemine Vivone, Luciano Alparone, Jocelyn Chanussot, Mauro Dalla Mura, Andrea Garzelli, Giorgio A Licciardi, Rocco Restaino, and Lucien Wald. A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5):2565–2586, 2015.

Charles J Vörösmarty, Ellen M Douglas, Pamela A Green, and Carmen Revenga. Geospatial indicators of emerging water stress: an application to africa. *AMBIO: A journal of the Human Environment*, 34(3):230–236, 2005.

Tobias Zahner, Tobias Lochbühler, Grégoire Mariethoz, and Niklas Linde. Image synthesis with graph cuts: a fast model proposal mechanism in probabilistic inversion. *Geophysical Journal International*, 204(2):1179–1190, 2015.