Use of multi-temporal and multi-sensor data for continental water body extraction in the context of the SWOT mission

Nicolas Gasnier

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2022IPPAT002

Thèse de doctorat

Use of multi-temporal and multi-sensor data for continental water body extraction in the context of the SWOT mission

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat : Images

Thèse présentée et soutenue à Palaiseau, le 11 janvier 2022, par

NICOLAS GASNIER

Composition du Jury :

Hugues Talbot Professeur, CentraleSupélec (Centre de Vision Numérique)	Président
Abdourrahmane M. Atto Maitre de Conférence, Université Savoie Mont Blanc, (LISTIC)	Rapporteur
René Garello Professeur Émérite, IMT Atlantique (Lab-STICC)	Rapporteur
Lionel Bombrun Maître de Conférences, Université de Bordeaux (IMS)	Examinateur
Sylvie Le Hégarat Professeure, Université Paris-Saclay (SATIE)	Examinatrice
Florence Tupin Professeure, Télécom Paris (LTCI)	Directrice de thèse
Loïc Denis Maître de Conférences, Université de Saint-Étienne (LHC)	Co-directeur de thèse
Roger Fjørtoft Ingénieur, Centre National d'Études Spatiales	Invité
Frédéric Liège Ingénieur, CS Group France	Invité

À ma mère

Remerciements

Je souhaite tout d'abord remercier le Centre National d'Études Spatiales (CNES) et l'entreprise CS Group France qui ont financé ces travaux, m'ont fait confiance pour les réaliser et m'ont donné les moyens de travailler dans de très bonnes conditions.

Je remercie également les membres de mon Jury, présidé par M. Hugues Talbot, qui ont accepté d'évaluer mes travaux et avec qui j'ai eu des échanges très intéressants lors de la soutenance. Merci en particulier aux rapporteurs : M. Abdourrahmane Atto et M. René Garello qui ont accepté de s'adapter à un calendrier contraint pour préparer leurs rapports et qui sont venus de loin pour assister à ma soutenance.

Je souhaite tout particulièrement remercier mes encadrants, à qui je sais que je dois beaucoup : Florence Tupin, pour tous ses conseils, pour sa disponibilité et pour avoir su concilier exigence et bienveillance dans son encadrement. Loïc Denis, pour ses nombreuses idées toujours brillantes, son optimisme et sa gentillesse. Merci aussi à Roger Fjørtoft qui m'a ouvert les portes du CNES et de la mission SWOT. Je n'aurais pas pu trouver un meilleur spécialiste du sujet. Ton souci du détail m'a aussi permis de beaucoup améliorer ce manuscrit. Merci enfin à Frédéric Liège, qui a tout fait pour que la thèse se passe dans les meilleures conditions, et qui n'a cessé de m'encourager dans mes travaux.

J'ai effectué la majeure partie des travaux de ma thèse au sein de l'équipe IMAGES du département IDS à Télécom Paris. Je tiens à en remercier le personnel administratif pour son aide et sa disponibilité. Merci aussi, bien sûr, aux enseignants-chercheurs de l'équipe : Jean-Marie, avec qui j'ai eu des discussions passionnantes et qui a essayé de me transmettre sa passion pour les fonctions de Meijer, à Christophe[1], à Pietro et Saïd avec qui j'ai beaucoup aimé plaisanter[2], à Henri pour ses conseils et pour son aide lorsque j'ai eu besoin de préparer des vérités terrain, à Michel, Yann, Isabelle, Kiwon, Amal, Tamy, Jean-Marc et Alasdair avec qui j'ai eu le plaisir d'échanger au cours de ma thèse.

J'ai aussi eu la chance de croiser de nombreux collègues doctorants, post-doctorants et stagiaires sans qui ces années de thèse n'auraient pas été aussi riches. Merci à Clément, Ivan, Sylvain et à Nicolas G[3] qui m'ont accueilli dans l'équipe, à Mateus et à son énergie débordante, à Emanuele[4], arrivé peu après moi dans l'équipe et avec qui j'ai eu la chance de participer

[1] Qui a rejoint l'équipe RADAR vers la fin de ma thèse, mais avec qui j'espère avoir l'occasion de faire de la séparation de sources un jour.

[2] et râler, aussi, parfois...

[3] Qui sait qu'ici, pour une fois, il ne sera pas confondu avec moi

[4] Que je recommande vivement comme voisin de bureau et comme fournisseur de café ou de méthodes de débruitage.

à une école d'été sur le radar en Allemagne[5], à défaut d'avoir pu participer physiquement à une conférence pendant ma thèse. À Antoine et Xiangli avec qui aussi j'ai partagé un bureau rue Barrault. Merci également à Chenguang[6], Arthur[7], Alessio, Bastien, Corentin, Nicolas C, Élie, Raphael, Matthis, Inès[8], Zoé[8], Giammarco, Chloé, Pierrick, Guillaume, Florence, Rassim, Gwilherm et à tous les autres. Je suis très content d'avoir pu vous rencontrer et j'espère avoir l'occasion de vous recroiser bientôt. Je souhaite bon courage pour la fin de leur thèse à tous les doctorants qui n'ont pas encore soutenu. Je sais en tout cas qu'ils seront très bien entourés, comme je l'ai été[9]. En arrivant au laboratoire, j'ai eu la chance d'être « élu » représentant des doctorants au conseil du département IDS, ce qui a été une expérience très enrichissante pour moi. Je souhaite remercier ses membres et en particulier Gaël.

Du coté de CS Group, je tiens à remercier Frédéric, bien sûr, mais aussi tous ceux avec qui j'ai pu échanger, à distance ou pendant mes passages à Toulouse. Je pense en particulier à Kévin, Lucie[10], Cécile et Manon[11], Luc, sans qui je n'aurais jamais réussi à faire fonctionner le code C++ des méthodes GrabCut, Sébastien, dont l'expertise en licences logiciel a été indispensable pour pouvoir publier les codes que j'avais écrits.

Au CNES aussi, j'ai eu la chance de travailler avec des gens passionnés dans une très bonne ambiance. Merci à Roger, Damien, Emmanuelle, Claire, Nadine, Céline, Santiago, Thierry[12] et à tous ceux que j'ai eu l'occasion de croiser lors de mes trop rares missions au CNES. J'en garde de très bons souvenirs et me réjouis de pouvoir travailler dans ce cadre à l'issue de ma thèse. Plus largement, merci aussi aux membres de la *Science Team* et de l'*Algorithm Development Team* de SWOT. Grâce à eux, j'ai pu mieux comprendre le contexte dans lequel s'inscrit ma thèse. Merci en particulier à Hervé Yésou du SERTIT, qui m'a fourni des données pour tester mes algorithmes.

En dehors du cadre de la thèse, je souhaiterais également remercier tous ceux qui, rencontrés il y a longtemps à Châteaubriant ou plus récemment à Cachan, ont été à mes côtés pendant ces trois années.

Enfin, je tiens à remercier ma famille. Merci en particulier à mes parents, à ma sœur et à ma grand-mère Marie pour leur soutien et leurs encouragements.

[5] De très bons souvenirs là aussi. J'espère avoir l'occasion de recroiser les organisateurs et les participants

[6] Thanks you for the time and energy you have put in planning our SAR meetings.

[7] Qui a apporté d'autres genres de radar à nos SAR meetings et sa bonne humeur à l'équipe.

[8] Je suis content de voir que la relève en imagerie SAR est assurée !

[9] Et pourtant, j'étais un des rares doctorants qui ne pratiquaient pas de magie noire avec TensorFlow.

[10] Merci pour ton aide et ta patience lorsque j'étais bloqué avec le cluster du CNES !

[11] Qui, comme moi, se souviendront sans doutes longtemps de la réunion ADT à Paris.

[12] Merci en particulier pour ton aide pour la préparation des données Sentinel-1 !

Cette thèse a été financée par le Centre National d'Études Spatiales et l'entreprise CS Group France

Contents

Chapter 1

Introduction

Un résumé détaillé en français est disponible en annexe E page 179.

1.1 Context

Surface water in lakes and rivers account for respectively 0.006% and 0.0002% of the total amount of water on Earth and occupy only 3.7% and 0.6% of its non-glaciated land surfaces (Verpoorter et al., 2014; Allen and Pavelsky, 2018). Yet they play a key role for some of the major challenges facing humanity. Water is a critical resource for agriculture, domestic, and industrial use. Its demand is growing more rapidly than the world population and its shortage represents a threat to the health and food safety of more than half of the world population (FAO, 2020, 2020; Mekonnen and Hoekstra, 2016). Surface water also represents a direct threat, with a growing proportion of the world population at risk of floods (Tellman et al., 2021), and also an indirect threat as lakes and dams are major factors of malaria transmission (Kibret et al., 2021). Rivers and lakes are also key for the production of renewable electricity as the flexibility of hydropower is critical for the stability of electric grids. Yet, globally half of the hydropower's economically viable potential is still untapped and a better understanding of water systems could help with the difficult task of assessing their opportunities and risks (IEA, 2021).

Furthermore, in the context of global warming caused by greenhouse gas emissions, major changes in the water cycle are expected (IPCC, 2021). For this reason, our knowledge of water systems has to be constantly updated.

Lakes and rivers also play a role that has long been underestimated on greenhouse gas fluxes, especially CO_2, CH_4, and NO_2 (Khalil and Rasmussen, 1983; Beaulieu et al., 2011) and an accurate global evaluation of their surfaces is needed for an accurate assessment of these fluxes.

To address these challenges, and also encourage the settlement of water-related conflicts through negotiation (Bernauer and Böhmelt, 2020), efficient monitoring and management of freshwater resources is needed, yet only a few developed countries have achieved it (UN-Water, 2021).

In this context, improving hydrological models and data collection is crucial, and spaceborne remote sensing is vital, as it enables data acquisition at the global scale. Spaceborne data

have been used for hydrology applications since they became available. Optical and Synthetic Aperture Radar (SAR) images have enabled global mapping of water bodies, for example with Landsat since 1972 and ERS since 1991, and more recently with Sentinel-1 and Sentinel-2.
SAR and optical instruments are and will remain precious assets for hydrology but lack the information on water elevation that is needed to evaluate river discharge and lake storage change. Nadir altimeters such as the Poseidon series (TOPEX then Jason satellites) or Sentinel-3's SRAL (Sar Radar ALtimeter) provide water elevation information on a rough spatial scale and have revolutionized oceanography. However, their low spatial resolution limits the hydrological application to continental waterbodies.

The SWOT (Surface Water and Ocean Topography) Mission aims at breaking through this limit with a swath altimeter that will measure water elevation on a high resolution spatial 2D grid by performing interferometric operations on a pair of Ka-band SAR images acquired simultaneously in near-nadir configuration. SWOT will also provide data for oceanographic applications that will not be mentioned in this thesis.

SWOT mission is a collaboration between the French Centre National d'Études Spatiales (CNES) and NASA's Jet Propulsion Laboratory (JPL), with contributions from the Canadian Space Agency (CSA) and United Kingdom Space Agency (UKSA). The tasks are distributed between these spatial agencies, with support from local academic and industrial partners. CNES is responsible for water detection algorithms, which are a key step in the processing of SWOT's High Rate (HR) data. The baseline method for detection of water was developed during Sylvain Lobry's Ph.D. work at Télécom Paris (Lobry, 2017) in collaboration with CNES, and is being calibrated and validated on simulated images with the support of CS Group France.

The SWOT mission relies on a sensor technology that has never been used on a spaceborne instrument before. In this context, there is still uncertainties on the performances in terms of signal-to-noise ratio (SNR), with consequences on the ability to detect water surfaces and to retrieve their elevation. The goal of this work is to provide robust alternative water detection methods that will be able to detect the water surfaces even in situations where the baseline method fails. This approach is part of a risk mitigation action for the SWOT mission, with potential application to other SAR sensors as well. To this end, we have focused on three strategies in order to make the detection more robust, which will be covered in detail later:

- Use of exogenous guiding information
- Multitemporal and multi-sensor approaches
- Use of a prior denoising step

Beyond the SWOT mission, our work on water detection in SAR images can be used for other SAR sensors such as Sentinel-1 that will remain useful for hydrology, as a complement to the SWOT mission: SWOT will not make other sensors unnecessary for hydrology applications. They will rather be combined in data and services centre, such as CNES's incoming HYSOPE II data hub. Our work may also be useful in the context of the WiSA interferometric altimetry mission (concept) that may replace SWOT after its end of life, but with a less complex and potentially noisier sensor.

The methods we propose are designed to be run on single-channel images, such as SWOT coherent power images. For dual-polarized Sentinel-1 images, they can be run on either a single-polarized channel or the pixel-wise geometric mean of the channels, as presented in appendix A.

1.2 Contributions

The main contributions of this PhD study are linked to three main applications:

1. A method for robust detection of narrow rivers guided by an exogenous river centerline database.
2. Some improved approaches for the detection of lakes on SAR images with exogenous information, multitemporal data, or the combination of information from optical and SAR images.
3. In the perspective of using a denoising step before water detection, we worked on the use of the temporal geometric mean of a SAR image time series.

1.2.1 Narrow river detection guided by an exogenous database

Contribution (1): Narrow river detection guided by exogenous database

This contribution is presented in our article (Gasnier et al., 2021b) and in chapter 5

Given the availability of global database of rivers, we proposed a framework for the guided detection of narrow rivers whose detection is the most crucial, robust to both a low signal-to-noise ratio in the images and errors in the prior information.
This method relies on three steps:

1. A novel linear structure detector (contribution 1-A)
2. A repositioning step for the river centerline based on exogenous information and the linear structure detector response, and a least-cost-path algorithm
3. A segmentation around the centerline using a new Conditional Random Field (CRF) model (contribution 1-B)

Contribution (1-A): Linear structures detection

This contribution is presented in our articles (Gasnier et al., 2021c; Gasnier et al., 2021b) and in section 5.2.2

We propose a new linear structure detector for SAR images, based on the Generalized Likelihood Ratio (GLR). It compares the likelihood of a patch considering the estimated reflectivities under two hypotheses:

- $\mathcal{H}_0$: there is no linear structure
- $\mathcal{H}_1$: there is a linear structure

The actual application of this method for SAR images is made possible by the optimized approach we present. We compare this method to (Tupin et al., 1998)'s linear structure detector and show that it results in fewer false positive detections.

Contribution (1-B): CRF segmentation model for narrow river detection around a centerline

This contribution is presented in our article (Gasnier et al., 2021b) and in section 5.2.4

For the last step of the proposed narrow river detection framework, we need to detect the water surface around the retrieved centerline, without prior knowledge of the reflectivities of land and water. We proposed a new CRF model for that task. This new CRF model is based on 4 terms:

- A data term that is different for the water class and for the land class. For water, it derives from a statistical model using a water reflectivity estimated from the centerline pixels. For land, as we have no information on the underlying reflectivities, the data term is spatially uniform and its value has been chosen so that it does not introduce any bias towards one of the classes.
- A centerline term that prevents the centerline from being classified as land.
- An asymmetric CRF-based regularization term, that is designed to take into account the fact that water is brighter than land (for SWOT) or darker than land (for conventional SAR systems such as Sentinel-1).
- A gradient flux term that compensates for the consequences of the regularization in some situations where it is needed.

1.2.2 Lake segmentation approach derived from GrabCuts

To improve the detection of lakes, especially with small areas or irregular shapes, we proposed to utilize *a priori* information in the form of a rough bounding polygon for each lake and to combine multitemporal and multi-sensor data. To use a rough bounding polygon as an input,

an approach derived from the GrabCut method (Rother et al., 2004) is very suited.

For this task, we proposed three methods:

1. An adaptation of the GrabCut method for water detection in single-date SAR images. This method does not use any prior knowledge on the water and land reflectivities, but takes a rough bounding polygon as input.

2. A multitemporal extension of the previous method.

3. A multitemporal and multi-sensor method that processes a combined time series of both SAR and optical images. As opposed to the two previous methods, it does not take a bounding polygon as input but instead uses prior statistical models for the water and land classes in both SAR and optical images.

Contribution (2): Guided multitemporal and multi-sensor approaches for the detection of lakes in SAR images

This contribution is presented in chapter 6

We adapted the GrabCut method, originally proposed for the detection of any given object in natural RGB images to water detection in SAR images. To this end, we adapted the mixture models used to the statistics of SAR images and added a flux term to the segmentation model in order to encourage the detection of structures darker (for Sentinel-1 or other conventional SAR sensors) or brighter (for SWOT) than their background.

Then, we adapted this model to SAR time series and added a temporal regularization term that improves the localization of the borders of the lake for a given date while preserving the temporal changes of the lake surface.

Finally, we proposed an unguided segmentation approach for combined time series of SAR and optical images. The statistical distributions for land and water that are taken as an input can be determined for SAR images by using the previous method, and for optical images, by an external clustering approach such as (Cordeiro et al., 2021). This unguided approach could also be embedded into a combined SAR and optical multitemporal GrabCut method.

1.2.3 Denoising the temporal geometric mean to facilitate water detection

A prior denoising step can facilitate the detection of water structures. We presented at the 2021 IGARSS conference (Gasnier et al., 2021a) preliminary results on the improvement of narrow river detection through a denoising preprocessing step on single-date SAR images. This denoising step could benefit from the temporal information in the time series, and the properties of the geometric mean make it appropriate for this purpose, for example through denoising by ratio (Zhao et al., 2019).

Contribution (3): Statistical properties and denoising of the temporal geometric mean of SAR images

This contribution is presented in our article (Gasnier et al., 2021d) and in section 7

We studied the statistical properties of the temporal geometric mean and compared it with its arithmetic counterpart. We showed that the geometric mean prevails in situations with transient bright outliers and with a temporally fluctuating reflectivity.
As no closed-form expression is available for the distribution of the temporal geometric mean, we proposed a numerical approach for its estimation and used it within a variational framework to denoise the geometric mean image.
We used the denoised temporal geometric mean in some applications such as change detection or denoising-by-ratio and showed that it improves the denoising result in certain situations compared to the arithmetic mean.

The distribution of our contributions between the three strategies is outlined in **Figure** 1.1

1.3 Organization of the manuscript

This manuscript is divided into three parts. In the first part, we present some background information on the images in chapter 2, on the context of the water detection problem in chapter 3, and on the methodological foundations of the methods we propose in chapter 4. The second part is dedicated to the proposed methods. First, the framework for the guided detection of narrow rivers is presented in chapter 5, including our proposed linear structure detector in section 5.2.2 and the CRF model in section 5.2.4. Then we present a segmentation approach derived from the GrabCut methods in chapter 6, first on a single image in section 6.1.3, and then on a temporal stack of images in section 6.3. A non-guided segmentation method combining SAR and optical images is then presented in section 6.4. Finally, chapter 7 presents our work on the denoising of the temporal geometric mean of SAR images, and chapter 8 draws a conclusion and introduces some perspectives.

Appendix A introduces the combination between VV and VH channels that we used for some of our experiments. Appendix B compares the Graph Cut models we use with some reference methods. Appendix C summarizes the notations used in this documents and appendix D presents the publications done during this PhD work.

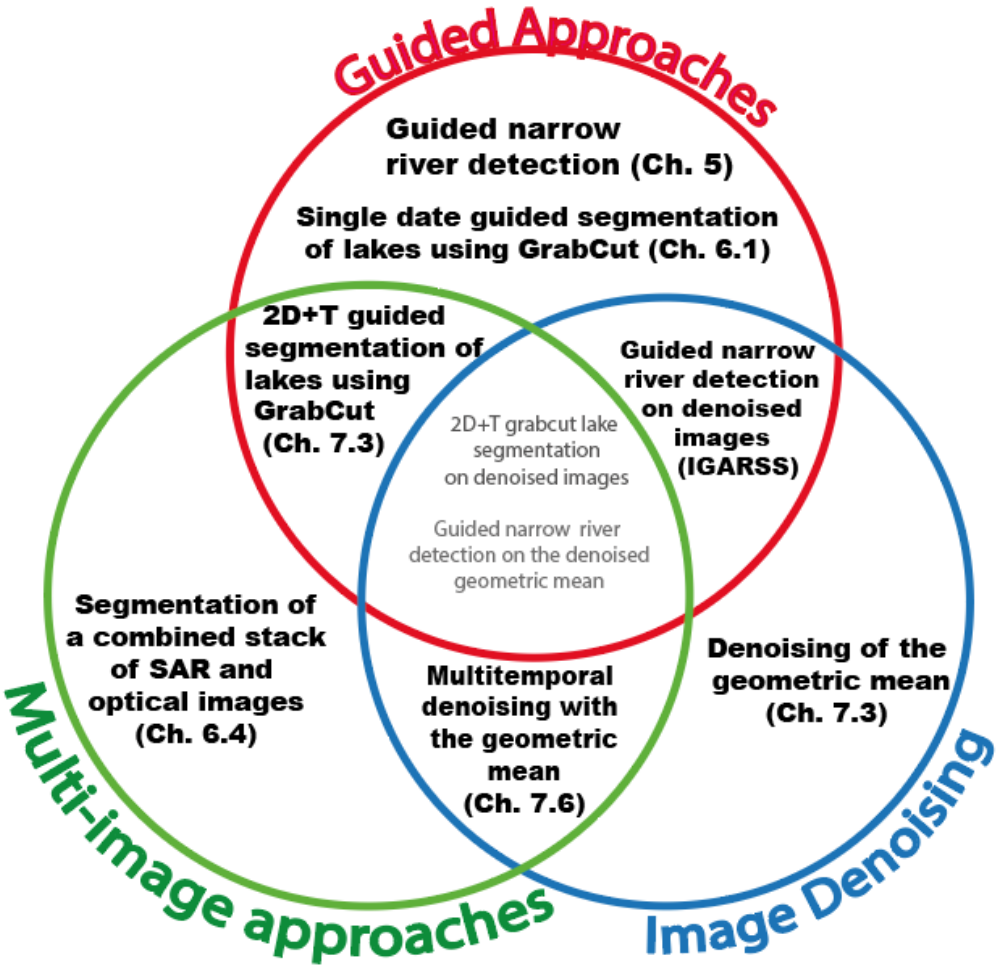


Figure 1.1 – Overview of some of the proposed approaches, associated with the corresponding chapter. The guided narrow river detection in denoised images is not presented in this document, but has been presented at the IGARSS 2021 conference. Approaches belonging to the perspectives, but not yet published nor presented, are in small gray font.

Part I

Background on SAR remote sensing and water surface monitoring with SAR images

Chapter 2

SAR images

This chapter provides background information on the acquisition of Synthetic Aperture Radar (SAR) images and their characteristics, and in particular for the Surface Water and Ocean Topography (SWOT) and Sentinel-1 images that are used in the work presented in the following chapters.

Organization of this chapter Section 2.1 briefly presents the acquisition of SAR images and introduces the notion of polarimetric images. Then, the main statistical distributions for SAR images are presented and their limitations are introduced. Section 2.2 and 2.3 present the specific properties of SWOT and Sentinel-1 satellites.

2.1 Physics and statistics of SAR images

2.1.1 Acquisition and synthesis of SAR images

SAR is an imaging technique in which a moving antenna (airborne or space-borne) emits a frequency modulated electromagnetic pulse (or "chirp") in the direction of an area of the ground and then measures the wave that is backscattered by the targets on this area. The measured echos for each pulse forms a 1D signal that forms one line in the two-dimensional complex raw image. These pulses are repeated multiple times and the concatenated lines form the raw image.

The L1 single look complex (SLC) image is then reconstructed from the raw data using a method that is presented in (Cumming and Wong, 2005) and involves range and azimuth compression of the raw image with the range and the azimuth reference functions.

Note that in some cases, there can be one emitting and two receiving antennas, thus creating two images of the same area with different paths, which can be used for interferometric applications. This operating mode is called bistatic and limits the power consumption.

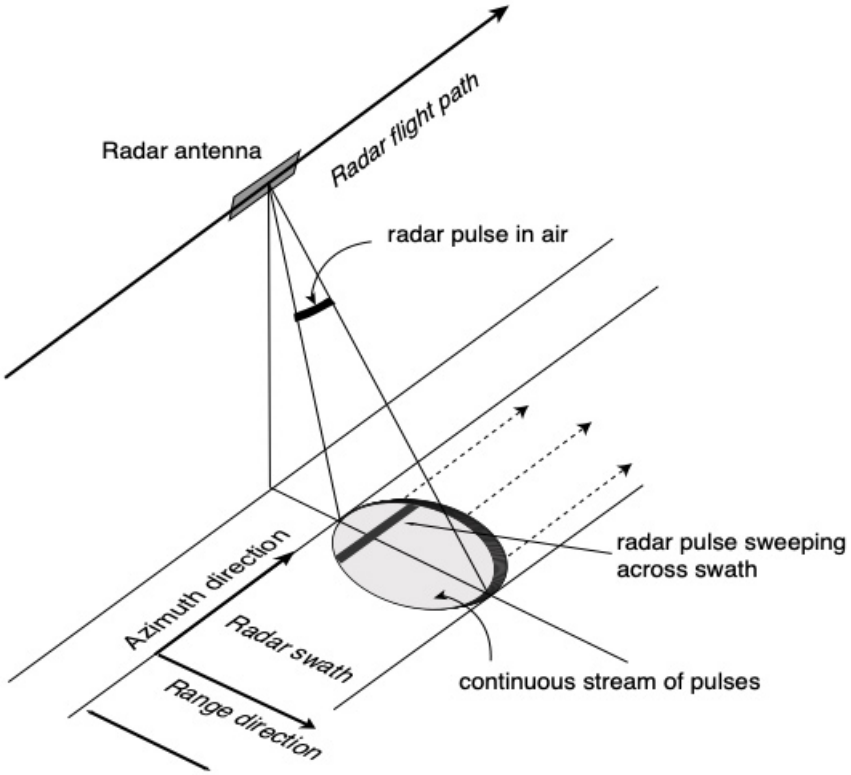


Figure 2.1 – Schematic view of the principle of an imaging radar. The vertical direction is called nadir. The direction of the radar pulse is called line of sight (LOS). The angle between the nadir and the LOS is called the incidence angle. Courtesy NASA/JPL-Caltech.

Multiple microwave frequency bands can be used depending on the objectives of the measurement as the scattering properties of the targets and the atmosphere depend on the frequency. The frequency bands presented below are defined by IEEE nomenclature (IEEE Std 521-2002, 2020) :

Table 2.1 – Operating frequency bands used by some SAR instruments

Band	Frequency	Vacuum wavelength	Example
C	4 – 8 GHz	7.5 – 3.75 cm	Sentinel-1, RADARSAT-2
X	8.0 – 12.0 GHz	3.75 – 2.5 cm	TerraSAR-X
Ka	27–40 GHz	11.1–7.5 mm	SWOT

2.1.1.1 Polarization

The polarization of an electromagnetic wave is defined by the direction of oscillation of its electric field vector: vertical, horizontal, or even elliptical in the presence of both vertically and horizontally polarized fields. Radar antennas can be designed to emit and receive an electromagnetic pulse according to a given polarization. This allows measuring the dependence

of the backscattering with the polarization and the ability of some targets to change the polarization of the backscattered wave, to extract more information from the scene.

For this reason, several SAR instruments emit and receive according to two polarization directions. This leads to up to 4 images for fully-polarized sensors such as RADARSAT-2:

1. The co-polarized image I_{VV} emitted in vertical polarization and received in vertical polarization.
2. The cross-polarized image I_{VH} emitted in vertical polarization and received in horizontal polarization.
3. The co-polarized image I_{HH} emitted in horizontal polarization and received in horizontal polarization.
4. The cross-polarized image I_{HV} emitted in horizontal polarization and received in vertical polarization.

In contrast, Sentinel-1 data are only dual-polarized: as only the vertically polarized pulse is emitted, only I_{VV} and I_{VH} are available. For SWOT, the images are single-polarized: the right swath is acquired only in VV polarization and the left swath only in HH.

The exploitation of polarimetric information in SAR images is a whole research field and its comprehensive presentation is beyond the scope of this thesis. The reader can find more information on this subject in (Lee et al., 2004), (Lee and Pottier, 2009) or (Moreira et al., 2013).

2.1.2 Interferometric processing

Interferometric approaches (Goldstein et al., 1988; Li et al., 2007) use the phase information of two or more SAR SLC images to estimate the difference of distance between the scatterers of a pixel and the positions of the antennas. One application of interferometry is the measurement of the ground elevation by using two images acquired at the same time by two antennas separated by a distance B called baseline. This method has been used to build accurate global digital elevation models (DEM), first with the Shuttle Radar Topography Mission (SRTM) mission (Farr et al., 2007) in which C-band and X-band SAR instruments with two antennas each have been mounted on the Endeavour space shuttle with a 60m baseline. More recently, the TanDEM-X constellation of two X-band SAR satellite with a variable baseline (Krieger et al., 2007) is used to produce a worlwide DEM.

KaRIn operates according to the same principle with its two antennas on a mast with a 10m baseline.

The phase difference $\Phi(i,j)$ between the two signals $\boldsymbol{z}_1(i,j)$ and $\boldsymbol{z}_2(i,j)$ acquired on each antenna of the same pixel (i,j) on the ground, is defined as:

$$\Phi(i,j) = \arg(\boldsymbol{z}_1(i,j).\boldsymbol{z}_2^*(i,j)), \tag{2.1}$$

and is related to the geometry of the acquisition (Fjørtoft et al., 2014):

$$\Phi(i,j) = \frac{2\pi}{\lambda_r}(r_2 - r_1), \tag{2.2}$$

where r_1 and r_2 are the distances between the scatterers within the pixel (i,j) and the antennas 1 and 2. $r_2 - r_1$ is related to the baseline, to the incidence angle θ_i, and to the position of the pixel. This allows to estimate the water elevation using a triangulation approach that is presented in (Fjørtoft et al., 2014). Because the phase $\Phi(i,j)$ is measured modulo 2π, the elevation for every pixels (x,y) can only be estimated modulo an altitude of ambiguity when using only the phase measured in (x,y). The determination of the actual elevation combines phase information for other pixels (phase unwrapping) with exogenous data (see (Desroches et al., 2016) for more detail on the elevation estimation approach for SWOT HR mode).

Another way of computing interferometric measurements uses two or more images collected by the same spaceborne instrument at different times. It can be used both for estimating ground deformations or elevation mapping (Massonnet and Feigl, 1998).

2.1.3 Statistical modeling of SAR images

This section introduces the fully developed speckle model, which is the most frequently used model for SAR images modeling, and the image statistics that derive from it.

The limits of this model will be presented along with alternative models for some specific situations.

2.1.3.1 Fully developed speckle

In most situations, the speckle phenomenon in SAR images can be modeled through the fully developed speckle model described in (Goodman, 1976). It allows for a simple expression of the measured amplitude in a pixel corresponding to a resolution cell on the ground that contains a large number of individual scatterers. Each of these scatterers backscatters an elementary wave with an amplitude a_n and phase ϕ_n.

For this model, four assumptions are made (Goodman, 1976):

1. There is a large number of elementary contributions.
2. The amplitude a_n and the phase ϕ_n of the n-th elementary contribution are statistically independent of each other.
3. The amplitude and the phase of any elementary contribution are independent of the amplitude and the phase of every other elementary contribution.
4. The phase ϕ_k of any elementary contribution follows a uniform distribution in the range (-π, π).

The last assumption requires the surface to be sufficiently rough compared to the wavelength.

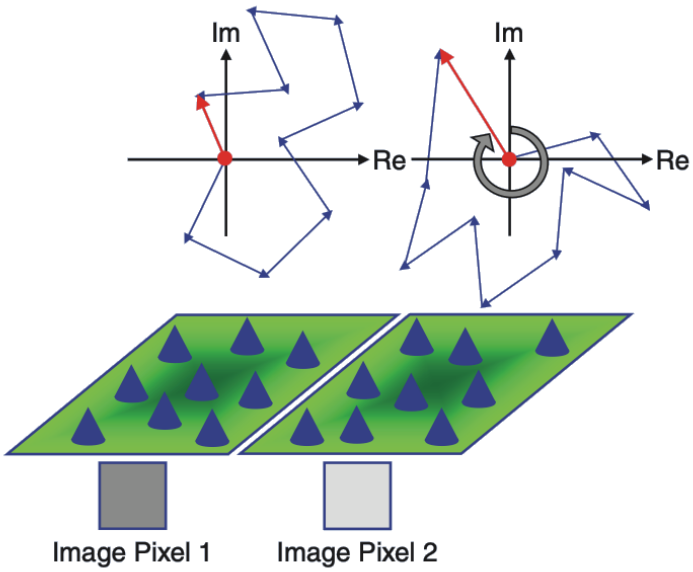


Figure 2.2 – Illustration of the speckle phenomenon. Under Goodman hypotheses (Goodman, 1976), the contributions of the elementary scatterers in the resolution cells are independent and identically distributed. Figure taken from (Moreira et al., 2013). © 2013 IEEE

Under these assumptions, the real and imaginary parts of the resulting vectorial sum z of the elementary contributions are independent and follow the same centered Gaussian distribution. As a consequence, the phase follows a uniform distribution and the intensity $\boldsymbol{I} = \Im(z)^2 + \Re(z)^2$ follows an exponential distribution:

$$p(I) = \frac{1}{R} e^{\frac{-I}{R}}, \tag{2.3}$$

where R is the reflectivity. The reflectivity increases with the radar backscattering coefficient σ_0, which is a parameter that depends on the physical interactions between the pulse and the surface, and the area of the surface. Note that the mean and the standard deviation of the distribution are both equal to R. For those reasons, speckle can be considered as a multiplicative noise,

The amplitude $A = \sqrt{I}$ follows a Rayleigh distribution:

$$p(A) = \frac{2A}{R} e^{\frac{-A^2}{R}}. \tag{2.4}$$

Single Look Complex images Single Look Complex images follow a Gaussian distributions for the real and imaginary parts, and their intensity follows an exponential distribution.

Multi-Look images In order to reduce the level of speckle noise, SAR images are often multi-looked: the pixel values are averaged along one or both spatial directions or along the temporal direction and the sampling is adapted accordingly.
Assuming uncorrelated, fully developed speckle, and a homogeneous reflectivity R, the intensity of a multi-looked intensity image follows a Gamma distribution:

$$\mathrm{p}(I|R) = \frac{L^L I^{L-1}}{\Gamma(L) R^L} \exp\left(-L\frac{I}{R}\right) \tag{2.5}$$

The variable L is the equivalent number of independent looks (ENL). The case $L = 1$ corresponds to single-look images.

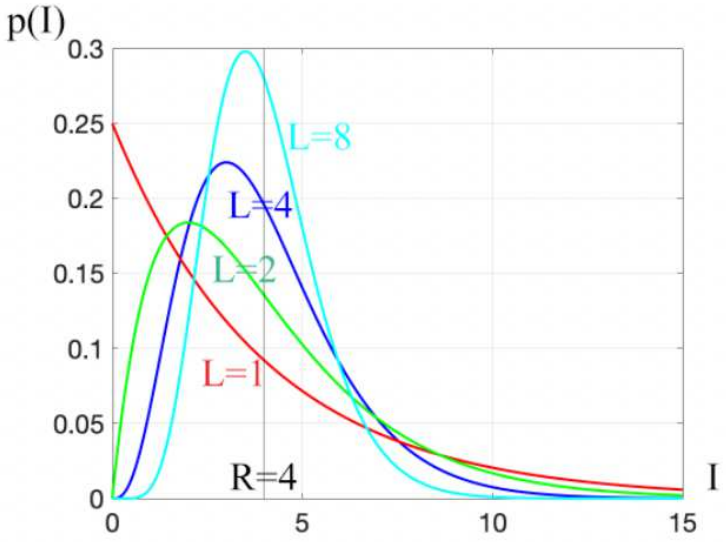


Figure 2.3 – Gamma distributions of the intensity I for a reflectivity $R = 4$ and multiple values of L. The red graph for $L = 1$ is a particular case of Gamma distribution that corresponds to an exponential distribution.

In practice, due to the spatial correlation of the speckle, the actual ENL can be lower than the averaging factor.

The mean μ_I of the intensity distribution over a homogeneous area (uniform value of R) is $\mu_I = R$ and its standard deviation σ_I is $\sigma_I = \frac{R}{\sqrt{L}}$. The ratio $\gamma_I = \frac{\sigma_I}{\mu_I}$ between the mean and the standard deviation is then equal to $\gamma_I = \frac{1}{\sqrt{L}}$.

The amplitude A follows a Nakagami distribution:

$$\mathrm{p}(A|R) = \frac{2L^L A^{2L-1}}{\Gamma(L) R^L} \exp\left(-L\frac{A^2}{R}\right) \tag{2.6}$$

Logarithmically transformed images SAR images are sometimes processed after a logarithmic transformation that converts the multiplicative speckle noise into additive noise. The statistics of logarithmically-transformed SAR images with fully developed speckle have been described by (Hua Xie et al., 2002).

The logarithmically transformed intensity follows a Fisher-Tippett distribution (also known as generalized extreme value (GEV) distribution) defined by the following expression, where

$y = \log(I)$ and $x = \log(R)$:

$$\mathrm{p}(y|x) = \frac{L^L}{\Gamma(L)} e^{L(y-x)} \exp(-Le^{y-x}). \tag{2.7}$$

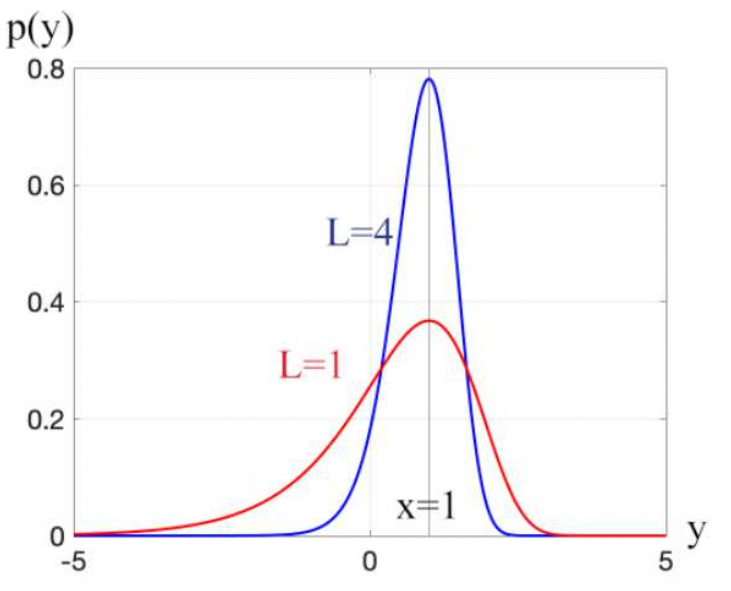


Figure 2.4 – Fisher-Tippett distributions for a log-reflectivity $x = 1$ and two values of L.

The expectation and the variance of $\boldsymbol{y}$ are:

$$\mathbb{E}[\boldsymbol{y}] = x - \log(L) + \psi(L) \tag{2.8}$$

$$\mathrm{Var}[\boldsymbol{y}] = \psi'(L) \tag{2.9}$$

$$\tag{2.10}$$

where $\psi(.)$ is the digamma function and $\psi'(.)$ is its derivative called trigamma function [1], as defined for example in (Olver et al., 2010).

Note that the mean $\mathbb{E}[y]$ is a biased estimator of $x = \log(R)$.

Even though the Fisher-Tippett is heavy-tailed, a Gaussian approximation for the Fisher-Tippett distribution of the logarithmically-transformed SAR images can be used if the value of L is high enough ($L = 4$ is enough for most applications).

$$\mathrm{p}(y|\mu_{GA}) \simeq \frac{1}{\sigma_{GA}\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{y-\mu_{GA}}{\sigma_{GA}}\right)^2}, \tag{2.11}$$

where $\sigma_{GA} = \sqrt{\psi'(L)}$, and $\mu_{GA} = x - \log(L) + \psi(L)$ to match the standard deviation and the expected value of the previous distribution.

Figure 2.5 compares the distribution of the log-transformed speckle for $L = 4.4$ (corresponding to Sentinel-1 full-resolution GRD images acquired in IW mode) with its Gaussian approximation. The two distributions are very similar.

[1] Other sources use the notation $\psi(1, .)$ for the trigamma function

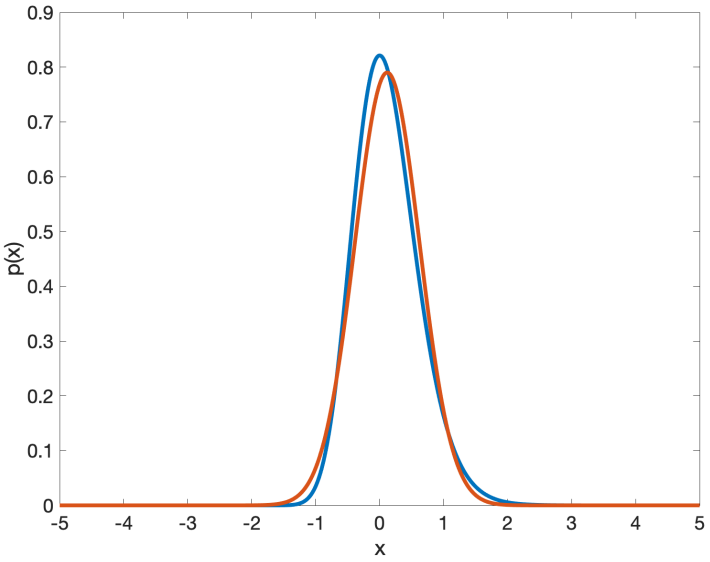


Figure 2.5 – Fisher-Tippett distribution (orange) and its Gaussian approximation (blue) for L=4.4.

A fully developed speckle model is not adapted to certain situations where some of its conditions are not satisfied as described in the following sections.

2.1.3.2 Water surfaces specific characteristics

Ripples can appear at the surface of water bodies under the influence of wind and water currents. The geometric characteristics of these riddles can have regularities that are incompatible with Goodman's fully developed speckle conditions (Migliaccio et al., 2007).

In particular, for a certain wavelength λ_S of the ripples on the water surface, the path difference for two successive scatterers is a multiple of the radar wavelength λ_r, as presented in **Figure** 2.6. The phase of all the elementary contributions are then not independent as assumed by Goodman's model but are identical. This causes the interferences to be constructive and result in a very strong scattering. The phenomenon is called Bragg resonance (Valenzuela, 1978) and has been widely studied in the context of SAR oceanography and maritime oil spill detection (Delignon et al., 1992; Garello et al., 1993; Brekke and Solberg, 2005; Weinberg and Tran, 2018).

For a wavefront perpendicular to the slant direction, the wavelength of a horizontal periodic structure that creates a Bragg resonance is given by:

$$\lambda_S = \frac{\lambda_r}{2 \sin \theta_i}, \tag{2.12}$$

with θ_i the incidence angle and λ_r the central wavelength of the SAR sensor. For Sentinel-1, with $\lambda_r = 5.6$ cm, the resonance wavelength is close to 5 cm. For SWOT, the resonance wavelength goes from 7cm in far range to 29cm in near range.

If the ripples wavefront is not perpendicular to the slant direction but makes an angle θ_v, the resonance wavelength is multiplied by $\sin(\theta_v)$:

$$\lambda_S = \frac{\lambda_r}{2 \sin \theta_i} \sin \theta_v \tag{2.13}$$

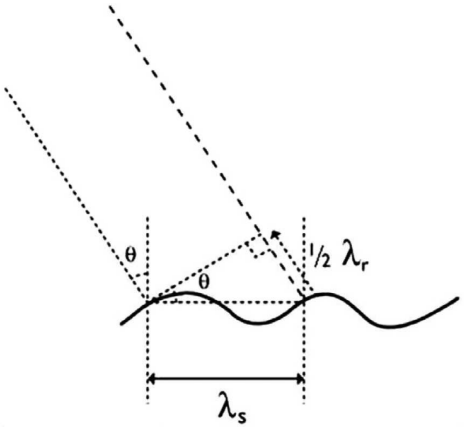


Figure 2.6 – Schematic view of the principle of Bragg phenomenon. For ripples distant of λ_S, the difference in optical path is a multiple of the radar wavelength λ_r and there is constructive interference between the elementary contributions. On this view, the ripples wavefront is perpendicular to the slant direction ($\theta_v = 0$). Figure taken from (Chaturvedi et al., 2019). License CC BY-NC-ND

As the Bragg phenomenon does not cause a change in the polarization of the returned pulse, the reflectivity for these areas is much higher in co-polarized images (VV and HH) than in cross-polarized images (VH and HV). This is illustrated in **Figure** 2.7.

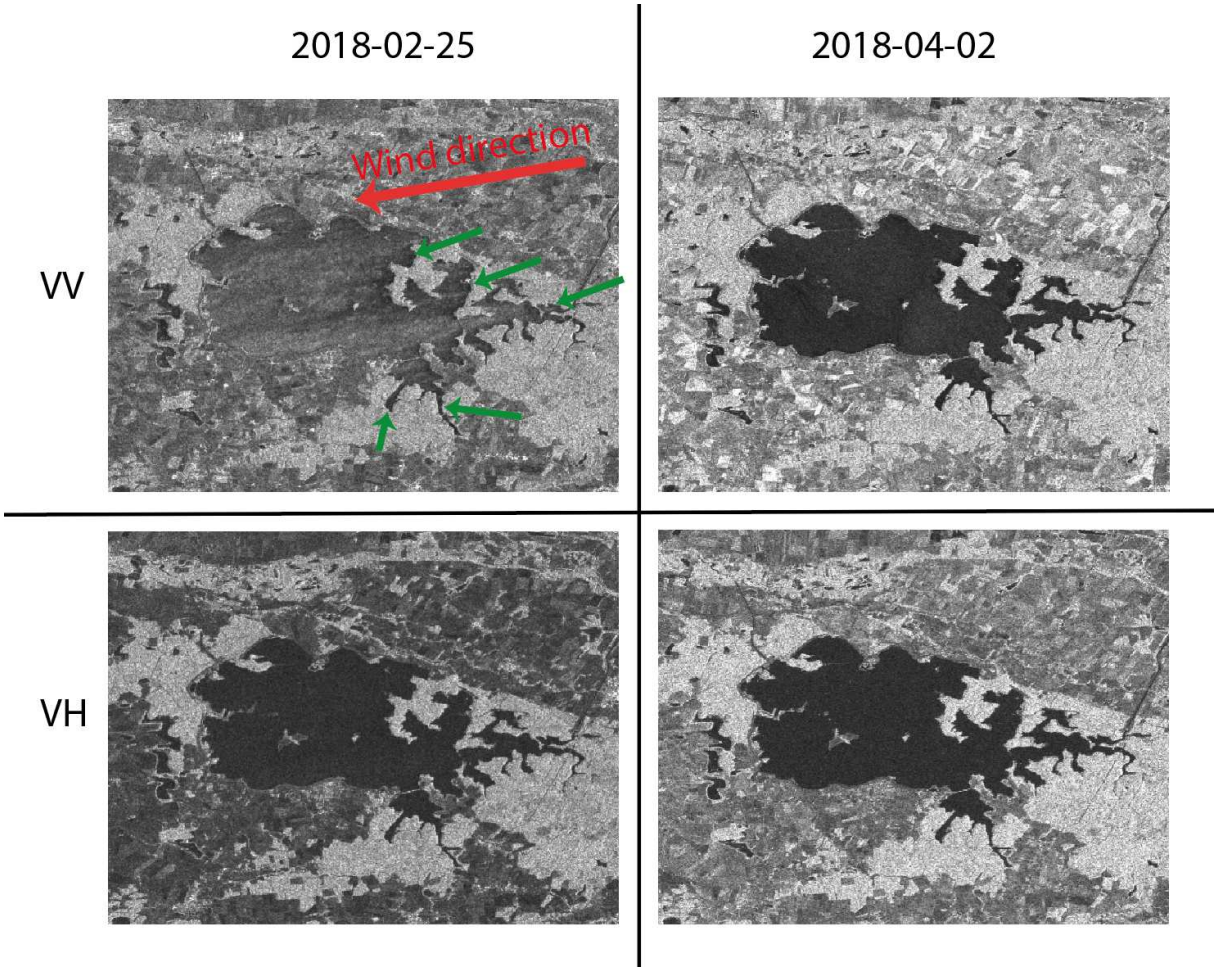


Figure 2.7 – Illustration of the consequence of Bragg phenomenon on the contrast between water and land in Sentinel-1 GRD VV images of Lake Der. Images on the left have been acquired on 2018-02-25, and images on the right on 2018-04-02, with a very similar water level. Top two images are VV polarized, bottom two images are VH polarized. On the 2018-02-25 images, the size and the orientation of the ripples caused a Bragg resonance. This phenomenon is weaker in areas sheltered from the wind (green arrows).

Multiple statistical distributions (such as K or Weibull distributions) have been proposed to model the backscattered intensity, which depends on the distribution of the waves geometry (Sun et al., 2018; Weinberg and Tran, 2018). An exhaustive overview of these distributions is beyond the scope of this section, but it can be noted that the Bragg scattering effects that can be encountered in SAR images of continental water bodies can lead to areas of the image in which the noise does not follow a gamma distribution. In particular, in these areas, the standard deviation of the intensities might be below what is expected from the equivalent number of looks of the image.

2.1.3.3 Specular reflection

The most obvious situation is the case of specular reflection on a very smooth surface at the scale of the wavelength. It can happen on very smooth, mill-pond like water or oil spills. For larger

wavelengths than Ka-band (like Sentinel-1's C-band), it can also happen on tarmac surfaces. On horizontal surfaces, the pulse is backscattered away from the sensor. As a consequence, such areas in the image appear very dark, even sometimes to the point where the only signal corresponds to thermal noise.

2.1.3.4 Double bounce scattering

Besides surface and volume scattering, a much stronger return can be seen in the presence of corner-like geometrical structures in the resolution cell. In that case, most of the signal is backscattered towards the antenna, which results in a very strong scatterer that overpowers all the other scatterers in the resolution cell in such a way that Goodman's conditions presented in section 2.1.3.1 are not met. Thereby, the distribution of the intensity in such pixels does not follow a Gamma distribution.

Instead, a modeling based on a Nakagami-Rice distribution (Nicolas and Tupin, 2020; Tison et al., 2004) should be used:

$$p(I|R) = \frac{1}{R} \exp\left(-\frac{I + s_c^2}{R}\right) I_0\left(2\frac{\sqrt{I s_c^2}}{R}\right), \tag{2.14}$$

where I_0 is the first-kind modified Bessel function with order zero (see (Olver et al., 2010)) and s_c is the cross-section of the corner reflector. However, for applications that do not require accurate modeling of such strong scatterers, this specific distribution may not be needed.

These structures are most often man-made (buildings, boats, towers, bridges,...). See for example the very bright pixels in image 2.8, that correspond to buildings.

Figure 2.8 – Sentinel-1 VH SLC image of Nantes Airport in the south of Nantes (France). The airport buildings appear as very bright pixels due to the double bounce effect. In contrast, the tarmac and the Loire River in the upper part of the image appear dark because of specular reflection.

Summary: Statistics of SAR images

In this section, we presented the distribution that derived from Goodman's fully-developed speckle model:

1. Single Look Complex: both real and imaginary parts follow a Gaussian distribution.
2. Single look intensity image $\boldsymbol{I}$ follows an exponential distribution.
3. Multi-looked intensity image $\boldsymbol{I}$ follows a Gamma distribution.
4. $\boldsymbol{y} = \log(\boldsymbol{I})$ follows a Fisher-Tippett distribution.

We also presented some cases when Goodman's model does not apply:

1. Specular reflexion: very flat surfaces are likely to act as a mirror and backscatter the signal away from the sensor through specular reflection, causing very low intensities (except for surfaces perpendicular to the line of sight).
2. Multiple bounce: corner structures yield a very strong scatter whose intensity follows a Rice distribution.
3. Bragg phenomenon: on water, Bragg phenomenon can create backscattering regimes that do not meet the Goodman conditions. The distribution of intensity may therefore not follow a Gamma distribution. Several other statistical distributions have been proposed.

2.2 The SWOT mission

The Surface Water and Ocean Topography (SWOT) mission is currently being developed by NASA and the Centre National d'Études Spatiales (CNES) since 2007. Its goal is to provide scientists with the first detailed global survey of the Earth's surface water, both for oceanographic and hydrological applications.

While nadir altimeters have been in use since 1992 with TOPEX/Poseidon (Zieger et al., 1991) followed by the JASON series (Carayon et al., 2002) and several others, SWOT marks a breakthrough in altimetry as its SAR interferometric altimeter provides a much higher spatial resolution than previous nadir altimeters.

This will enable the study of sub-mesoscale phenomenon in oceanography and new applications in hydrology including monitoring of river discharge and lake storage change.

As the launch of the SWOT satellite is projected in late 2022, actual SWOT images have not been available for the work presented here and the experiments have been conducted on simulated SWOT images. These images have been obtained with the Jet Propulsion Laboratory (JPL) HR simulator (JPL D-79123, 2014) using an accurate digital earth model, a water mask, and models for the ground reflectivity and the water roughness.
The SWOT platform carries several instruments:

- The Ka-band Radar Interferometer KaRIn (see next section).
- A dual-frequency (C and Ku band) nadir altimeter, similar to the Poseidon altimeter used for the Jason missions.
- An Advanced Microwave Radiometer (AMR) similar to that of Jason, which measures the microwave transmission property of the troposphere to calibrate the wet tropospheric delay correction over ocean.
- Several instruments (DORIS, GPS,...) that accurately measure the orbit of the satellite.

2.2.1 The KaRIn sensor

The main sensor in the SWOT mission is the interferometer KaRIn which is a Ka band bistatic near-nadir SAR interferometer. It consists of two parallel 5m long antennas that are 10m apart (see Fig. 2.9). It acquires images alternatively from two swaths: one looking to the right and one looking to the left (see Fig. 2.10).

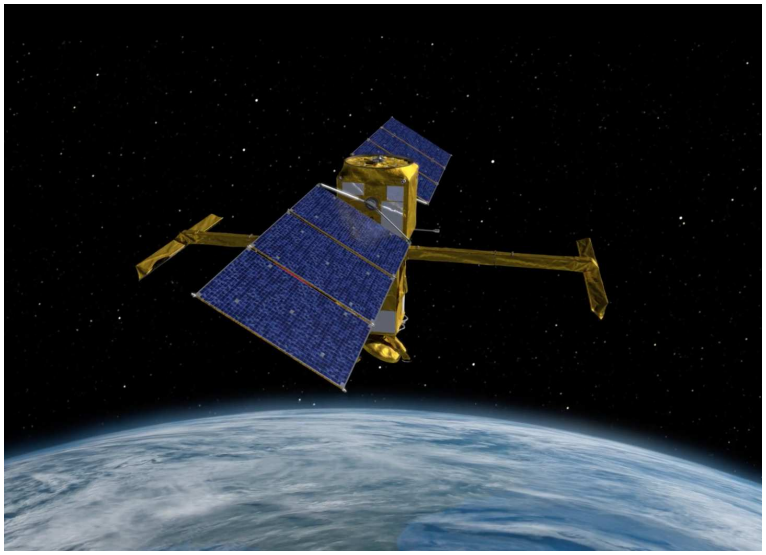


Figure 2.9 – Artist's representation of the SWOT satellite in operation. The two parallel antennas of KaRIn sensor are visible. Courtesy NASA/JPL-Caltech.

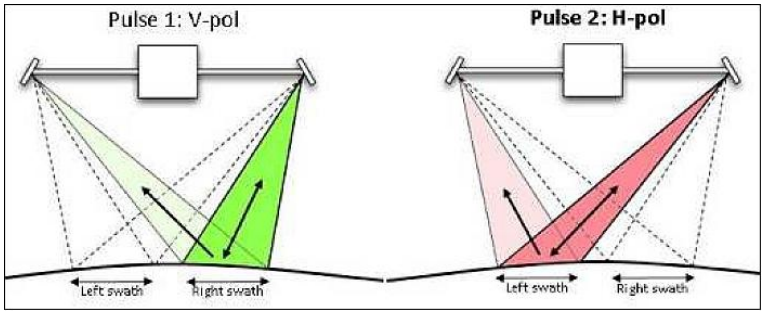


Figure 2.10 – Schematic view of KaRIn's alternating acquisition over two swaths. Courtesy NASA/JPL-Caltech

KaRIn is run in two different modes depending on the situation:

- Low Rate (LR) dedicated to oceanography, with $\sim$ 500m resolution and on-board processing to reduce the data transmission rate to around 0.2 Mbps.
- High Rate (HR) mode for continental surfaces. The on-board pre-processing is limited to an azimuth multilooking of factor of $\sim$ 2 which leads to a very high data rate (around 300 Mbps) (Fjørtoft et al., 2010) but keeps the resolution high enough for hydrological applications.

This work will focus on the HR mode, that is the only one for which water detection is relevant.

In HR mode, the azimuth resolution is uniform ($\sim$ 5m) but the range resolution varies from $\sim$10m in far-range to $\sim$70m in near range within the same image because of the variation of the incidence angle along the swath.

2.2.2 Specific characteristics of near-nadir Ka-band KaRIn SAR images

Besides its two antennas and its bistatic operational mode, KaRIn sensor is a breakthrough compared to previous SAR instruments because of its Ka wavelength and its incidence angle.

In contrast to most SAR sensors (Sentinel-1, TerraSAR-X,...) that typically have an incidence angle of 30° to 45°, the incidence angle of KaRIn is only 0.6° to 4.1°. This has major consequences on backscattering properties over water and significantly increases the layover phenomenon (Fjørtoft et al., 2014).

Another consequence of the small incidence angle is a large difference in range sampling within the image because it is proportional to $\frac{1}{\sin(\theta_i)}$, where θ_i is the incidence angle. Indeed, the range resolution goes from 10m in far-range to 70m in near-range.

As KaRIn operates in Ka-band, its wavelength (8.6mm) is much smaller than for other SAR spaceborne sensors, and its interactions with the ground are not as well known as for C-band or X-band SAR sensors. A more detailed overview can be found in (Fjørtoft et al., 2014):

1. As more surfaces appear rough, specular reflections appear less frequently. However, specular reflection can still happen for very smooth water, in low wind situations (see next section).
2. The penetration into vegetation, soil, snow,... is very weak.
3. The acquisitions are relatively sensitive to the tropospheric conditions (rain).

Several experiments (see for example (Fjørtoft et al., 2014)) have studied Ka-band with a low incidence angle on multiple targets (land, water,...). These results have been confirmed by airborne measurements on rivers using JPL's AirSWOT instrument (see (Altenau et al., 2017)).

These experiments have confirmed that the reflectivity of water surfaces strongly depends on their roughness and thus on the wind conditions. The reflectivity of water is generally much higher than the reflectivity of the land (see table 2.2.2 and (Fjørtoft et al., 2014)), except for some very smooth and flat, surfaces can yield a very strong backscattering (tarmac, muddy fields, river banks, roads) if the surface is perpendicular to the radar line of sight within a very small margin (0.1°).

Surface	Backscattering coefficient σ_0
Rock, soil, vegetation,...	-5dB to -10dB for θ_i from 0° to 5°
Tarmac	17dB to 0dB θ_i from 0° to 5°
Water ($\sim$ 1 m/s wind speed)	23–25 dB for $\theta_i = 0°$, 17–20 dB $\theta_i = 2°$, 5–10 dB for $\theta_i = 4°$
Water ($\sim$ 2 m/s wind speed)	14-20dB in the 0°-4° range
Water ($\sim$ 4 m/s wind speed)	10-15dB in the 0°-4° range
Water ($\sim$ 10 m/s wind speed)	$\sim$10dB in the 0°-4° range

The thermal noise level of KaRIn is very high, with a noise equivalent σ_0 of 0 dB, while it is as low as -20dB to -30dB for most spaceborne SAR instruments (Sentinel-1,...). The very low signal-to-noise ratio is due to power constraints. As a consequence, the measured signal for the darkest pixels (with σ_0 below 0dB such as vegetation) does not correspond to the backscattered signal but to thermal noise. The resulting statistical distribution in the SAR image nevertheless corresponds to that of fully developed speckle.

It should be noted that while these elements provide information about the performances that can be expected for KaRIn instrument, the actual performances can not be precisely known before the actual SWOT data become available after launch. As a consequence, the processing algorithms have to be robust to image quality below the expectations: While the nominal land-to-water contrast is of 10dB-20dB, algorithms have to be tested for weaker contrast (as low as 3dB in the Worst Case scenario).

Coherent power The water detection processing is done on the coherent power image instead of the two individual SLC images. The coherent power is a combination of the two phase-flattened SLC images (Lobry et al., 2019):

$$\boldsymbol{z}_c = \frac{\boldsymbol{z}_1 + \boldsymbol{z}_2 \cdot \exp(\phi_{ref} \cdot i)}{2} \tag{2.15}$$

where $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are the complex SLC image of the two antennas, ϕ_{ref} is the flattening reference phase and $\boldsymbol{z}_c$ is called the coherent average of $\boldsymbol{z}_1$ and $\boldsymbol{z}_1$. Here, $i = \sqrt{-1}$

$$\boldsymbol{I} = 2\boldsymbol{z}_c\boldsymbol{z}_c^* \tag{2.16}$$

Here, $\boldsymbol{I}$ is the coherent power of the two SLC images. For the sake of simplicity, the notation $\boldsymbol{I}$ will always designate the coherent power image when referring to SWOT images. Its statistical properties have been described in (Lobry et al., 2019). It can be modeled as an intensity SAR image with an equivalent number of looks $L = 4$, and follows the corresponding Gamma distribution.

2.2.2.1 Dark Water

Even if KaRIn uses a shorter wavelength than other spaceborne SAR instruments and a smaller incidence angle than other SAR instruments, specular reflection will still occur over water in very low wind situations, leading to very weak signal, i.e. comparable to that of surrounding land surfaces and vegetation.

In these situations, the low reflectivity can make the water detection operation difficult and the signal be so weak that the interferometric processing that is needed to compute water elevation (see 2.1.2) can be impossible for the affected pixels.

2.2.2.2 Layover

The layover is a phenomenon that occurs when the terrain slope exceeds the incidence angle. All SAR instruments are concerned by the effect, but the small incidence angle of KaRIn makes it dramatically more exposed to it (see Fig. 2.11).

This effect could be an issue for water detection, as a large area of terrain may fall within a single resolution cell, resulting in a bright pixel and thereby false detection of water. Besides this land/land layover, land areas can also fall within the same resolution cell than water, as in the figure (land/water layover), which is not an issue for water detection but for the height estimation.

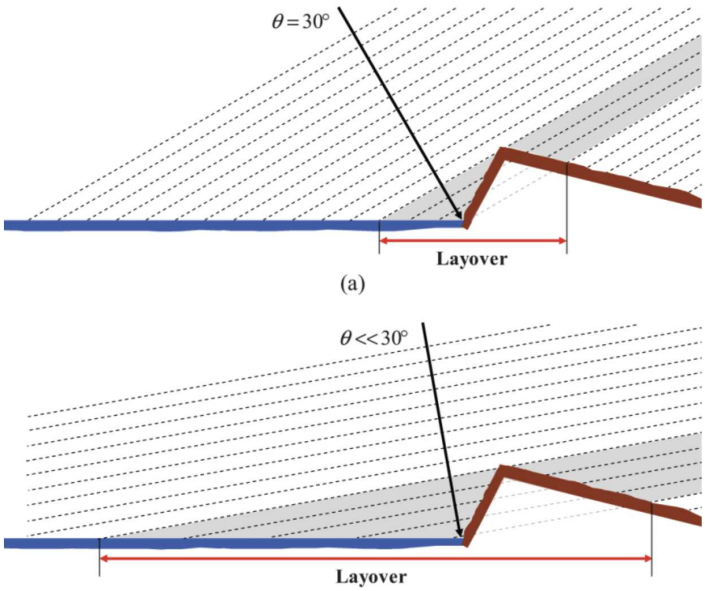


Figure 2.11 – Schematic view of land/water phenomenon. Figure taken from (Fjørtoft et al., 2014). © 2014 IEEE

2.2.2.3 Localization uncertainty

Another consequence of the very low angle of incidence is that the estimated position of water bodies in SWOT HR images based on prior knowledge of the water elevation may be inaccurate. A difference of 1 meter between the prior and the actual water elevation can lead to a change in the river position of up to 95m in near range.

This uncertainty is a cause of inaccuracy in the projection of exogenous information in the SWOT SAR image, and algorithms that rely on such exogenous information have to be robust to these projection errors. Note that the interferometric processing will provide more accurate water elevation and geolocation, but water detection takes place prior to this.

Summary: Particular characteristics of SWOT

SWOT's KaRIn instrument constitutes a breakthrough compared to nadir altimeters and traditional spaceborne SAR instruments. Its images will have significant differences compared to those of other sensors and their characteristics are not known perfectly before the launch. These specific properties have to be taken into account for the design of the processing algorithms. The algorithms should also be robust to image quality below expectations.

2.3 Sentinel-1

Sentinel-1[2] is a constellation of C-band dual-polarized SAR satellites, currently Sentinel-1A and Sentinel-1B launched in 2014 and 2016 as a part of the Copernicus program. The two satellites provide a combined 6-12 days revisit period globally and the images are freely available.

The images are available in dual (VV and VH) polarization in several acquisition modes and product levels.

Most continental Sentinel-1 images, except for small isolated islands or disasters that are acquired in Strip Map (SM) mode, are acquired in Interferometric Wide Swath (IW) mode.

IW mode images are acquired in three separate swaths (IW1, IW2, and IW3) using the Terrain Observation with Progressive Scanning SAR (TOPSAR) technique that makes the image characteristics (SNR) across the whole swath more homogeneous than the conventional ScanSAR approach.

For the IW mode, Sentinel-1 images are distributed as raw, Single Look Complex (SLC), and Ground Range Detected (GRD) products. GRD products are multilooked by a factor 5 in the range direction and projected from slant to ground geometry onto an ellipsoid that is corrected using the terrain height, which varies in azimuth but is constant in range.

This projection does not provide a very accurate pixel localization, hence the need for an orthorectification step to register Sentinel-1 images with other data. A comprehensive description of Sentinel-1 products characteristics is presented in (S1-RS-MDA-52-7440, 2016). SLC images have a resolution of 2.7m to 3.5m in range by 22m in azimuth and a 2.3m by 14.1m pixel size. GRD images have a 20x22m mid-range resolution and a 10x10m pixel spacing. Their equivalent number of looks L is 4.4.[3]

Because of their acquisition mode, Sentinel-1 images are very different from SWOT images. In particular, the radiometric characteristic is almost opposite with SWOT being generally dark

[2] https://sentinels.copernicus.eu/en/web/sentinel/missions/sentinel-1

[3] These values are the average of the three swaths. There can be slight differences between IW1, IW2, and IW3

on land areas and Sentinel-1 being bright, while water is generally bright for SWOT images and dark for Sentinel-1.

While the use of Sentinel-1 images is not planned as a support for water detection within the SWOT mission processing chain, but rather in hydrological data hubs, and is relevant as a part of the SWOT downstream program operations to promote the use of remote sensing data for hydrology more generally.

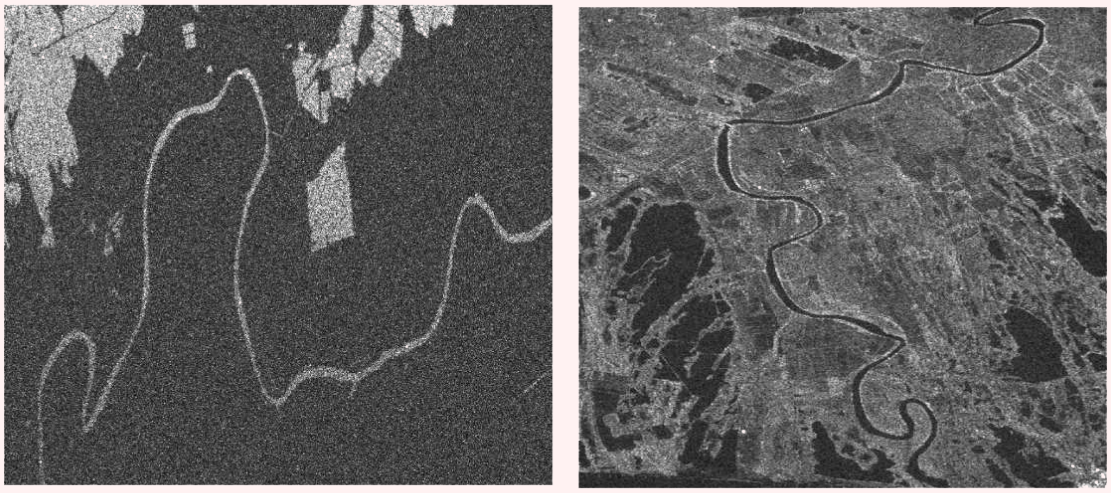
Moreover, as similar approaches can be applied to SWOT and Sentinel-1, existing Sentinel-1 images enable the test of water detection algorithms on actual images, instead of only simulated data.

Summary: SWOT and Sentinel-1 comparison

Some of the main differences between SWOT and Sentinel-1 can be summarized as below:

	SWOT	Sentinel-1 IW
Wavelength (cm)	0.86	5.55
Incidence angle	0.6°-4.1°	20°-46°
Polarization	Single (HH or VV)	Dual (VV and VH)
Azimuth resolution	20m (coherent power)	22m (SLC and GRD)
Range resolution	10 - 70m (coherent power)	2.7 - 3.5m (SLC) 20m(GRD)
ENL	$L = 4$ (coherent power)	$L = 1$ (SLC) $L = 4.4$ (GRD)
Land	Dark	Dark to very bright
Rippled water	Bright	Dark to quite bright (Bragg)
Calm water	Dark	Very dark

While not directly related to water detection in SWOT processing, studies on water detection in Sentinel-1 images are still relevant as a part of the SWOT mission.



SWOT simulated image (RADAR geometry) and Sentinel-1 VH GRD image (ground geometry). Both image show downstream Petit Rhone River (Camargue, France). The river is bright in SWOT image and dark in Sentinel-1.

Chapter 3

SAR water detection and hydrological prior

This chapter provides background and context for the water detection within the SWOT data processing chain.

Organization of this chapter Some water detection approaches for SAR images are introduced in section 3.1, with a focus on the existing baseline water detection method for SWOT. Then, section 3.2 outlines the context of water detection in SWOT data processing. The requirements related to water detection are then introduced as they define the performances that have to be achieved by the water detection algorithm. Then, section 3.3 presents some background about existing hydrological data.

3.1 Water detection in SAR images

While simple pixel-based thresholding is used for water detection in composite optical images (NDWI or MNDWI, see 3.3.1), the same cannot be directly done for SAR images. Indeed, the strong speckle noise makes direct pixel-based segmentation unusable. Instead, other types of approaches can be used to efficiently deal with the problem:

1. Segmentation after denoising.
2. Segmentation with a regularization.

Both region-based and edge-based segmentation method can be adapted to SAR images using these approaches.

3.1.1 Pre-processing of the SAR data

To apply pixel-based segmentation to SAR images, the simplest approach is to first apply a denoising step. A denoised image enables the use of pixel-based detection methods, which can be very simple. For example, a maximum likelihood or even maximum a posteriori estimation of the pixel classification boils down to a simple thresholding operation. However, more

sophisticated approaches are often needed to compute the threshold.

Multiple water detection algorithms for SAR images based on this principle have been proposed, for example, by Liu and Jezek (Liu and Jezek, 2004) or Cazals et al. (Cazals et al., 2016), who applied thresholding on a denoised SAR image. For (Liu and Jezek, 2004), the denoising method is a Lee filter (Lee, 1981; Lee, 1983) followed by anisotropic diffusion (Perona and Malik, 1990; Sohn and Jezek, 1999) that enhances the edges of the image. Then, a Levenberg-Marquardt algorithm (Moré, 1978) is applied on the histogram to fit the parameters of a bimodal Gaussian distribution from which the threshold is computed. (Cazals et al., 2016) uses a Perona-Malik filter (Perona and Malik, 1990) for denoising followed by hysteresis thresholding on a temporal time series of Sentinel-1 images. (Martinis et al., 2009) proposes a multi-scale approach that applies a minimum error (Kittler and Illingworth, 1986) threshold on TerraSAR-X images denoised with a Gamma-MAP algorithm (Lopes et al., 1990). (Huang et al., 2018) also uses a preliminary denoising step but then detect water using a random forest approach.

Some of these methods combine the thresholding segmentation with a subsequent post-processing step that can eliminates small objects or isolated pixels using mathematical morphology (Liu and Jezek, 2004).

An overview of denoising algorithms published before 2013 can be found in (Argenti et al., 2013). Note that more recent denoising methods could be used as well. These new approaches generally feature major improvements in the preservation of small details, the level of remaining noise, and in the bias induced by the denoising. Such denoising methods include patch-based approaches (Deledalle et al., 2014), or the variational framework MuLoG (Deledalle et al., 2017a) and its multi-temporal extension RABASAR (Zhao et al., 2019), which enable the denoising of SAR images by plugging in Gaussian denoisers such as BM3D (Dabov et al., 2007). More recently, denoising approaches based on deep learning such as SAR2SAR (Dalsasso et al., 2021) or Speckle2Void (Molini et al., 2021) have been proposed with good denoising results.

3.1.2 Segmentation with regularization

A direct segmentation of noisy SAR images is possible provided that the segmentation methods contain regularization terms that limit the consequences of the speckle noise.

One way to regularize is to use a Markov Random Fields (MRF) approach, as introduced by (Geman and Geman, 1984) and (Greig et al., 1989) for binary segmentation. MRF methods are presented from a theoretical point of view in section 4.1. They have been used extensively for the segmentation of SAR images: see for example (Rignot and Chellappa, 1992; Fjortoft et al., 2003) or (Deng and Clausi, 2005; Pelizzari and Bioucas-Dias, 2007) for ice and oil spills detection on water. SWOT baseline water detection method of Lobry et al. (Lobry et al., 2019), presented in section 3.2.3 also uses an adapted MRF approach.

A variation of the MRF is the Conditional Random Field approaches (Lafferty et al., 2001). The conditional random fields enable an adaptation of the regularization to take into account the edges of the image.

Another way to apply a regularization is with active contour approaches such as level sets, for example (Silveira and Heleno, 2009) that is based on a simplified version of (Chan and Vese,

2001)'s "active contour without edges". This active contour approach minimizes an energy that takes into account two terms:

1. Two data terms that ensure the homogeneity of the inside and the outside regions.
2. A regularization term that limits the length of the boundary

The original Chan-Vese approach also included a third term that limits the area of the inside region. These methods use strong regularization priors to avoid speckle-induced false detection, which impairs the detection of small structures such as narrow rivers or small lakes.

More recently, methods based on deep learning approaches have been proposed to detect water directly on SAR images in a classification framework (Isikdogan et al., 2017; Nemni et al., 2020)].

3.1.3 Edge-based approaches

Unlike (Chan and Vese, 2001)'s "active contour without edges", older "snakes" active contours methods rely on the image edges to ensure fidelity of the boundaries to the image (Kass et al., 1988). Their goal is to ensure that the boundary of the detected regions corresponds to edges that are present in the image while adding constraints on the length and the smoothness of the boundary. Such approaches on SAR images require an edge detector that is robust to their statistics. For example, (Li et al., 2011) used on snakes based on the ROEWA (Fjortoft et al., 1998) method to compute the gradient and detect the edges. More recent approaches for edge detection in SAR images include a contrario methods (Liu et al., 2020a) or deep learning methods (Liu et al., 2020b).

3.1.4 Specific approaches for river detection

Specific approaches for river detection have also been proposed such as the one developed by Cao et al. (Cao et al., 2011) for SWOT images that combines the elementary segments detected with a linear structure detector. (Valero et al., 2010) proposed an approach based on mathematical morphology for road detection in high-resolution images. This approach has been adapted for rivers and automated using machine learning by Klemenjak et al. (Klemenjak et al., 2012). Sghaier et al. (Sghaier et al., 2017) combines it with structural feature sets. Other river-specific approaches based on active contours have also been proposed, such as (Han and Wu, 2018).

3.1.5 Water detection guided by ancillary data

The combination of SAR images with exogenous data has been used to improve water detection. For example, topographic information provided by a Digital Earth Model (DEM) are often used (such as in (D'Addabbo et al., 2016)) to improve water detection and limit the false detection rate. Other approaches such as (Hong et al., 2015) even add optical images to the combination.

Summary: Water detection in SAR images

The strong speckle noise in SAR images requires appropriate methods for water detection. Multiple approaches can be used:

- Segmentation after a preliminary denoising.
- Segmentation using a regularization such as MRF approaches.

As the spatial regularization or even some denoising approaches tend to impair the detection of small structures, specific approaches have been proposed for rivers.

The use of ancillary data such as topographic information can help with water detection.

3.2 SWOT processing and products

3.2.1 SWOT processing pipeline

This section focuses on the processing of High Rate data, from the raw data acquired by the sensor to the final hydrological products. The final products are the pixel cloud and more specific vector products for lakes and rivers. The pixel cloud is a sparse array (or cloud) of geolocated points in 3-D space with many pieces of information per point. Points correspond to water pixels and their vertical position corresponds to the estimated water height.

The processing chain is presented here as it currently exists and does not involve any of the risk mitigation approaches that are proposed in this thesis.

Three stages may be distinguished in the processing flow:

1. Processing of the two L0 raw images to create the two L1 SLC images.
2. Processing of the two L1 SLC images to create the L2 pixel cloud.
3. Processing of the pixel cloud to create river-specific and lake-specific products, using the RiverObs and LOCNES processing chains.

The water detection takes place in the second stage and is performed on a coherent power image that is a combination of the two phase-flattened L1 SLC images. It can be modeled as an intensity SAR image with an equivalent number of looks of $L = 4$.

The water detection step is followed by a water height estimation step to determine the water pixel's vertical position in the pixel cloud as well as its horizontal geolocation (latitude, longitude). The water height estimation and geolocation processing are presented in (Desroches et al., 2016). These two steps happen before the processing of the final hydrological products and are crucial for their quality. Pixels that are not detected as water because they are not bright enough but likely to be water based on prior data (Pekel masks,...) are flagged as dark water. These pixels are taken into account for the water surface determination but not used for

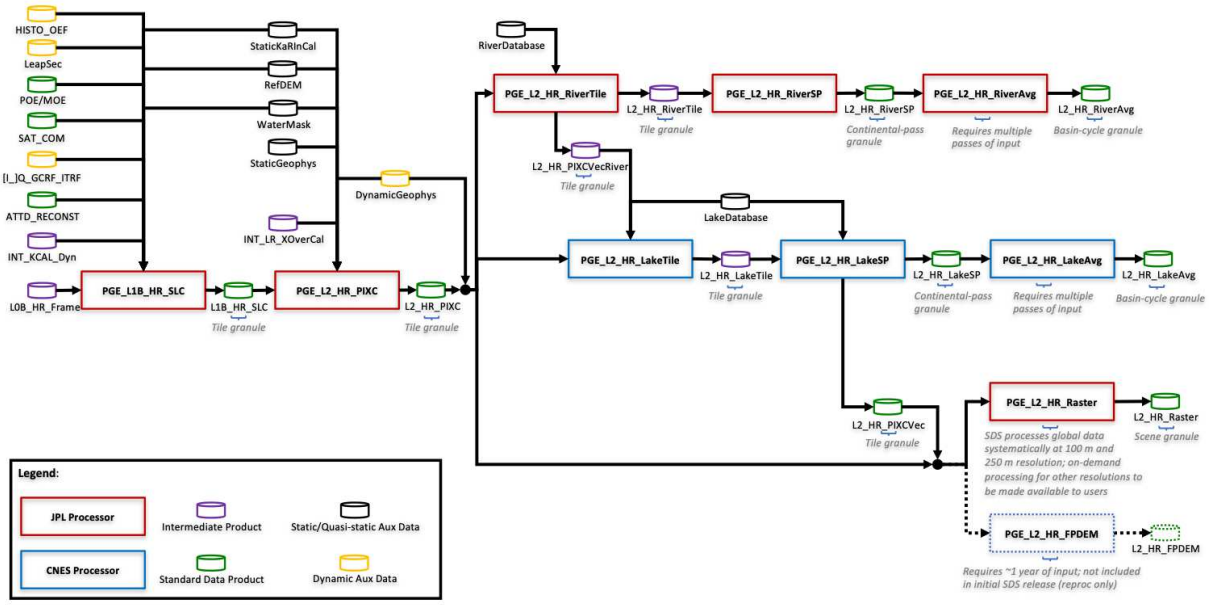


Figure 3.1 – SWOT High Rate Algorithm Flow. The water detection algorithms take place in the "PGE_L2_HR_PIXC" processor. Courtesy NASA/JPL-Caltech, CNES

height determination through phase unwrapping.

It can be noted that the baseline water detection step does not use a prior water mask nor a Digital Earth Model in order to prevent false detection even though they are available, and does not use the prior river database nor the prior lake database that are used in later processing. As these databases are available within the processing framework, their use for water detection as a part of a risk mitigation approach would require limited changes in the global architecture. In contrast, using multi-temporal processing or combining SWOT data with other sensors would be much more difficult.

The methods we present in this dissertation are part of a risk mitigation approach to improve the detection for situations in case the baseline algorithm does not achieve the required performances.

The river-specific and lake-specific processing are done only for objects that correspond to an item in the SWORD river database or the lake database.

3.2.2 SWOT water surface area requirements

The required performances for the SWOT mission are defined in the SWOT Science Requirements Document (JPL-D-61923-rev-B, 2018).

The main hydrologic science requirements are defined as follows in this document:

> 1. To provide a global inventory of all terrestrial surface water bodies whose surface area exceeds $(250\text{m})^2$[1] (goal: $(100\text{m})^2$, threshold: 1km^2 (lakes, reser-

[1]The reference surface area of the waterbodies are given in the document by the dimension of its sides

Reference	Type	Maximum error (1 σ)	Water bodies area	River width
2.8.2.a	Baseline Requirement	15%	>(250 m)^2	>100 m
2.8.2.b	Threshold Requirement	15%	>1 km^2	>170 m
2.8.2.c	Goal	25%	(100 m)^2 - (250 m)^2	50 m-100 m

voirs, wetlands) and rivers whose width exceeds 100 m (goal: 50 m, threshold: 170 m).

2. To measure the global storage change in terrestrial surface water bodies at sub-monthly, seasonal, and annual time scales.
3. To estimate the global change in river discharge at sub-monthly, seasonal, and annual time scales.

Concerning the evaluation and characterization of the performances, the requirement 2.6.3.a of the document states that the required performance will be evaluated using non-vegetated water bodies with an area greater than (250m)^2 and rivers of width greater than 100 m and with negligible layover effects. The characterization is done on smaller water bodies ((100m)^2) and narrower rivers (50 m).

Additional requirements concern the accuracy of the estimated areas of the detected water surfaces. The 1σ relative error has to be below a certain limit:

With the area being defined as the non-vegetated surface area and the river width being defied as the average river width over a reach whose length exceeds 10 km. The 1σ relative error value is so that 1/3 of the observed errors are worse and 2/3 are better.

These requirements set the performances that the water detection algorithms will have to achieve. In particular, the surface accuracy requirements (2.8.2) directly concerns the output of water detection. If the actual quality of the images turns out to be worse than expected, risk mitigation approaches for water detection might be needed to achieve these requirements.

3.2.3 Baseline water detection method for SWOT

The baseline water detection method for the SWOT processing uses the MRF-based approach presented in (Lobry et al., 2019) and in Sylvain Lobry's Ph.D. dissertation (Lobry, 2017). This method has been developed to be robust to the spatial variations in class parameters induced by the non-uniform antenna pattern that cannot be compensated. To this end, it starts from an initial set of two parameter maps (one for land and one for water), known from exogenous data (antenna pattern, expected water and land reflectivities), and a SWOT coherent power SAR image, and alternatively runs two steps a given number of times (see Fig. 3.2):

- Perform classification based on the current set of parameter maps.
- Estimation of a new set of parameter maps using the current classification.

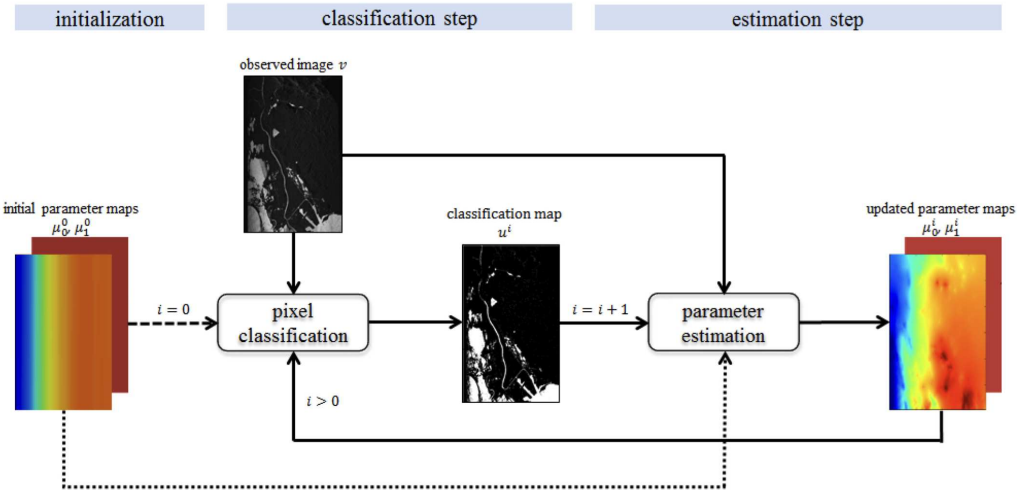


Figure 3.2 – The baseline approach alternatively estimates a classification map u^i from the parameter map μ_1^{i-1} for water and μ_0^{i-1} for land and then two parameters map: μ_1^i and μ_0^i from the classification map u^i. The iterations are repeated for a given number of times $i \in \{0, 1, ...n\}$. Note that these notations are specific to this figure. Figure taken from (Lobry et al., 2019). © 2019 IEEE

For both the classification and the parameter estimation steps, an MRF-based regularization is used.

Summary: SWOT water detection

Water detection is a crucial step in the processing of SWOT data. The required performances are defined in the SWOT Mission Science Requirements document and specific approaches might be needed to achieve them.
The rivers and lakes whose detection is needed are reported in databases used in further processing. These databases might be useful to guide water detection in a robust approach.

3.3 Prior water masks and databases

This section introduces some existing hydrological databases and water masks that are relevant in the context of the SWOT mission, and could potentially be used to facilitate and guide automated water detection.

3.3.1 Water masks

Following the increasing availability of spatial data, for example with the opening of Landsat optical images time series in 2008 (Woodcock et al., 2008; Wulder et al., 2012) and more recently the Sentinel data from the Copernicus program, several works have sought to provide a global overview of the occurrence and temporal dynamic of continental water bodies.

For these studies, the water detection usually involves thresholding of water and vegetation indices (NDVI, NDWI (McFeeters, 1996), and MNDWI(Xu, 2006; Feng et al., 2016)) and takes the topography into account. An example of a state-of-the-art approach for water surface extraction in optical images can be found in (Cordeiro et al., 2021).

Following (Yamazaki et al., 2015), the currently most used water mask has been proposed in (Pekel et al., 2016) and provides several 30-meter resolution rasters. Pekel masks give information on water occurrence (between 0 and 100%) such as in **Figure** 3.3, water seasonality, recurrence, transitions in the presence of water, and maximum water extent. These masks are based on 30-year time series of Landsat optical images.

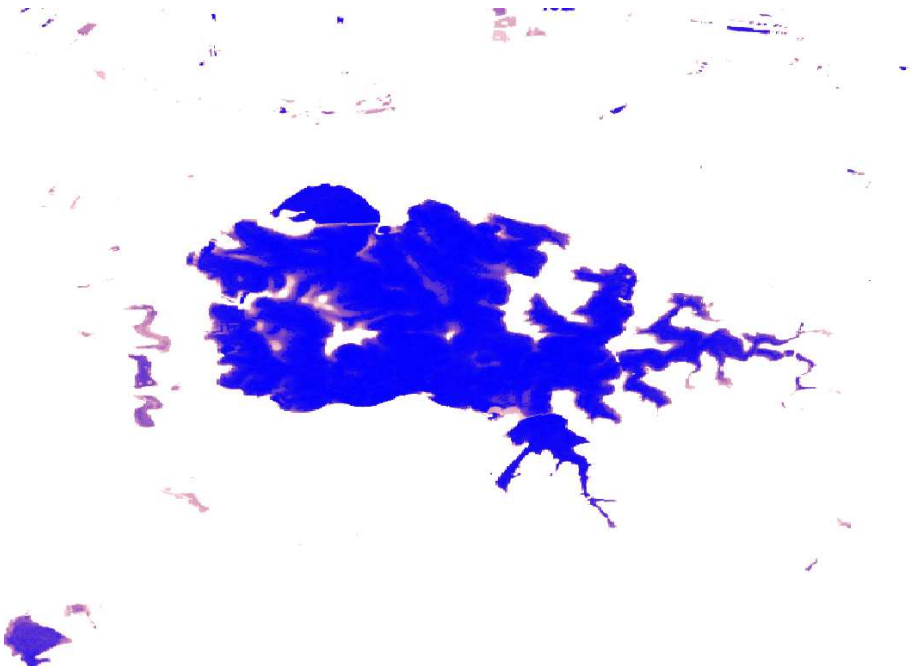


Figure 3.3 – Example of Pekel occurrence mask for Lake Der (Grand Est, France). Occurrence are displayed from blue (100%) to white (0%).

Before that, lower resolution water masks have been published, such as (Carroll et al., 2009), with a resolution of 250 m that were derived from MODIS optical images and the 90 m SWBD (SRTM Water Body Dataset) mask obtained with the SRTM (Shuttle Radar Topography Mission) data. Older masks with a 1 km resolution have also been published before, but such a low resolution does not meet the needs for most hydrological applications.

Vector data and datasets, dedicated to lakes and rivers have also been derived from remote sensing observations, such as those presented in the following sections.

3.3.2 River databases

Compared to raster products, river-specific products enable the processing of high-level products such as river width, slope, and discharge, which are critical for a good understanding of the water cycle and its evolution. The estimation of the global Rivers and Streams Surfaces Area

(RSSA), which plays a major role in greenhouses gas fluxes, is also a goal of these river-specific products (Downing et al., 2012). While previous databases were limited to very large rivers (Global Runoff Data Center database of the German Federal Institute of Hydrology (BfG) (GRDC, 2020), or Global Width database for Large Rivers (Yamazaki et al., 2014)), or to limited geographic areas, current river databases are more exhaustive. The Global River Width from Landsat database (GRWL) (Allen and Pavelsky, 2018) is based on Landsat images and contains more than 2 million kilometers of rivers of width above 30-90 m.
The GRWL database contains multiple attributes for each river reach, and in particular a centerline. The centerline is a set of nodes (one node every 30 m). Each node is associated with an accurate location determined using Landsat data, the local width of the river, and other attributes.

The HydroSHEDS (Lehner et al., 2008) databases have been built from Shuttle Radar Topography Mission (SRTM) elevation data. HydroSHEDS databases such as HydroBASINS (Lehner and Grill, 2013) or HydroRIVERS (Linke et al., 2019) are even more comprehensive than GRWL (36 million km of rivers for HydroRIVERS), but lack important information of GRWL such as the width or the accurate centerline. The MERIT hydro database (Yamazaki et al., 2019) is derived from MERIT (Multi-Error-Removed Improved-Terrain) DEM (Yamazaki et al., 2017).

Specific databases have been developed for reservoirs and dams built on rivers such as the GRanD database which contains 6862 records of reservoirs and their associated dams worldwide (Lehner et al., 2011), and the Global River Obstruction Database (GROD) (Whittemore et al., 2020). GROD contains about 30000 records of obstructions on rivers (dams, locks,...). Another database, HydroFALLS[2] records major waterfalls over rivers worldwide.

In the context of the SWOT mission, the GRWL database has been improved with information from GRanD and HydroSHEDS to form the SWOT *a priori* River Database (SWORD) (Altenau et al., 2021). Similar to GRWL, SWORD database contains many attributes for each river reach, including the centerline. However, the centerline location in the SWORD database cannot be simply projected on SAR images to directly help with river detection. Indeed, beyond the issues associated with elevation and projection, the actual position, and shape of the river can evolve over time (Coulthard and Van De Wiel, 2012), especially for meandering rivers (Hooke, 1984). Such changes can be very brutal in case of major flood events or earthquakes, or when caused by human activity. In addition, rivers can also undergo seasonal or inter-seasonal changes that the database does not fully take into account.

[2] https://wp.geog.mcgill.ca/hydrolab/hydrofalls

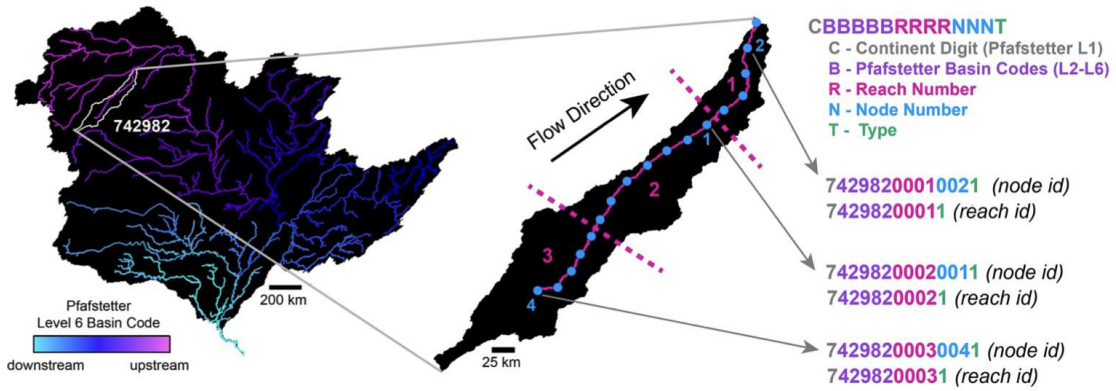


Figure 3.4 – Illustration of how the river centerlines are stored in the SWORD database. Each basin (part of a river or tributary between two tributaries) is given a 6-digits code (742982 in this example) and is divided into multiple reaches. The median reach length is about 10 km, not to scale on the figure. Each reach is given an identifier, that increases in the upstream direction. Nodes are located every 200 m of the reach and are given a node identifier. Figure from (Altenau et al., 2021)

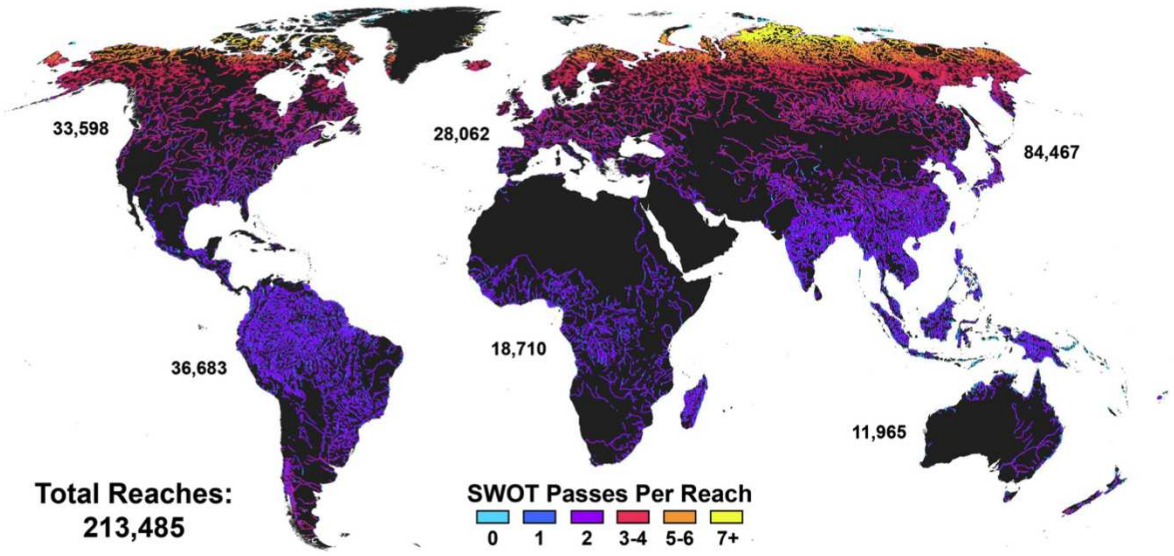


Figure 3.5 – Global repartition of SWORD reaches, with the frequency of SWOT observation associated with each reach. Figure from (Altenau et al., 2021)

3.3.3 Lake databases

In the same way as for rivers, specific databases for lakes have been developed. Following the 2004 Global Lakes and Wetlands Database and the 2016 HydroLAKES database (Messager et al., 2016), the UCLA Circa-2015 lake database (Sheng et al., 2016) will serve as a prior database for the prior lake database (PLD) that will be used for the processing of SWOT data over lakes. This

database will contain several attributes for each lake, including a water extent polygon. Note that defining a lake can be complex as lakes can merge and split depending on the water level, and their area can undergo important changes over time. Similarly to the issue for rivers, this calls for robust approaches if the lake database is used as a prior information to guide water detection.

Summary: Hydrological products

Multiple global raster products and datasets have been produced for hydrology. These datasets could be used to guide the water detection provided the approach is robust enough. An example of a water detection method that relies on such prior data is presented in chapter 5.

Chapter 4

Methodological background

This chapter provides background information from a methodological point of view on approaches that are used in the following chapters:

1. Markov random fields (MRFs), and their adaptations: Conditional Random Fields (CRF) and GrabCut, used in chapters 5 and 6.
2. Variational methods and in particular the MuLoG framework, used in chapter 7.

Organization of this chapter Section 4.1 first briefly presents the principle and the use of MRFs. Then, CRFs are introduced in 4.1.2. The GrabCut method is presented later, in section 6.1.2. Section 4.2.1 presents the background on variational approaches and in particular the MuLoG method.

4.1 Markov random fields

MRFs are a theoretical framework that allows to include both pixel-based information and spatial information in a global model. First introduced by (Geman and Geman, 1984), this formalism has attracted substantial interest in image processing where it is beneficial to take into account the spatial relationship between neighboring pixels. This is especially true for processing of images with a high noise level that calls for the use of regularization. MRF models have been used for various tasks such as denoising, segmentation, or stereo disparity processing.

4.1.1 Definition of the model

MRF formalism borrows the vocabulary of statistical mechanics. The elementary object is a **site**, that can correspond to a pixel for image processing. A set of directly connected, neighboring sites is called a clique. Multiple definitions can be considered for the neighborhood. For example in 2 dimensions, two connectivities can be used:

- 4-connectivity in which only the 4 direct neighbors of the pixel are considered: upper, lower, right, and left pixels.
- 8-connectivity, that adds the diagonal neighbors to those of the 4-connectivity.

In this way, 4-connectivity only allows for cliques of order 1 (a singleton containing only the considered site) and 2 (two 4-connectivity neighbors) while 8-connectivity also allows order 3 and 4 cliques.

The MRF formalism for a random process requires the local conditional probability in any site to respect the Markov property: it can only depend on the local configuration of the cliques the site belongs to. (and not on the configuration of sites outside its neighborhood). The other condition is that there is a non-zero probability for any possible configuration.

It can be shown that given these conditions, there is an equivalence between a MRF and a Gibbs field (Hammersley-Clifford theorem) (Hammersley and Clifford, 1971; Geman and Geman, 1984; Boykov et al., 1998). In a Gibbs field, the global probability of a configuration u is proportional to $\exp(-\sum_{C_s} V_{C_s}(u))$, where $V_{C_s} > 0$ is called the clique potential and is related to the probability of a particular configuration of the sites in the clique C_s. Maximizing this probability amounts to minimizing this sum of clique potentials, i.e. the total energy of the system. Hence, the maximum a posteriori estimation with a MRF model with cliques of order 1 and 2 boils down to minimizing the following global energy (Greig et al., 1989), as summarized by (Boykov and Kolmogorov, 2004):

$$\mathscr{E}(\boldsymbol{\ell}) = \sum_{k} U_k(\boldsymbol{\ell}(k)) + \sum_{k \sim k'} U_{k,k'}(\boldsymbol{\ell}(k), \boldsymbol{\ell}(k')), \tag{4.1}$$

where k is any pixel in the image, and k' is any neighbor of k, $\boldsymbol{\ell}$ is a label field, U_k is an energy term (e.g. data penalty function) and $U_{k,k'}$ is an interaction term (clique potential). When dealing with binary classification like our water detection problem, the label field is binary: $\boldsymbol{\ell}(k) \in \{0, 1\}$. We will not mention the particular characteristics of MRF with non-binary labels or with cliques of more than 2 sites.

When using MRFs for regularization as it is done in image processing, the interaction term $U_{k,k'}$ encourages spatial coherence by penalizing discontinuities between neighboring pixels. For example, for a binary field, the attractive Ising model (Greig et al., 1989) for the interaction energy is defined as:

$$U_{\text{Ising}}(\boldsymbol{\ell}(k), \boldsymbol{\ell}(k')) = \beta |\boldsymbol{\ell}(k) - \boldsymbol{\ell}(k')| \tag{4.2}$$

where $\beta \geq 0$ is a constant penalization that is added if the labels of k and k' are different.

This energy is said to be sub-modular because it satisfies the following condition:

$$U_{k,k'}(0,0) + U_{k,k'}(1,1) \leq U_{k,k'}(1,0) + U_{k,k'}(0,1) \tag{4.3}$$

(Kolmogorov and Zabih, 2004) showed that this condition is necessary and sufficient for the global energy $\mathscr{E}(\boldsymbol{\ell})$ defined equation 4.1 to be graph-representable.

In this case, the global energy is equal to the cost of a given cut on a flow network. A flow network is a directed graph where each edge has a given capacity and receives a flow that cannot exceed its capacity. For each node (or vertex), the sum of the outward flows has to be equal to the sum of the inward flows except for the source ($\mathcal{S}$), which only has outward flow, and the sink ($\mathcal{T}$), which only has inward flow (see Fig. 4.1).

The flow network that corresponds to the global energy function has one non-terminal node N_k for every pixel k in the image. There is one pair of arcs (k, k') and (k', k) with respectively $U_{k,k'}(1,0) + U_{k',k}(0,1)$ and $U_{k',k}(1,0) + U_{k,k'}(0,1)$ capacity between the nodes N_k and N'_k two neighboring pixels k and k'.

One arc with a capacity $U_k(0)$ is added from the source $\mathcal{S}$ to every non-terminal node N_k and one arc with a capacity $U_k(1)$ is added from N_k to the sink $\mathcal{T}$.

The global energy associated with a label field $\boldsymbol{\ell}$ corresponds to the cost of the cut in which all the arcs between N_k and $\mathcal{S}$ are severed if $\boldsymbol{\ell}(k) = 1$ and the arcs between N_k and $\mathcal{T}$ are severed if $\boldsymbol{\ell}(k) = 0$. In addition, the arcs (k, k') are severed if and only if $\boldsymbol{\ell}(k) = 1$ and $\boldsymbol{\ell}(k') = 0$.

The max-flow/min-cut theorem (Dantzig and Fulkerson, 1955) states that in a flow network, the maximum amount of flow passing from the source $\mathcal{S}$ to the sink $\mathcal{T}$ is equal to the total weight of the edges in a minimum cut, which is the smallest total weight of the edges which if removed would disconnect the source from the sink. In this way, the label field that minimizes the global energy can be computed using a maximum flow algorithm. Maximum flow is a classical problem in operational research and multiple methods have been proposed to find its optimal solution. They fall into two principal categories:

Augmenting path methods, first proposed by (Ford and Fulkerson, 1956). Augmenting paths are paths with remaining available capacity (i.e. not saturated). The augmenting paths are iteratively selected and saturated by increasing their flow. Several algorithms using this approach have been proposed such as (Edmonds and Karp, 2003). A new algorithm, still based on augmenting paths but with better empirical performance for the kind of graphs that are used in image processing applications, has been proposed in (Boykov and Kolmogorov, 2004). (Liu and Sun, 2010) proposed an improvement of the this algorithm with a parallel implementation.

Push-relabel algorithms, first proposed by (Goldberg and Tarjan, 1988) are based on the notion of "preflow", which is "like a flow, except that the total amount flowing into a vertex is allowed to exceed the total amount flowing out" (Goldberg and Tarjan, 1988). The methods iteratively push the excess flow to the sink through the estimated shortest path. While the theoretical performances of these algorithms are better than those of augmenting path methods (Goldberg, 2008), the empirical performance for image processing problems is not as good as those of the (Boykov and Kolmogorov, 2004) method.

Note that other optimization approaches find an approximate, non-optimal solution for MRF and can be relevant when the interaction term is not sub-modular (in this case finding the global optimum would be a NP-hard problem (Boykov and Kolmogorov, 2004)). These approaches have been mainly used when graph-cut-based methods were not available.

As the computation of the solution using a graph-cut method can be long and require large memory resources, several methods have been proposed to make this processing faster and more scalable. For instance, (Lermé et al., 2010) proposes to remove nodes that are not likely to be useful (i.e. near a border). (Delong and Boykov, 2008) introduces a way to parallelize the processing. Approximate approaches that run faster than the previous exact approaches have also been proposed, such as methods based on electrical flows (Christiano et al., 2011; Lee et al., 2013; Yim, 2015) that are only available when dealing with undirected graphs.

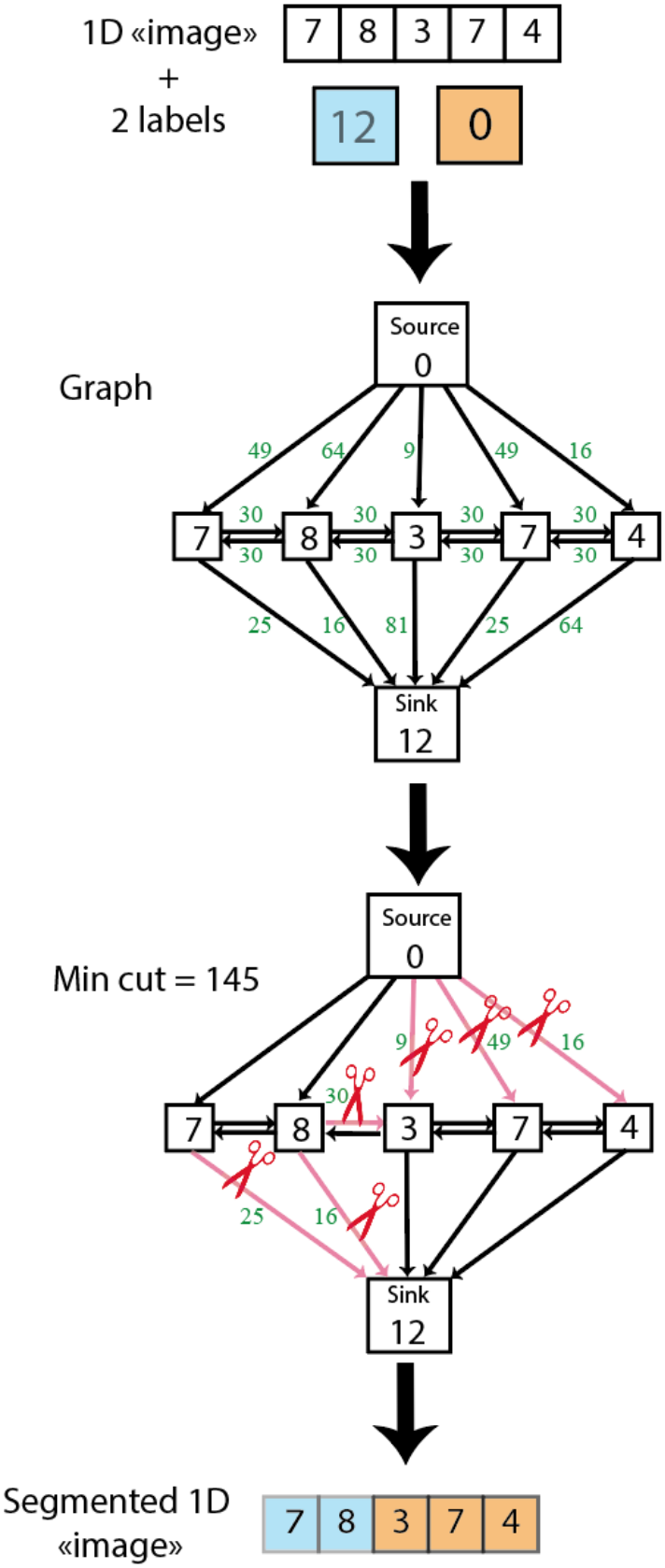


Figure 4.1 – Schematic view of the principle of Graph Cut segmentation on a 1D image with only 5 pixels. The data term is quadratic and the regularization term is based on an Ising model with $\beta = 30$. Severed arcs are shown in red on the Min cut figure. The total cost of the cut is the sum of the capacities of the severed edges. Note that the regularization caused the 4th pixel to fall into the orange class, while a maximum likelihood would have classified it in the blue class.

4.1.2 Conditional random fields

A limitation of traditional MRF models for images segmentation is that the regularization term is the same for all pairs of pixels of the image (e.g. for Ising regularization) and does not take into account the context of the image. For example, the penalization for a boundary in the label field $\boldsymbol{\ell}$ is the same when aligned with an edge of the image and in the middle of a homogeneous region. This is an issue for the detection of small structures.

A simple solution for that issue is to use a regularization term that depends on the local properties of the image. For example, (Boykov and Jolly, 2001) propose an interaction term of the form:

$$U_{k,k'}(\boldsymbol{\ell}(k), \boldsymbol{\ell}(k')) = \beta \exp\left(-\frac{(\boldsymbol{y}(k') - \boldsymbol{y}(k))^2}{\lambda}\right) \cdot \frac{|\boldsymbol{\ell}(k) - \boldsymbol{\ell}(k')|}{dist(k, k')}, \tag{4.4}$$

with two tuning hyper-parameters $\beta \geq 0$ and $\lambda > 0$ and a function $dist(k, k')$ that gives the Euclidean distance between k and k', i.e. 1 or $\sqrt{2}$ if k and k' are connected only by a diagonal. Under these conditions, the function is sub-modular and the global energy can be represented and optimized through a graph.

This approach is related to the general theoretical framework of CRF, introduced by (Lafferty et al., 2001) that has been widely used for applications in images segmentation, see for example (He et al., 2004; Kohli et al., 2008).

Summary: Markov Random Fields

Markov Random Fields (MRF) approaches and derived approaches (CRF) are very valuable for segmentation as they allow to combine multiple energies (data term, regularization term,...) and to efficiently find the optimal partition through graph cut methods.

4.2 Variational methods for image denoising

This section will briefly present variational methods for image denoising. In particular, the aim is to introduce the MUlti-channel LOgarithm with Gaussian denoising (MuLoG) framework.

4.2.1 General background on variational approaches for image denoising

The basic concept behind variational approaches for image processing is to consider an image $\boldsymbol{x}$ not as a set of pixels but as a function $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ is the domain over which the image is defined. A functional $\mathcal{F}$ can be defined on the space of all functions f to map any of these functions to a value $\mathcal{F}(f)$.

For example, a functional that can be minimized for the denoising of a noisy image $\boldsymbol{y}$ can be of the following form:

$$\mathcal{F}_{ex}(f) = \int_{u \in \Omega} E_D(f(u), \boldsymbol{y}(u)) du + E_R(f) \tag{4.5}$$

where $E_D(f(u), \boldsymbol{y}(u))$ is a function that depends on the local value of f and $\boldsymbol{y}$ at spatial location u and that ensures a good fidelity of the function f to the image $\boldsymbol{y}$. $E_R(f)$ is a

functional that ensures a good regularity for the function f, for example by penalizing its spatial variations.

As an illustration, we can consider the total variation (TV) denoising for additive white Gaussian noise proposed by (Rudin et al., 1992). It minimizes the integral of the gradient ∇f of f under two constraints: the mean of f and $\boldsymbol{y}$ have to be equal and the mean squares difference has to be equal to σ_n^2, where σ_n is the standard deviation of the noise.

$$\begin{aligned} \hat{f} &= \arg\inf_f \int_{u\in\Omega} ||\nabla f(u)|| \mathrm{d}u \\ \text{subject to } &\int_{u\in\Omega} f(u)du = \int_{u\in\Omega} \boldsymbol{y}(u)\mathrm{d}u \\ \text{subject to } &\int_{u\in\Omega} (f(u)-\boldsymbol{y}(u))^2 du = \sigma_n^2 \int_{u\in\Omega} \mathrm{d}u \end{aligned} \tag{4.6}$$

Unlike quadratic ("Tikhonov") regularization that minimizes the integral of $||\nabla f(u)||^2$, TV regularization is not differentiable in 0 but is much better at preserving image edges that are blurred by quadratic regularization.

This TV minimization under constraint can be written as the minimization of a functional as in 4.5 with the following energies (Chambolle and Lions, 1997):

$$\begin{aligned} E_D(f, \boldsymbol{y}(u)) &= (f(u)-\boldsymbol{y}(u))^2 \\ E_R(f) &= \int_{u\in\Omega} \beta||\nabla f(u)|| du \end{aligned} \tag{4.7}$$

where β is a hyper-parameter that adjusts the strength of the regularization. Hence the global energy can he written as a functional:

$$E_{global}(f) = \int_{u\in\Omega} (f(u)-\boldsymbol{y}(u))^2 + \beta||\nabla f(u)|| \mathrm{d}u \tag{4.8}$$

This global energy can be minimized using numerical optimization methods that find a local minimum of the energy, or even the global minimum in the above example as the functional E_{global} is convex. For example, (Rudin et al., 1992) derives an explicit scheme from the Euler-Lagrange equation associated with the problem.

Beyond the classical formulations for additive Gaussian noise, adaptations of variational methods have been proposed for multiplicative noise. (Aubert and Aujol, 2008) introduce a functional $\mathcal{F}_A$ that combines a total variation regularization with a data term adapted for multiplicative Gamma distributed noise:

$$E_{\text{global}} = \mathcal{F}_A(f) = \int_{u\in\Omega} \log(f(u)) + \frac{\boldsymbol{y}(u)}{f(u)} + \beta||\nabla f(u)|| \mathrm{d}u \tag{4.9}$$

Here, the term $\log(f(u)) + \frac{\boldsymbol{y}(u)}{f(u)}$ derives from the neg-log-likelihood $-\log(p(\boldsymbol{y}(u)|f(u)))$

However, such approaches still have a major limitation in the choice of the criterion for the regularization, that has to be chosen simple enough to be able to solve the minimization ((Aubert and Aujol, 2008) choose a convex regularization).

This limits its performances compared to more recent denoising approaches in terms of detail preservation and remaining noise. Unfortunately, such denoisers are often built for Gaussian additive noise and cannot be directly used on SAR images.

4.2.2 MuLoG framework for SAR image denoising

To address this issue, the MuLoG method (Deledalle et al., 2017a) proposes a new variational formulation in which the regularization term is given by a state-of-the-art Gaussian denoiser such as BM3D (Dabov et al., 2007). This denoiser can be non-local. MuLoG operates on log-transformed SAR images, whose noise follows a Fisher-Tippett distribution (see section 2.1.3.1). To prevent bias caused by the Gaussian denoising, a data fidelity term is derived from the exact distribution.

Using the same notations as above, the global energy for the denoising of $\boldsymbol{y}$ is given by the following functional:

$$\mathcal{F}(f) = -\log(\mathrm{p}(\boldsymbol{y}|f)) + \mathcal{F}_{\mathrm{reg}}(f) \tag{4.10}$$

Here, $-\log(\mathrm{p}(\boldsymbol{y}|f))$ derives from the Fisher-Tippett distribution of $\boldsymbol{y}$ and is convex (Deledalle et al., 2017a):

$$-\log(\mathrm{p}(\boldsymbol{y}|f)) = L\sum_k f(k) + e^{\boldsymbol{y}(k)-f(k)} + Cst \tag{4.11}$$

As the Fisher-Tippett distribution of log-transformed images is defined on $\mathbb{R}$, the minimization problem is unconstrained. In contrast, Gamma-distributed intensity would have required the positivity constraint $f(k) \geq 0$ on f.

To find the optimal solution, the MuLoG framework follows the same variable splitting for mono-channel SAR image as MIDAL (multiplicative image denoising by augmented Lagrangian) (Bioucas-Dias and Figueiredo, 2010). This variable splitting introduces two new variables $\widehat{\boldsymbol{d}}$ and $\widehat{\boldsymbol{z}}$. The resulting problem is then minimized by using the Plug and Play Alternating direction method of multipliers (ADMM) (Chan et al., 2017) implementation of the ADMM. ADMM method was first proposed by (Glowinski and Marroco, 1975) and improved by (Gabay and Mercier, 1976; Eckstein and Bertsekas, 1992; Boyd et al., 2011).

The minimization of $\mathcal{F}$ defined equation 4.10 boils down to the iterative computation of the following steps:

$$\widehat{\boldsymbol{z}} \leftarrow \underset{\boldsymbol{z}\in\mathbb{R}^n}{\arg\min} \ \frac{\beta}{2}||\boldsymbol{z} - f + \widehat{\boldsymbol{d}}||^2 + \mathcal{F}_{\mathrm{reg}}(\boldsymbol{z}), \tag{4.12}$$

$$\widehat{\boldsymbol{d}} \leftarrow \widehat{\boldsymbol{d}} + \widehat{\boldsymbol{z}} - \widehat{f}, \tag{4.13}$$

$$\widehat{f} \leftarrow \underset{f\in\mathbb{R}^n}{\arg\min} \ \frac{\beta}{2}||\widehat{\boldsymbol{z}} - f + \widehat{\boldsymbol{d}}||^2 - \log \mathrm{p}(\boldsymbol{y}|f), \tag{4.14}$$

Here, the minimization of equation 4.12 corresponds to a denoising problem under additive white Gaussian noise ($\mathcal{F}_{\text{reg}}$ depends on the choice of the Gaussian denoiser). In this way, this step can be handled by embedding any off-the-shelf Gaussian denoiser (Venkatakrishnan et al., 2013; Buzzard et al., 2018). This makes this approach both flexible and efficient for non-Gaussian denoising.

The data fidelity correction of equation 4.14 has a closed-form solution for SAR intensity despeckling which involves special functions (Bioucas-Dias and Figueiredo, 2010). (Deledalle et al., 2017a) and (Bioucas-Dias and Figueiredo, 2010) recommend a fast approximation of the solution by applying the Newton method, i.e. using the following iteration:

$$\widehat{f}(k) \leftarrow \widehat{f}(k) - \frac{\beta(\widehat{f}(k) - \widehat{\boldsymbol{d}}(k) - \widehat{\boldsymbol{z}}(k)) + L(1 - e^{\boldsymbol{y}(k) - \widehat{f}(k)})}{\beta + L e^{\boldsymbol{y}(k) - \widehat{f}(k)}}. \tag{4.15}$$

Summary: Variational methods for image denoising

Variational denoising method work by finding the function f, which corresponds to an image, that minimizes the functional $\mathcal{F}(f)$ that combines a data fidelity term and a regularization term. The approach proposed in chapter 7 to denoise the temporal geometric mean is derived from the MuLoG method (Deledalle et al., 2017a) which is a variational method.

Part II

Proposed approaches

Chapter 5

Guided extraction of narrow rivers on SAR images using an exogenous river database

This chapter is based on an already published IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS) article. Besides this main publication, the method corresponding to section 5.2.2 on linear structures detection has also been presented in the 2021 EUSAR conference:

- N. Gasnier, L. Denis, R. Fjørtoft, F. Liège and F. Tupin, "Narrow River Extraction From SAR Images Using Exogenous Information," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 5720-5734, 2021 (Gasnier et al., 2021b)
- N. Gasnier, L. Denis and F. Tupin, "Generalized Likelihood Ratio Tests for Linear Structure Detection in SAR Images," EUSAR 2021; 13th European Conference on Synthetic Aperture Radar, 2021, pp. 1-6. (Gasnier et al., 2021c)

Some figures and substantial parts of the text of this chapter have been taken from these articles.

5.1 Introduction

This chapter presents a new approach for the detection of narrow rivers in SAR images using guiding information from an exogenous database such as GRWL or SWORD (see 3.3.2). Section 5.1.1 introduces the motivation for such a specific approach, then our new method is introduced in section 5.2 and some results for SWOT and Sentinel-1 images are presented in section 5.3.

5.1.1 Motivation

Limitation of SWOT's baseline water detection method While it is appropriate for the detection of large water bodies, the SWOT's baseline detection approach (presented in section 3.1) can hardly handle the detection of very narrow rivers because of its regularization term that tends to delete small structures. This is visible in **Figure** 5.1 where most of the Petit Rhone river is missing from the baseline detection. This is an issue as this river is in SWOT's *a priori* river database SWORD and its detection is needed for further processing. Besides, as its average width is above 100m, the relative error on the surface should be below 15% according to SWOT mission requirements (see 3.2.2). This brings the need for a specific approach to detect the rivers in SWOT images despite a worst-case situation for the sensor performances.

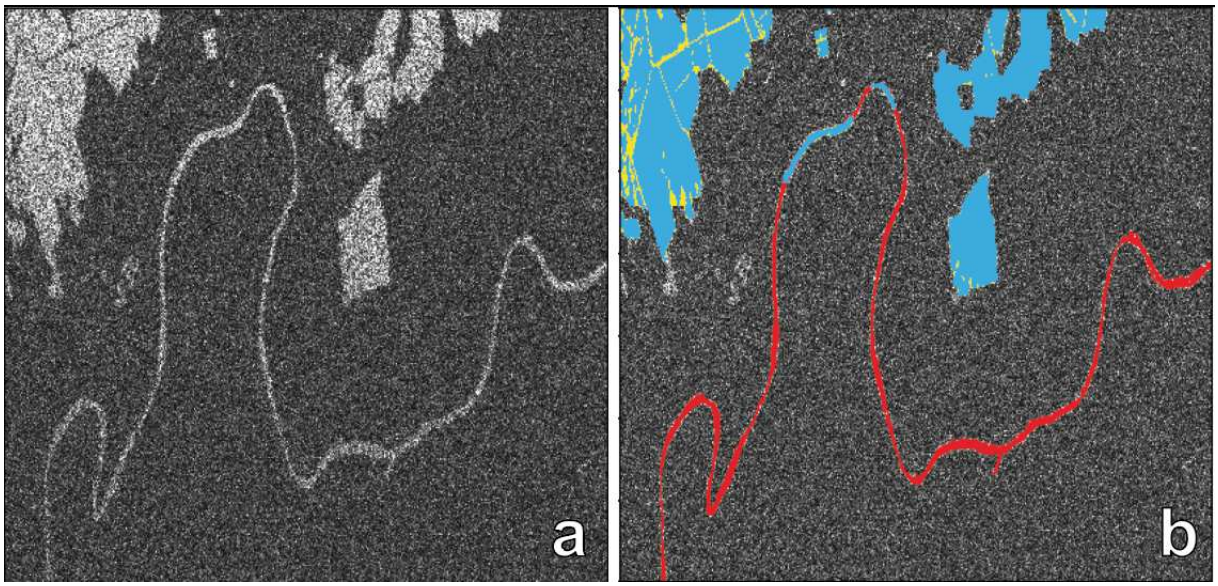


Figure 5.1 – Crop of (a) a simulated SWOT image of Petit-Rhone River and surrounding lakes, presented in Section 5.3. The image is simulated under worst-case hypothesis (low water/land contrast). (b) shows the detection results using baseline MRF approach: blue is correctly detected while red marks the missed detection and yellow marks the false detection. While the surrounding lakes are correctly detected, most of the river is missing from the detection.

Even using river-specific approaches, the detection of narrow rivers in SAR images with a limited false detection rate is very difficult without using any exogenous information. Indeed, beyond usual issues associated with speckle noise and low contrast, river detection is particularly complex because roads, terrain slope, and various artifacts can create structures resembling rivers such as in **Figure** 5.2. Distinguishing rivers from other visually similar structures such as the

large road in **Figure** 5.2(a) or the topography artifact in **Figure** 5.2(b) can be very difficult or even impossible when using only the information available in the image, especially when the contrast of an actual river can be very low (as in **Figure** 5.2(b)). To prevent false detection, prior information about the location and the direction of known rivers can be useful. It allows distinguishing linear structures corresponding to a known river from other visually similar linear features. For example, rough waypoints from exogenous data can give information about the course of the river that has to be detected.

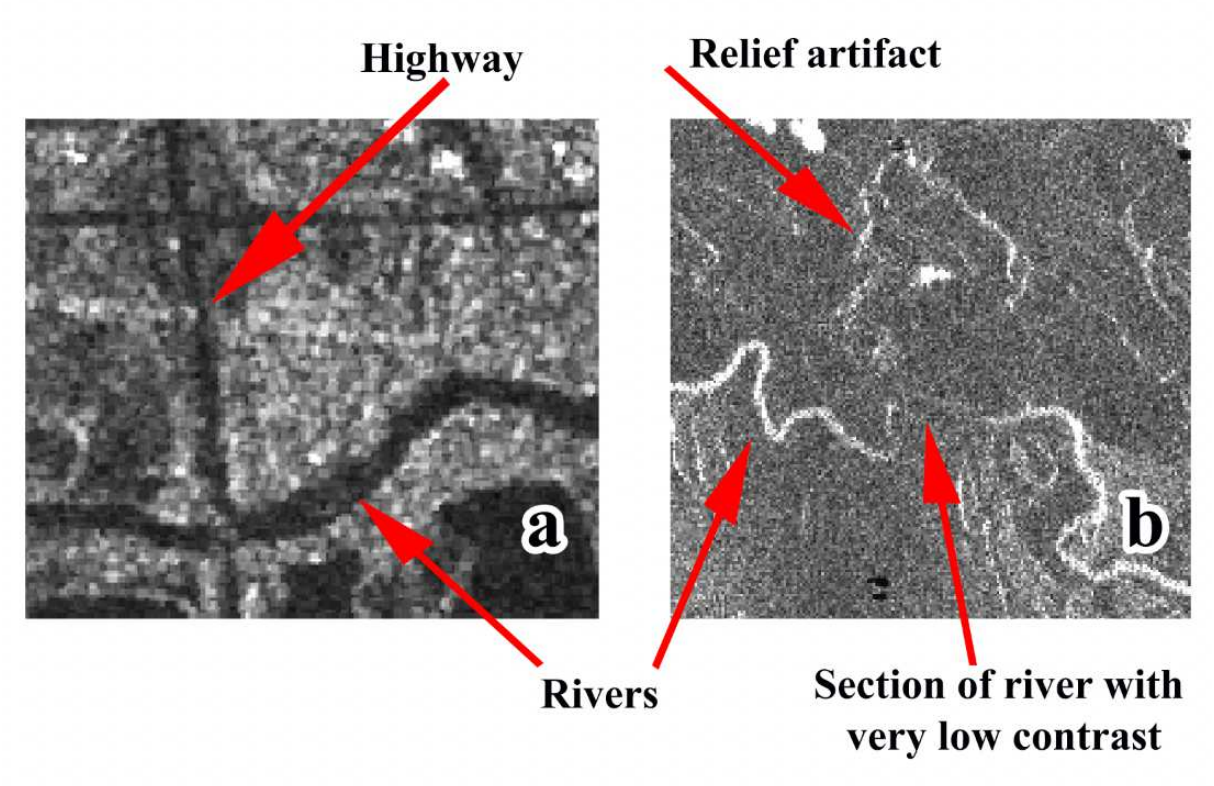


Figure 5.2 – Crop of (a) a Sentinel-1 image from Des Moines, and (b) a simulated SWOT image (Saline), presented in Section 5.3. Both images contain linear structures that correspond to actual rivers and linear structures that correspond to other structures: a large road for (a) and terrain slope layover effects for (b). Image (b) also shows a river section with very low contrast.

In this context, the global river databases provide, on a global scale, information that can be included within new approaches for river detection from SAR images. Before such global databases became available, the use of exogenous information was difficult and often required manual preparation of input and semi-automated approaches, such as (Gruen and Li, 1994; Dillabaugh et al., 2002) for optical images. In contrast, GRWL contains a centerline for each river that provides information about the course of the river. If this database centerline did perfectly correspond to the actual river centerline in the image after projection to the image coordinates, its use would be straightforward and only the third step of the proposed method would be needed. Unfortunately, direct use of the prior centerline of a river provided by the database to detect and segment the river in a SAR image remains problematic. Indeed, there are three main reasons why there can be a discrepancy between the database centerline projection

in the image and the actual river:

1. The actual position and shape of the river can evolve over time (Coulthard and Van De Wiel, 2012), especially for meandering rivers (Hooke, 1984). Such changes can be very quick in case of major flood events or earthquakes, or when caused by human activity. Rivers can also undergo seasonal changes that the database does not take into account.

2. There can be a positional error caused by the projection of the database centerline into the radar image. For Sentinel-1 Ground Range Detected images, it can be induced by the GRD image construction or ortho-rectification process (inaccurate digital elevation model, or errors in the water level). For SWOT images, as the water detection is done in radar geometry and before water height extraction, shifts could come from a difference between the prior water level used for the projection of the centerline and the actual water level. The near-nadir geometry of SWOT is very sensitive to this, as even a relatively small difference in elevation can lead to a major shift in position in the range direction as illustrated in **Figure** 5.3.

3. There may be some errors in the database itself, especially in areas with complex topology or dense vegetation.

This brings the need for an approach that can exploit the exogenous information provided by GRWL's river centerlines while being robust to discrepancies between the projection of these centerlines and the true river in the image. We, therefore, propose a robust approach that uses the database centerlines as a source of approximate waypoints that can be used in combination with the image to retrieve the actual river centerline. This centerline can then be used to accurately detect the river extent while avoiding confusion with other linear structures.

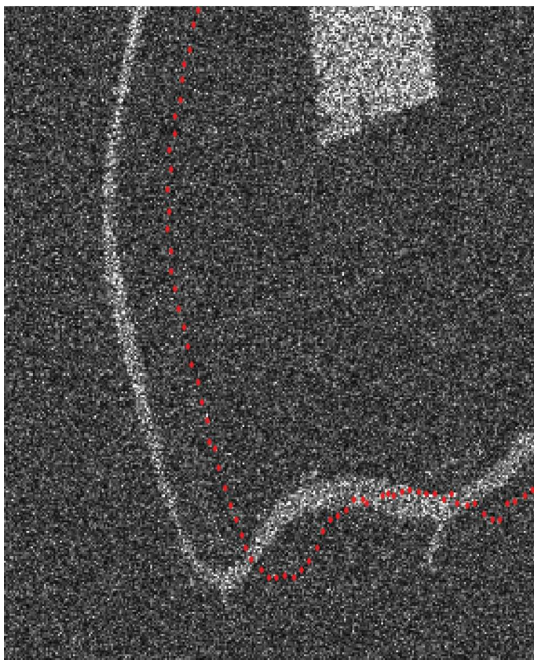


Figure 5.3 – Illustration of the displacement between the database centerline projected in radar geometry (red dotted line) and the river observed in a simulated SWOT image (where water is bright and land is dark). Such a displacement can be caused by variations in water elevation and inaccuracies in the digital elevation model used for projection: a few meters difference between actual and prior elevation can lead to shifts of hundreds of meters in ground range.

5.1.2 Existing methods for river extraction

As already mentioned in section 3.1 he problem of river detection in SAR images has strong similarities other widely studied topics that involve the detection of other narrow structures, in SAR images processing and outside:

- Roads detection in SAR remote sensing (Negri et al., 2006; Tupin et al., 2002; Sun et al., 2021; Tupin et al., 1998; Perciano et al., 2016; Chanussot et al., 1998)
- Roads crack detection (Amhaz et al., 2016; Oliveira and Correia, 2013; Delagnes and Barba, 1995)
- Blood vessels detection in retina fundus images (Rossant et al., 2011)

Approaches developed to address these problems can be adapted for the detection of rivers in SAR images. For example, (Valero et al., 2010) proposes an approach based on mathematical morphology for road detection in high-resolution images. This approach has been adapted for rivers and automated using machine learning by (Klemenjak et al., 2012). (Sghaier et al., 2017) combines it with structural feature sets. Other approaches based on active contours have also been used, such as (Han and Wu, 2018). For SWOT images, specific approaches have been proposed by (Cao et al., 2011) and (Lobry et al., 2017).

5.2 Proposed river segmentation pipeline

5.2.1 Technical overview

As mentioned in the introduction, the aim was to provide a new framework for river extraction in SAR images guided by a river database in order to overcome the limitations of blind detection while being robust to a discrepancies in river location and shape between the database and the actual images.
To achieve this purpose, we proposed a three steps framework presented in **Figure** 5.4.

- The first step consists in applying a GLRT (Generalized Likelihood Ratio Test) based line detector described in the next section to the SAR image. Its response shows the likelihood of the presence of a linear structure for every pixel of the image, irrespective of the cause of the linear structure (river, road, artifact, ...).
- The second step uses the Dijkstra algorithm to find the shortest path between two nodes through a cost array deriving from the response of the linear structure detector.
- The third step consists in segmenting the river reach on the image around the previously estimated centerline. A new conditional random field (CRF) approach has been proposed for this purpose.

The two first steps lead to an estimation of the actual river centerline on the image which is robust both to noise and low contrast in the image and to discrepancy in shape and position of the centerline between the database and the image.

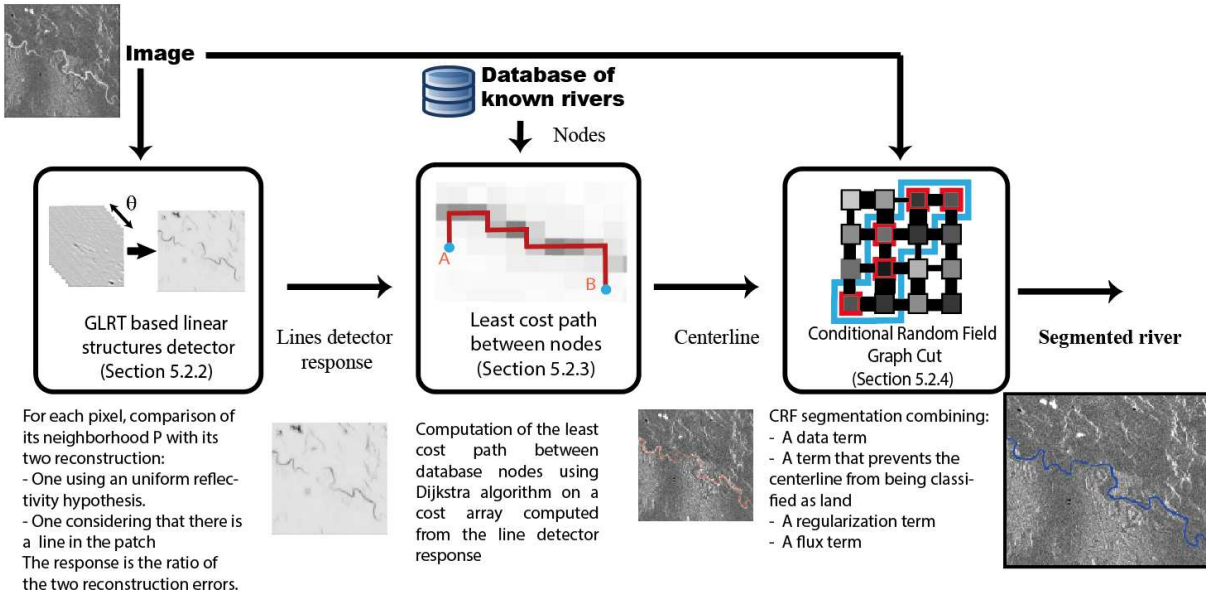


Figure 5.4 – Global overview of the proposed method: the first step consists in computing the linear structure detector response, that is then used in the second step with the nodes from the *a priori* database to retrieve the centerline. The river is then segmented around the centerline using a CRF approach in the third step.

5.2.2 Linear structure detector for SAR images

As mentioned before, the first step of our approach computes a map that indicates the likelihood of the presence of a linear structure in every pixel of the image. In our context, a linear structure can be defined as a set of contiguous pixels in a long and thin layout (width of a few pixels) whose reflectivity is significantly different from the reflectivity of the background. For river detection, the relevant linear structures can be dark, as for most sensors such as Sentinel-1 or RADARSAT, or bright for near-nadir sensors like KaRIN. This linear features detection on SAR images can be very difficult due to the high level of speckle and to the very low contrast of certain rivers. Because of this, methods developed for optical images such as (Geman and Jedynak, 1996; Fischler et al., 1987; Deschênes and Ziou, 2000; Vanderbrug, 1976; Movaghati et al., 2010) cannot be directly applied to SAR images, even after log transformation. Methods specific to SAR images have been proposed in the past, such as (Hellwich et al., 2002) that uses both intensity and coherence images, and (Tupin et al., 1998) that combines the results of two detectors: one being based on ratios in a neighborhood, the other being based on cross-correlation. (Chanussot et al., 1999) introduces a morphological line detector and multitemporal fusion approaches. We proposed a new line detector that improved the detection over (Tupin et al., 1998). This detector has been chosen to be integrated within our framework as it has shown better performances than (Tupin et al., 1998) in terms of ROC.

5.2.2.1 GLRT criterion for linear structure detection

Our detection criterion evaluates the likelihood of the presence of a linear structure centered at a given pixel k by comparing two hypotheses:

- $\mathcal{H}_0$: there is no linear structure
- $\mathcal{H}_1$: there is a linear structure

The comparison between these two hypotheses is done by determining which hypothesis best explains the observed patch $\boldsymbol{I}_k \in \mathbb{R}^{(2N+1)\times(2N+1)}$ (i.e., the vector formed by the concatenation of all the intensities inside a small square window of size $(2N+1)\times(2N+1)$ centered at the k-th pixel). In this paper, we assume that the null hypothesis $\mathcal{H}_0$ ("no linear structure") corresponds to a patch with a constant reflectivity $R_H k$. This assumption may seem too restrictive. Inhomogeneous reflectivity distributions inside the patch that can neither be modeled by a constant nor by a shift-invariant profile lead to similar likelihood values under $\mathcal{H}_0$ and $\mathcal{H}_1$, they, therefore, do not lead to false detection. The simplifying assumption of a constant reflectivity under $\mathcal{H}_0$, therefore, does not limit the applicability of the method to homogeneous or linear structures. Under the alternative hypothesis $\mathcal{H}_1$, a linear structure is present and the reflectivities inside the patch are shift-invariant in the direction of the structure, see **Figure** 5.5.

The likelihood of each hypothesis depends on several unknown parameters:

- The constant reflectivity $R_H k$, under $\mathcal{H}_0$
- The reflectivity profile $\boldsymbol{P}_k$ and the line direction θ_k under $\mathcal{H}_1$.

These unknown parameters can be obtained by the maximum likelihood estimator. The decision in favor of hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$ can be made based on the generalized likelihood ratio, i.e., the ratio of the likelihoods of each hypothesis where unknown parameters are replaced by their maximum likelihood estimates (Van Trees, 2004):

$$\mathrm{GLR}_k = \frac{\mathrm{p}(\boldsymbol{I}_k|\mathcal{H}_1, \widehat{\boldsymbol{P}_k}, \widehat{\theta_k})}{\mathrm{p}(\boldsymbol{I}_k|\mathcal{H}_0, \widehat{R_H k})} \,. \tag{5.1}$$

This GLR computation must be repeated at each pixel k of the image. Computing the maximum likelihood estimators is simplified when log-transformed intensities are considered. As presented in section 2.1.3.1, the log-transformed intensity image follows a Fisher-Tippett distribution that can be approximated by a Gaussian distribution, especially for relatively high ENL ($L = 4$ for SWOT, $L = 4.4$ for Sentinel-1 GRD). This Gaussian approximation can be useful to obtain closed-form expressions for the likelihood estimators and GLRs:

$$\mathrm{p}(y|\mu_{GA}) \simeq \frac{1}{\sigma_{GA}\sqrt{2\pi}} \, \mathrm{e}^{-\frac{1}{2}\left(\frac{y-\mu_{GA}}{\sigma_{GA}}\right)^2} , \tag{5.2}$$

where y is the log-intensity of any pixel in the patch, $\sigma_{GA} = \sqrt{\psi'(L)}$, and $\mu_{GA} = \log(\boldsymbol{R}_k) - \log(L) + \psi(L)$. $\boldsymbol{R}_k$ is the vector created by concatenating all the reflectivities values for the pixels belonging to the patch centered in k.

Under these assumptions, the biased log-reflectivity $\hat{r}_k = \mu_{GA}$, of the homogeneous background under $\mathcal{H}_0$ hypothesis can be estimated from the mean of the log-transformed intensities $\boldsymbol{y}_k$ of the patch. The estimation of the reflectivity profile and of the line orientation is described in more details in the next section.

When the estimates of the (biased) log-reflectivities of the k-th patch $\hat{r}_k\mathbf{1}$, under $\mathcal{H}_0$, and $\widehat{\boldsymbol{r}}_{k,\hat{\theta}}$, under $\mathcal{H}_1$, are substituted in the definition of GLR_k in equation (5.1), we obtain, under our Gaussian approximation:

$$\log(\mathrm{GLR}_k) = \frac{1}{2}||\boldsymbol{y}_k - \hat{r}_k\mathbf{1}||^2 - \frac{1}{2}||\boldsymbol{y}_k - \widehat{\boldsymbol{r}}_{k,\hat{\theta}}||^2 \,, \tag{5.3}$$

where $\boldsymbol{y}_k = \log(\boldsymbol{I}_k)$, $\mathbf{1}$ is a vector of ones with the same dimension as $\boldsymbol{y}_k$ (the number of pixels in a patch).

5.2.2.2 Modeling of a linear structure

Before describing the linear structure parameters, we first define how a linear structure can be characterized at the scale of a patch. Considering the patch of size $(2N+1) \times (2N+1)$, centered at the k-th pixel, with a dark line that crosses the patch as depicted in **Figure 5.5**, two ingredients define our model: (i) the reflectivity is lower (for dark lines) in the central line than farther from the line, and (ii) the reflectivity distribution is invariant in the direction of the line. The 1D distribution of the reflectivity along the direction orthogonal to the line is called the **reflectivity profile**.

Beyond the shift-invariance of the reflectivity in the direction of the line, we also require the profile to be symmetrical with respect to the median axis of the line. This is useful to improve the localization of the linear structure and to reduce the number of false positives.

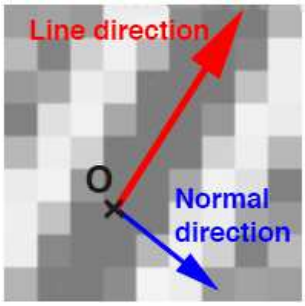


Figure 5.5 – The profile of the reflectivity of the linear structure is defined in the direction that is normal to the direction of the linear structure.

Estimation of the local orientation of the linear structure The maximum likelihood estimate $\widehat{\theta}$ of the angle of the linear structure in the patch is obtained by uniformly sampling the orientations (60 steps are used in the range $[0, \pi]$ in our experiments). The largest value of GLR_k is retained among all values computed for the set of orientations considered.

Estimation of the reflectivity profile of the linear structure The first step to estimate the reflectivity profile of the linear structure is to model the mapping from a 1D profile $\boldsymbol{p}_{k,\theta}$ to a 2D patch $\boldsymbol{r}_{k,\theta}$. In order to cover a patch of size $(2N+1) \times (2N+1)$ pixels for all orientations of the line, the 1D profile has to cover $\sqrt{2}(2N+1)$ pixels. Since we consider profiles that are symmetrical with respect to the central line, defining the profile only for the first $\sqrt{2}(N+1)$ pixels from the patch center is sufficient. The mapping operation amounts to interpolating the 1D profile at each pixel of the 2D patch according to the distance of the pixel to the line that goes through the patch center and that forms an angle θ with respect to the horizontal direction. This interpolation operation is a linear transform characterized by a matrix $\boldsymbol{M}_\theta$ of size $(2N+1) \times (2N+1) \times \sqrt{2}(N+1)$:

$$\boldsymbol{r}_{k,\theta} = \boldsymbol{M}_\theta \, \boldsymbol{p}_{k,\theta} \tag{5.4}$$

The second step is to compute the maximum likelihood estimate $\widehat{\boldsymbol{p}}_\theta$ of the reflectivity profile of a linear structure oriented in the direction θ. Under our Gaussian approximation, this corresponds to the least squares solution:

$$\widehat{\boldsymbol{p}}_{k,\theta} = \boldsymbol{M}_\theta^{\mathrm{pinv}} \boldsymbol{y}_k \tag{5.5}$$

where $\boldsymbol{M}_\theta^{\mathrm{pinv}}$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{M}_\theta$.

In order to force the reflectivity at the center of the line structure to be the minimum of the reflectivity profile, a thresholding operation is added after estimating the maximum likelihood profile:

$$\widehat{\boldsymbol{p}}^{+}_{k,\theta} = \max\left\{\widehat{\boldsymbol{p}}_{k,\theta}, [\widehat{\boldsymbol{p}}_{k,\theta}]_1\right\} , \tag{5.6}$$

where the maximum is applied component-wise and $[\widehat{\boldsymbol{p}}_k, \theta]_1$ is the value of the log-reflectivity at the center of the profile (first element of the vector). In the thresholded profile $\widehat{\boldsymbol{p}}^{+}_{\theta}$, no reflectivity can be lower than the reflectivity at the center of the profile. If, rather than dark lines, bright

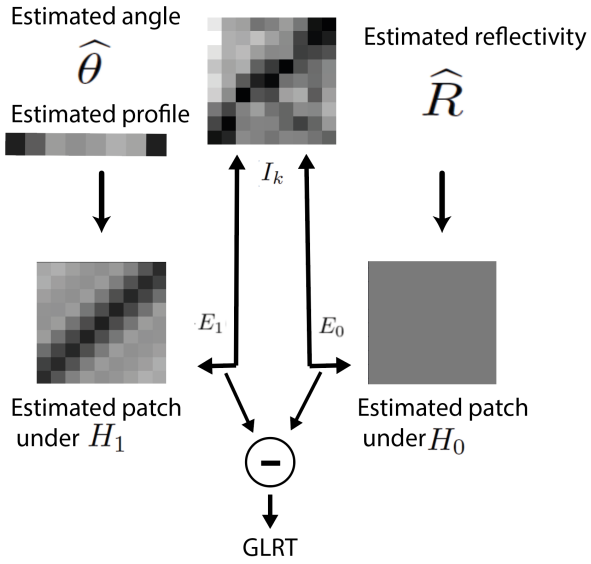


Figure 5.6 – General presentation of the linear structures detection performed on the log transformed images.

lines were to be detected, this maximum should be replaced by a minimum to define a profile $\widehat{\boldsymbol{p}}_\theta^-$ where no reflectivity is brighter than the reflectivity at the center of the line.

From this estimated profile, the estimated reflectivities inside the patch centered on pixel k are obtained by applying the interpolation operator $\boldsymbol{M}_\theta$:

$$\widehat{\boldsymbol{r}}_{k,\theta} = \boldsymbol{M}_\theta \, \widehat{\boldsymbol{p}}_{k,\theta}^+ \, . \tag{5.7}$$

The maximum likelihood orientation of the line structure in the k-th patch is obtained by:

$$\widehat{\theta} = \arg\max_\theta \; \|\boldsymbol{y}_k - \boldsymbol{M}_\theta \, \widehat{\boldsymbol{p}}_{k,\theta}^+\|^2 \, . \tag{5.8}$$

The computation of $\mathrm{GLR}_{k,\theta}$ is then obtained by the application of equation (5.3) for a given orientation θ. The detection criterion at pixel k and orientation θ can be expanded as follows:

$$\begin{aligned} \log \mathrm{GLR}_{k,\theta} &= \tfrac{1}{2}||\boldsymbol{y}_k - \hat{r}_k \mathbf{1}||^2 - \tfrac{1}{2}||\boldsymbol{y}_k - \widehat{\boldsymbol{r}}_{k,\theta}||^2 \\ &= \tfrac{1}{2}||\hat{r}_k \mathbf{1}||^2 - \boldsymbol{y}_k^{\mathrm{T}} \hat{r}_k \mathbf{1} + \boldsymbol{y}_k^{\mathrm{T}} \widehat{\boldsymbol{r}}_{k,\theta} - \tfrac{1}{2}||\widehat{\boldsymbol{r}}_{k,\theta}||^2 \end{aligned} \tag{5.9}$$

where the maximum likelihood estimate of the reflectivity $\hat{r}_k$ of a constant patch is the mean log intensity in the patch: $\hat{r}_k = \mathbf{1}^{\mathrm{T}}\boldsymbol{y}_k / \mathbf{1}^{\mathrm{T}}\mathbf{1} = \frac{1}{(2N+1)^2}\sum_i [\boldsymbol{y}_k]_i$.

The GLR in k boils down to the difference between the reconstruction errors $E_0 = \frac{1}{2}||\boldsymbol{y}_k - \hat{r}_k \mathbf{1}||^2$ and $E_1 = \frac{1}{2}||\boldsymbol{y}_k - \widehat{\boldsymbol{r}}_{k,\theta}||^2$, as presented in **Figure** 5.6

Efficient implementation A more efficient way to compute this GLR has been presented in (Gasnier et al., 2021c) and used in our framework. It allows to compute the GLR value in every pixel without having to compute the estimated patch $\widehat{\boldsymbol{r}}_k(\hat{\theta})$ for $\mathcal{H}_1$. The result can then be improved by combining different scales, in the $[S_{min}, S_{max}]$ range.

Straightforward implementation A straightforward implementation of (5.9) requires computing, at each pixel of the $H \times W$ pixels SAR image, norms or scalar products of $\hat{r}_k$ and $\widehat{\boldsymbol{r}}_{k,\theta}$. The estimate $\hat{r}_k$ is obtained in $(2N+1)^2 = \mathscr{O}[N^2]$ multiplications. The estimate $\widehat{\boldsymbol{r}}_{k,\theta}$ requires $2(2N+1)^2\sqrt{2}(N+1) = \mathscr{O}[N^3]$ multiplications. The total cost for evaluating log $\mathrm{GLR}_{k,\theta}$ at all pixels and for T angles θ is thus $\mathscr{O}[WHTN^3]$. Such an implementation would be too slow for practical use.

Improved implementation We show that the algorithmic complexity can be reduced to $\mathscr{O}[WHTN^2 \log(WH)]$ (and even to $\mathscr{O}[WHTN \log(WH)]$ if the constraint (5.6) is dropped) with discrete correlations computed in Fourier domain using fast Fourier transforms.

Note that any product of the form $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{y}_k$ corresponds to a 2D discrete correlation of the log-transformed image $\boldsymbol{y}$ with the 2D filter whose 1D representation in lexicographic order is $\boldsymbol{w}$: $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{y}_k = [\mathrm{correl}(\boldsymbol{y}, \boldsymbol{w})]_k$. This correlation can be computed efficiently using 2D fast Fourier transforms: $\mathrm{correl}(\boldsymbol{y}, \boldsymbol{w}) = \mathtt{FFT}_{2\mathrm{D}}^{-1}\left[\mathtt{FFT}_{2\mathrm{D}}(\boldsymbol{y}) \cdot \mathrm{conj}\left(\mathtt{FFT}_{2\mathrm{D}}(\boldsymbol{w})\right)\right]$ where the conjugate operation conj() and the product $\cdot$ are performed element-wise.
The first term $\frac{1}{2}||\hat{r}_k\mathbf{1}||^2$ in equation (5.9) corresponds to $\frac{1}{2(2N+1)^2}[\mathrm{correl}(\boldsymbol{y}, \mathbf{1})]_k^2$, the second term $-\boldsymbol{y}_k^{\mathrm{T}}\hat{r}_k\mathbf{1}$ to $-\frac{1}{(2N+1)^2}[\mathrm{correl}(\boldsymbol{y}, \mathbf{1})]_k^2$, their sum is thus equal to $-\frac{1}{2(2N+1)^2}[\mathrm{correl}(\boldsymbol{y}, \mathbf{1})]_k^2$, which can be computed in $\mathscr{O}[WH \log(WH)]$ operations with fast Fourier transforms.
The third and fourth terms require the computation of $\boldsymbol{M}_\theta$ and the estimation of the profile $\widehat{\boldsymbol{p}}_{k,\theta}^{+}$. The i-th element of $\widehat{\boldsymbol{p}}_{k,\theta}$ corresponds to the product of the i-th row of matrix $\boldsymbol{M}_\theta^{\mathrm{pinv}}$ and the log-transformed data $\boldsymbol{y}_k$, which can be expressed using a discrete correlation: $[\widehat{\boldsymbol{p}}_{k,\theta}]_i = [\mathrm{correl}(\boldsymbol{y}, [\boldsymbol{M}_\theta^{\mathrm{pinv}}]_{i,\bullet})]_k$, where the notation $[\boldsymbol{M}_\theta^{\mathrm{pinv}}]_{i,\bullet}$ indicates the i-th row of matrix $\boldsymbol{M}_\theta^{\mathrm{pinv}}$. All profiles $\widehat{\boldsymbol{p}}_{k,\theta}^{+}$ can thus be computed in $\mathscr{O}[WHTN \log(WH)]$ operations. Rather than computing $\widehat{\boldsymbol{r}}_{k,\theta}$ for all θ and k before deriving $\boldsymbol{y}_k^{\mathrm{T}}\widehat{\boldsymbol{r}}_{k,\theta}$, it is more efficient to compute $(\boldsymbol{M}_\theta^{\mathrm{T}}\boldsymbol{y}_k)^{\mathrm{T}}\widehat{\boldsymbol{p}}_{k,\theta}^{+}$ since $[\boldsymbol{M}_\theta^{\mathrm{T}}\boldsymbol{y}_k]_i = [\mathrm{correl}(\boldsymbol{y}, [\boldsymbol{M}_\theta]_{\bullet,i})]_k$. The third term is thus obtained in an additional $\mathscr{O}[WHTN \log(WH)]$ operations. The computation of the fourth term is the most costly. To reduce the cost, we use the singular value decomposition (SVD) of matrix $\boldsymbol{M}_\theta$: $\boldsymbol{M}_\theta = \boldsymbol{U}_\theta\boldsymbol{S}_\theta\boldsymbol{V}_\theta^{\mathrm{T}}$ where $\boldsymbol{U}_\theta$ is a $(2N+1)^2 \times (2N+1)^2$ unitary matrix, $\boldsymbol{V}_\theta$ is a $\sqrt{2}(N+1) \times \sqrt{2}(N+1)$ unitary matrix, and $\boldsymbol{S}_\theta$ is rectangular with zeros outside the main diagonal. The expansion $\|\boldsymbol{M}_\theta\widehat{\boldsymbol{p}}_{k,\theta}^{+}\|^2 = \widehat{\boldsymbol{p}}_{k,\theta}^{+\mathrm{T}}\boldsymbol{M}_\theta^{\mathrm{T}}\boldsymbol{M}_\theta\widehat{\boldsymbol{p}}_{k,\theta}^{+} = \widehat{\boldsymbol{p}}_{k,\theta}^{+\mathrm{T}}\boldsymbol{V}_\theta\boldsymbol{S}_\theta^2\boldsymbol{V}_\theta^{\mathrm{T}}\widehat{\boldsymbol{p}}_{k,\theta}^{+}$ shows that

$$\|\widehat{\boldsymbol{r}}_{k,\theta}\|^2 = \sum_{i=1}^{\sqrt{2}(N+1)} [\boldsymbol{S}_\theta]_{i,i}^2([\boldsymbol{V}_\theta]_{\bullet,i}^{\mathrm{T}}\widehat{\boldsymbol{p}}_{k,\theta}^{+})^2 \tag{5.10}$$

which can be computed for all k and all θ in $\mathscr{O}[WHTN^2]$ operations once $\widehat{\boldsymbol{p}}_{k,\theta}^{+}$ has been computed. In the absence of the non-linear thresholding operation (5.6), it would be possible to compute $\|\widehat{\boldsymbol{r}}_{k,\theta}\|^2$ by discrete correlations between the SAR image $\boldsymbol{y}$ and the columns of the SVD of matrix $\boldsymbol{M}_\theta\boldsymbol{M}_\theta^{\mathrm{pinv}}$ in $\mathscr{O}[WHTN \log(WH)]$ operations.

The complete algorithm has a complexity $\mathcal{O}[WHTN(N+\log(HW))]$ to compute the detection map. It is summarized below:

Algorithm to compute a line detection map

Input: $\boldsymbol{y}$ ($H \times W$ pixels SAR image)
Output: log GLR ($H \times W$ pixels detection map) {compute first two terms}
1. $\boldsymbol{c} \leftarrow \text{correl}(\boldsymbol{y}, \mathbf{1})$
2. **for** $k = 1$ to HW **do**
3. $\quad[\textbf{log GLR}]_k \leftarrow -\frac{1}{2(2N+1)^2}[\boldsymbol{c}]_k^2$
4. **end for**
{compute last two terms}
5. $\boldsymbol{d}_{\max} \leftarrow \mathbf{0}$ ($H \times W$ temporary map)
6. **for** $\theta = \theta_1$ to θ_T **do**
7. $\quad\boldsymbol{d} \leftarrow \mathbf{0}$ ($H \times W$ temporary map)
8. $\quad\{\boldsymbol{U}_\theta, \boldsymbol{S}_\theta, \boldsymbol{V}_\theta\} \leftarrow \texttt{SVD}(\boldsymbol{M}_\theta)$
9. $\quad$**for** $i = 1$ to $\sqrt{2}(N+1)$ **do**
10. $\quad\quad[\boldsymbol{p}]_{\bullet,i} \leftarrow \text{correl}(\boldsymbol{y}, [\boldsymbol{M}_\theta^{\text{pinv}}]_{i,\bullet})$
11. $\quad\quad[\boldsymbol{p}]_{\bullet,i} \leftarrow \max\big([\boldsymbol{p}]_{\bullet,i}, [\boldsymbol{p}]_{\bullet,1}\big)$
12. $\quad\quad\boldsymbol{d} \leftarrow \boldsymbol{d} + \text{correl}(\boldsymbol{y}, [\boldsymbol{M}_\theta]_{\bullet,i}) \cdot [\boldsymbol{p}]_{\bullet,i}$ (3rd term)
13. $\quad\quad\boldsymbol{t} \leftarrow \mathbf{0}$ (HW temporary array)
14. $\quad\quad$**for** $j = 1$ to $\sqrt{2}(N+1)$ **do**
15. $\quad\quad\quad\boldsymbol{t} \leftarrow \boldsymbol{t} + [\boldsymbol{V}_\theta]_{j,i}[\widehat{\boldsymbol{p}}]_{\bullet,j}$
16. $\quad\quad$**end for**
17. $\quad\quad\boldsymbol{d} \leftarrow \boldsymbol{d} - \frac{1}{2}[\boldsymbol{S}_\theta]_{i,i}^2 \boldsymbol{t}^2$ (4th term)
18. $\quad$**end for**
19. $\quad\boldsymbol{d}_{\max} \leftarrow \max(\boldsymbol{d}_{\max}, \boldsymbol{d})$
20. **end for**
21. **log GLR** $\leftarrow$ **log GLR** $+ \boldsymbol{d}_{\max}$

The computational cost can be further decreased by computing the FFT of the image only once for all orientations θ. This reduces the cost of all the subsequent convolutions in the Fourier domain that involve the image (lines 10 and 12 of the algorithm). However, the computing speed could still be dramatically improved by using parallel processing.

5.2.2.3 Results for the linear structures detector

An example of response of the linear structures detector, combining the results for different scales is presented **Figure** 5.7.

This section presents the results of the proposed line detector applied to various Sentinel-1 SAR images. The results are compared to the response of the linear structures detector presented in (Tupin et al., 1998). The results presented here have been obtained using our line detector [1] with a symmetry constraint on three scales (with $S_{min} = 1$ and $S_{max} = 3$ which correspond to rescaling factors of 3, 2, and 1) and summing the results.

All images are SWOT simulated coherent power images or Sentinel-1 high-resolution GRD images acquired in IW mode with dark linear structures corresponding to rivers. **Figure 5.8** shows a comparison of the two detectors on linear structures corresponding to the Esk River near Carwinley (United Kingdom). On **Figure 5.9** the linear structures correspond to the Vilaine and Oust rivers near Redon (France). On **Figure 5.12** the linear structures correspond to the Loire river in Angers (France) and to smaller rivers nearby. **Figure 5.10** shows one example over the city of Des Moines (Iowa, USA) with the Racoon River.

[1] The code of the line detector is available at https://gitlab.telecom-paris.fr/ring/glrt_based_lines_detector

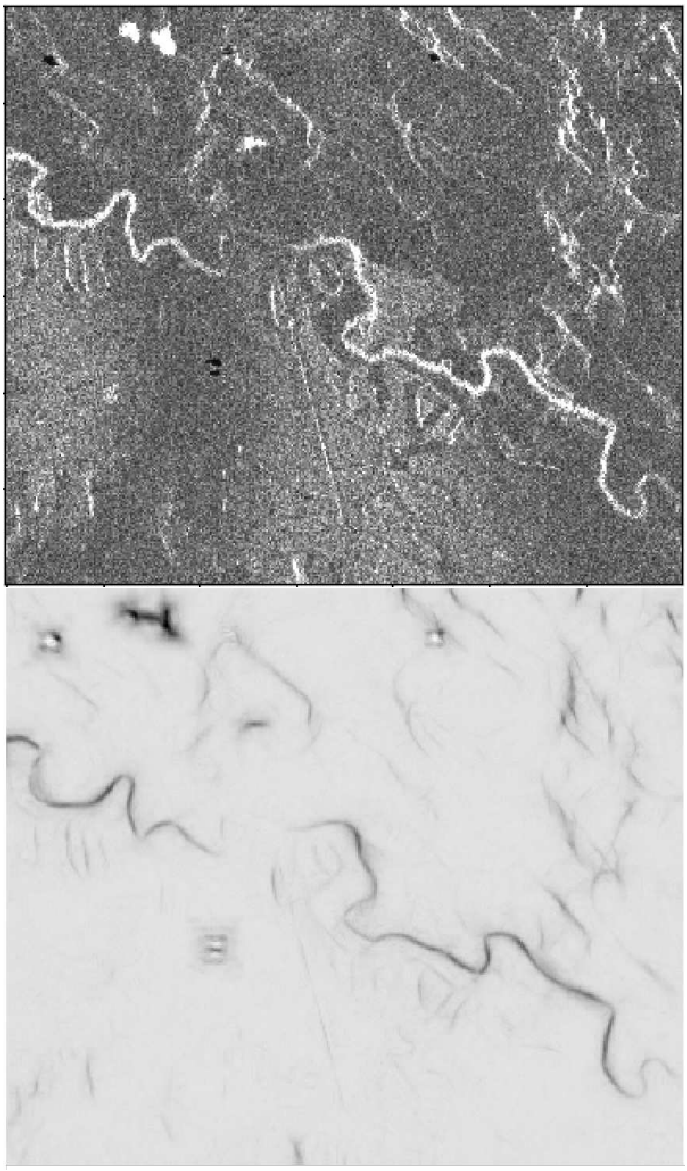


Figure 5.7 – Simulated SWOT image and linear structure detector response, combining the results for scales 1, 1/2 and 1/3. The response is displayed with inverted gray scale for better visualization.

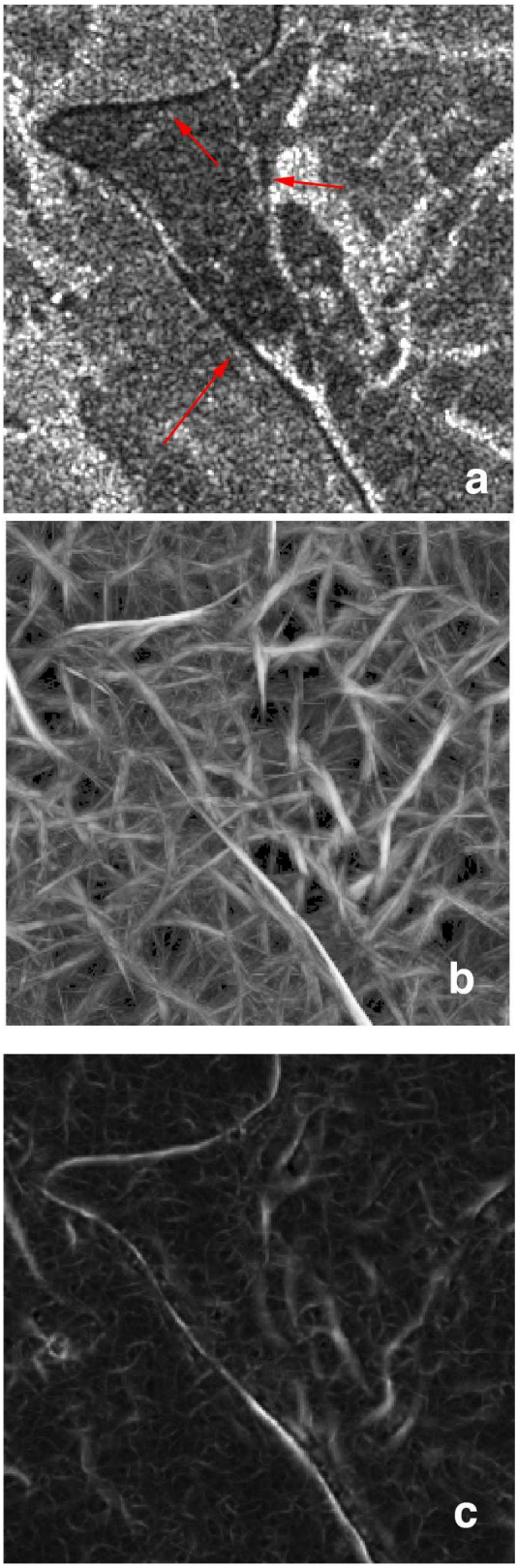


Figure 5.8 – Comparison for a Sentinel-1 GRD image in Gretna between the proposed detector (c) and the state-of-the-art detector (b) for one GRD image with linear structures highlighted by red arrows (a).

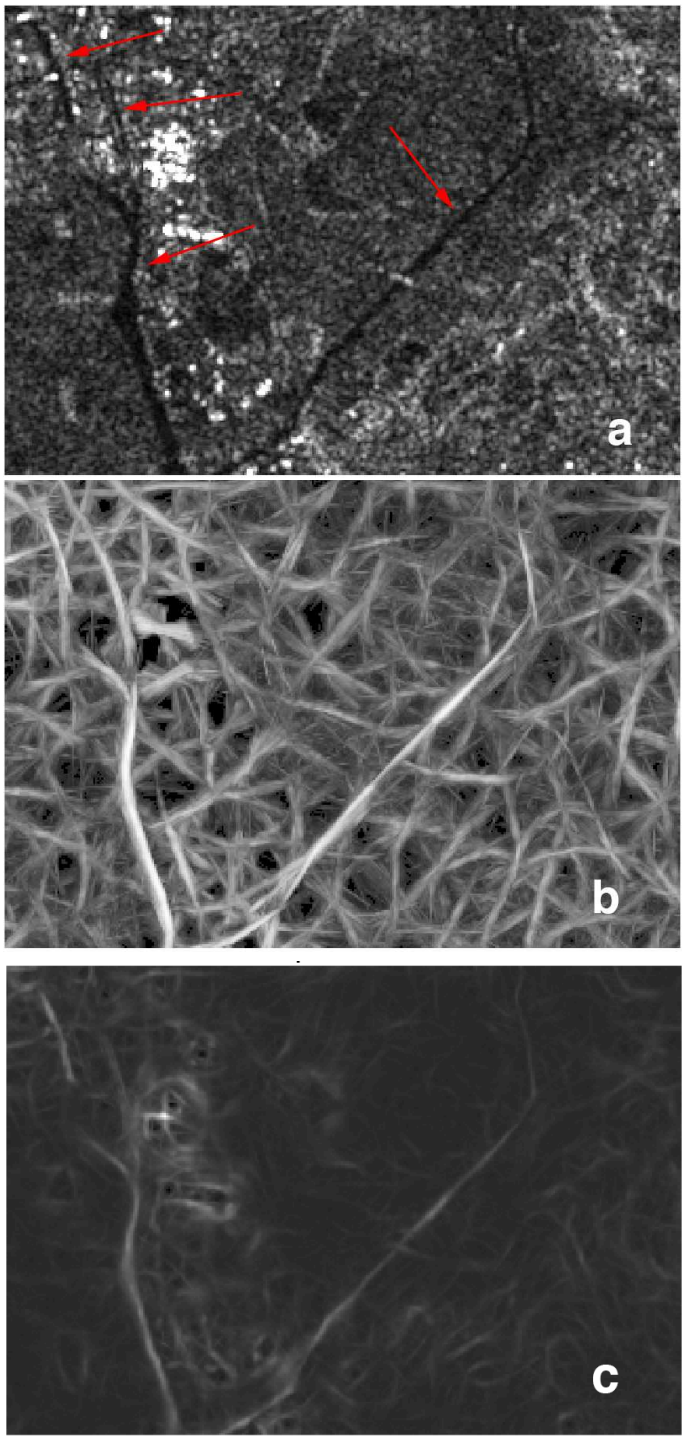


Figure 5.9 – Comparison for a Sentinel-1 GRD image in Redon between the proposed detector (c) and the state-of-the-art detector (b) for one GRD image with linear structures highlighted by red arrows (a).

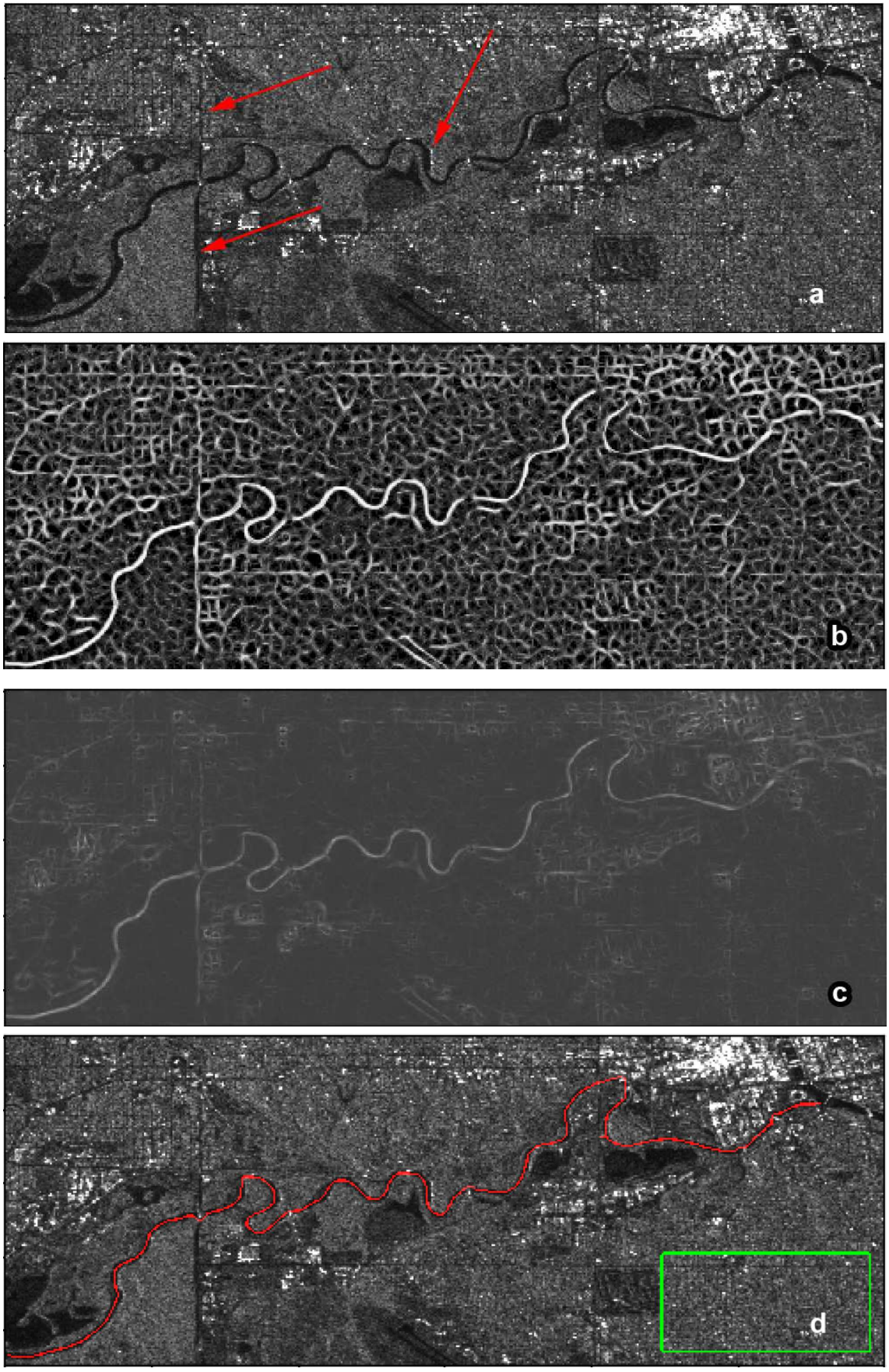


Figure 5.10 – Comparison between the proposed detector (c) and the state-of-the-art detector (b) for one GRD image with linear structures highlighted by red arrows (a). (d) shows the "line" area (red line) and the "no line" area (green rectangle) used to compute the ROC curve.

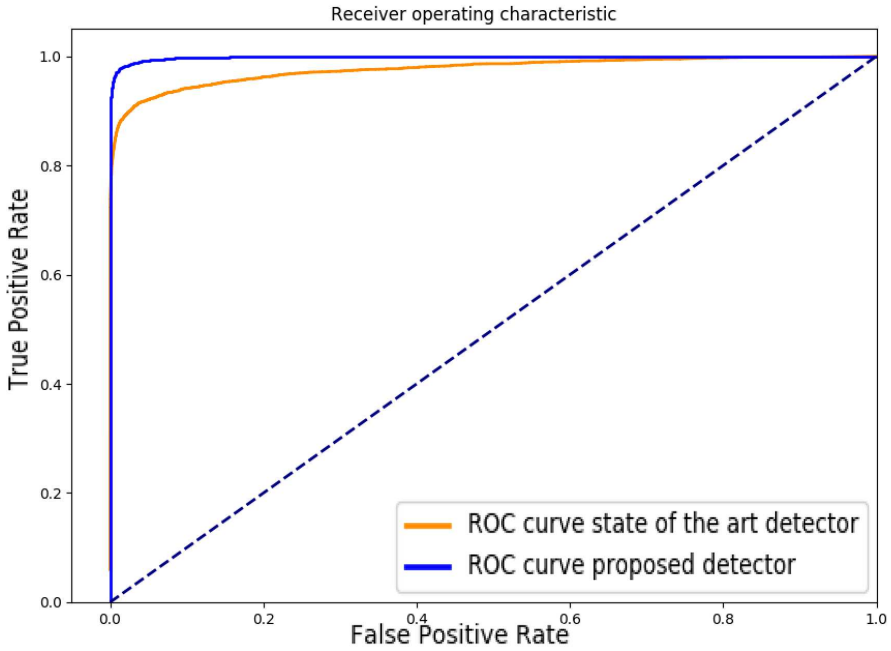


Figure 5.11 – ROC curves for both state-of-the-art detector (orange) and proposed detector (blue)

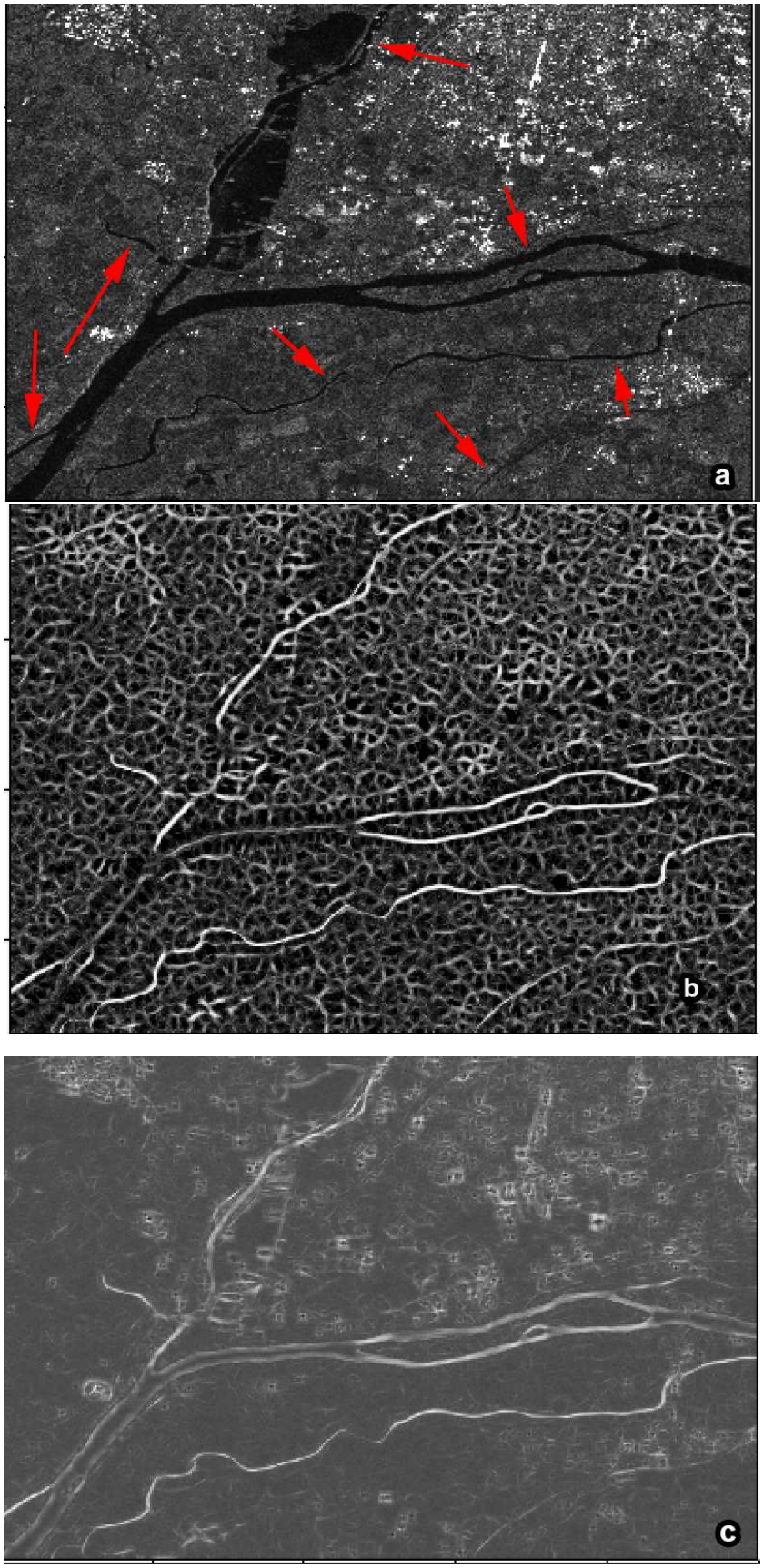


Figure 5.12 – Comparison between the proposed detector (c) and the state-of-the-art detector (b) for one Sentinel-1 GRD image with linear structures highlighted by red arrows (a).

Quantitative comparison A ground truth (d) for the "line" class (red line) and for the "no line" class (green rectangle) has been used to draw the receiver operating curves (ROC) for both detectors and is presented **Figure 5.11**. The ROC curve of both detectors is created by plotting their True Positive Rate (TPR) against their False Positive Rate (FPR) for multiple threshold values.

On the ROC curves, the proposed detector is better than the state-of-the-art detector, as for any given false positive rate, its true positive rate is higher.

Conclusion The proposed algorithm response clearly has fewer false positive while maintaining a good detection of the linear structures. More importantly, the artifacts created by the proposed method are not line-shaped unlike those of the reference method and will be less troublesome for the following steps of the method.

Summary: Linear structures detector

We proposed a new linear structures detector based on the generalized likelihood ratio (GLR). Its response corresponds to the generalized likelihood ratio (GLR) which boils down to the difference between the reconstruction error E_1 and E_2 for a local patch under two hypothesis (see 5.6):

- $\mathcal{H}_0$: there is no linear structure
- $\mathcal{H}_1$: there is a linear structure

We proposed an efficient algorithm to compute this difference indirectly. This new method produced better results than the reference linear structures detector (Tupin et al., 1998), both qualitatively and quantitatively.

5.2.3 Accurate centerline determination using least-cost path algorithm

The second step of the algorithm is to retrieve the actual centerline of the river reach using both the response of the linear structure detector and prior information on the river position. The external database that we use (GRWL) provides for each river reach (about 10 km long) nodes that are 200 m apart along the centerline. From this approximate centerline, at least two approaches can be considered to obtain the actual centerline:

- To apply an active contour approach such as snake (Kass et al., 1988) on the entire centerline using the detector response.
- To consider only some nodes in the centerline and to compute the minimum cost path between pairs of nodes on a cost image derived from the detector response.

A major issue with the snakes approach for this application is its sensitivity to the initialization and to the parameters that determine the evolution of the active contour. A preliminary study showed the difficulty to choose the right parameters and the lack of stability of the results. The proposed method is based on a minimum path between a subset of nodes of the centerlines using Dijkstra's algorithm. A similar method has been proposed by Dillabaugh et al. (Dillabaugh et al., 2002) for optical images, with user-specified start and end points. An overview of this second step of the proposed method is given by **Figure** 5.16.

We define the cost $\boldsymbol{C}(x, y)$ at every pixel (x, y) based on the line detector response $\boldsymbol{D}(x, y)$ as:

$$\boldsymbol{C}(x, y) = [1 - \boldsymbol{D}(x, y)/D_{max}]^{N_{pow}} \tag{5.11}$$

with D_{max} the maximum value of the detector response $\boldsymbol{D}$ on the whole image and N_{pow} a tuning parameter. N_{pow} adjusts the cost of crossing a pixel whose detector response is not maximal. It has to be high enough to penalize short paths that cut through a meander but not too high either to prevent the risk of being diverted by a road with a strong line detector response or having numerical computational issues.

In the situation where one or both nodes are outside of the river, and provided N_{pow} is high enough, the least-cost path is expected to go from one node to the other through the river via the minimum cost path, as presented in blue between nodes $B1$ and $B2$ in **Figure** 5.13. This approach is robust to situations where the *a priori* nodes are far away from the actual river (due to changes in the actual river or to projection errors). This has been assessed using nodes with a very exaggerated shift from the true position (over 1 km) in **Figure** 5.13 (and for other Sentinel-1 images in the supplementary materials). We see that the center part of the river segment is here correctly detected, but that close to node B1 an erroneous path has been chosen. This generally occurs in the presence of strong noise or when there are other linear structures in the area.

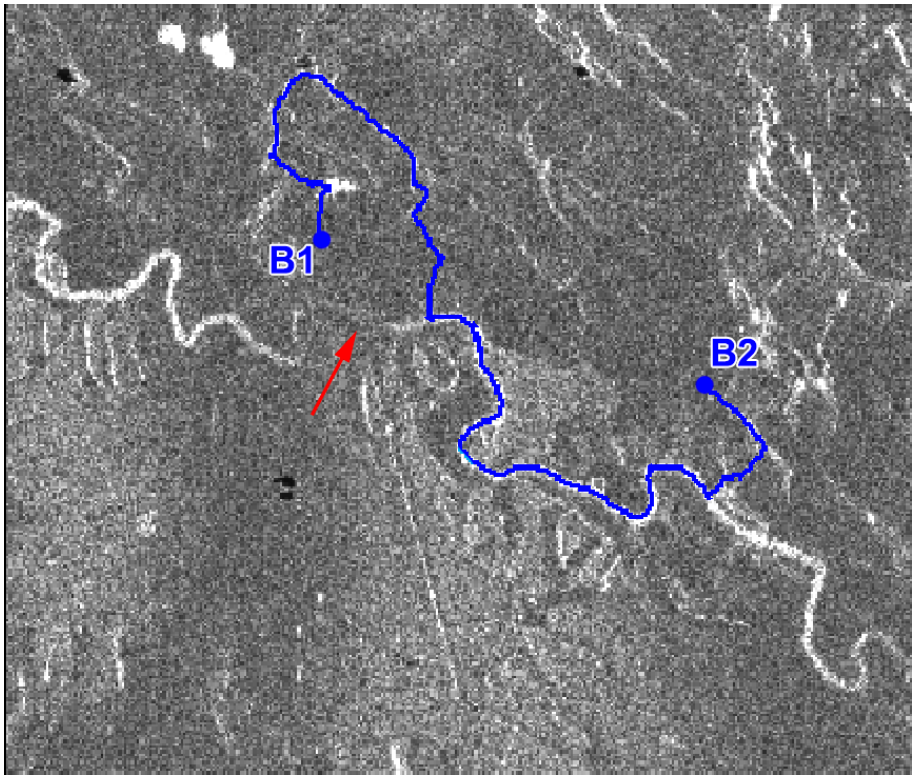


Figure 5.13 – Shortest path determination between nodes $B1$ and $B2$ displayed on the original image. The red arrow is pointing to the part of the river that has been missed by the detection. Indeed linear structures that do not correspond to the river caused the centerline to circumvent this part of the river.

To cope with this issue and in order to retrieve the entire centerline, we propose to use overlapping pairs of nodes as extremities for the minimum cost path search. Recall that GRWL has a node every 200 m, whereas the pairs of points that we use are in the order of 1 to 10 km apart. By combining the results for each pair of nodes (for example, the green, blue, and magenta lines in **Figure** 5.14), we obtain the estimated centerline for the whole reach plus one off-river branch between the centerline and every *a priori* node that does not belong to the actual centerline.

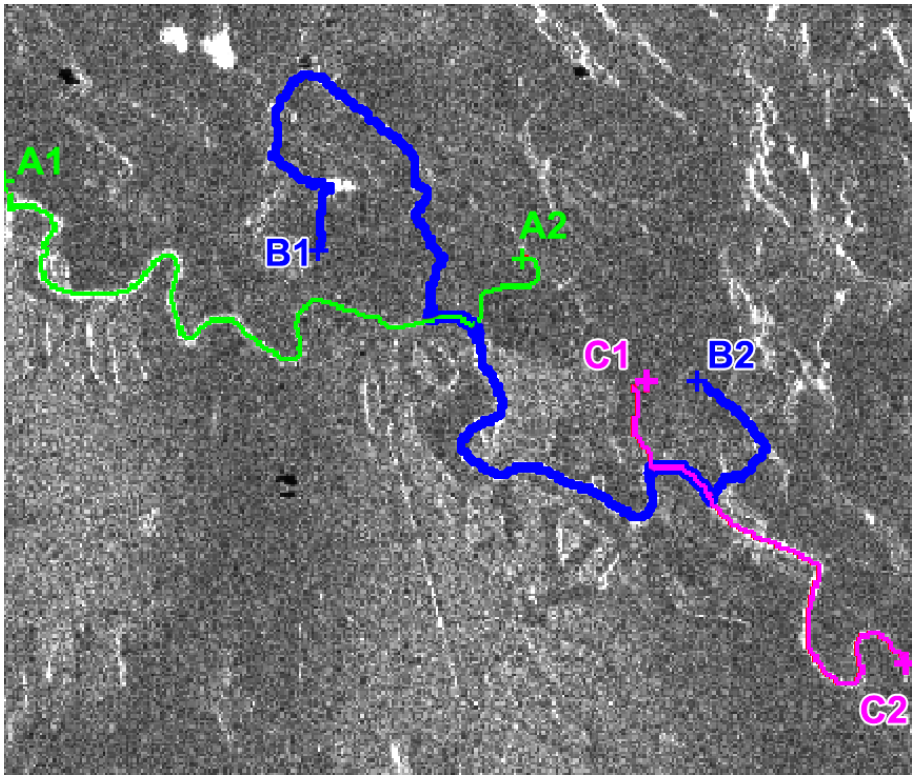


Figure 5.14 – Visualization on the same image of the result of the least-cost paths for 3 pairs of nodes: $A1 \rightarrow A2$ in green, $B1 \rightarrow B2$ in blue and $\boldsymbol{C}1 \rightarrow C2$ in magenta. The centerlines have been widened for better visualization. In this example, the *a priori* nodes have been chosen excessively far from the river to illustrate the robustness of the proposed approach.

The off-river branches can be easily eliminated using a pruning method. Because of the overlap of the reach nodes, only the pixels on the least-cost path between the end nodes of a reach and the previous reach are kept in the final central line. **Figure** 5.15 shows the result of the pruning of the centerlines in **Figure** 5.14. The final centerline for each river is then stored as a boolean raster $\boldsymbol{C}_L$ of the same size as the image that takes the value 1 on the centerline and 0 elsewhere.

Summary: Accurate centerline determination using least-cost path algorithm

We proposed a simple yet robust way to detect the river centerline on a linear structures detector response with the help of a set of overlapping pairs of nodes. It is based on a least-cost path computation on a cost array between pairs of nodes, followed by a pruning step.

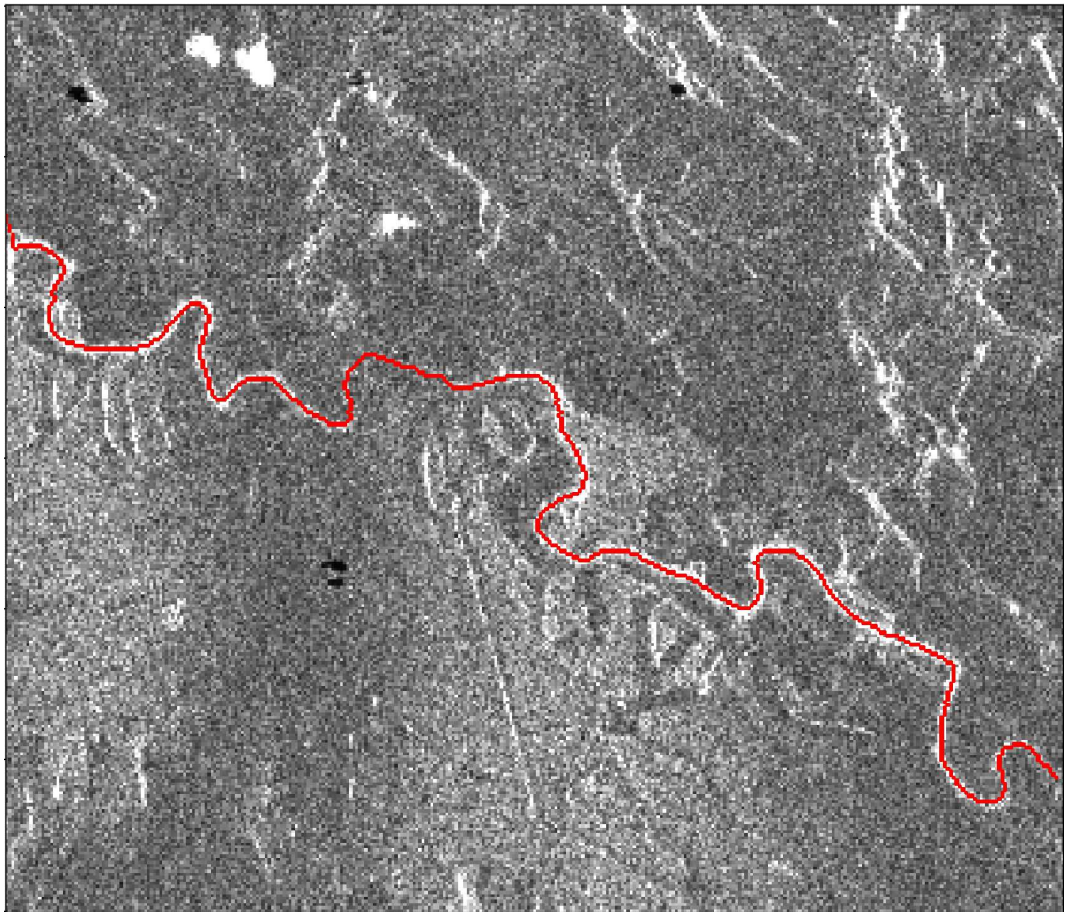


Figure 5.15 – Centerline obtained after pruning of the previous result. The centerline has been widened for better visualization.

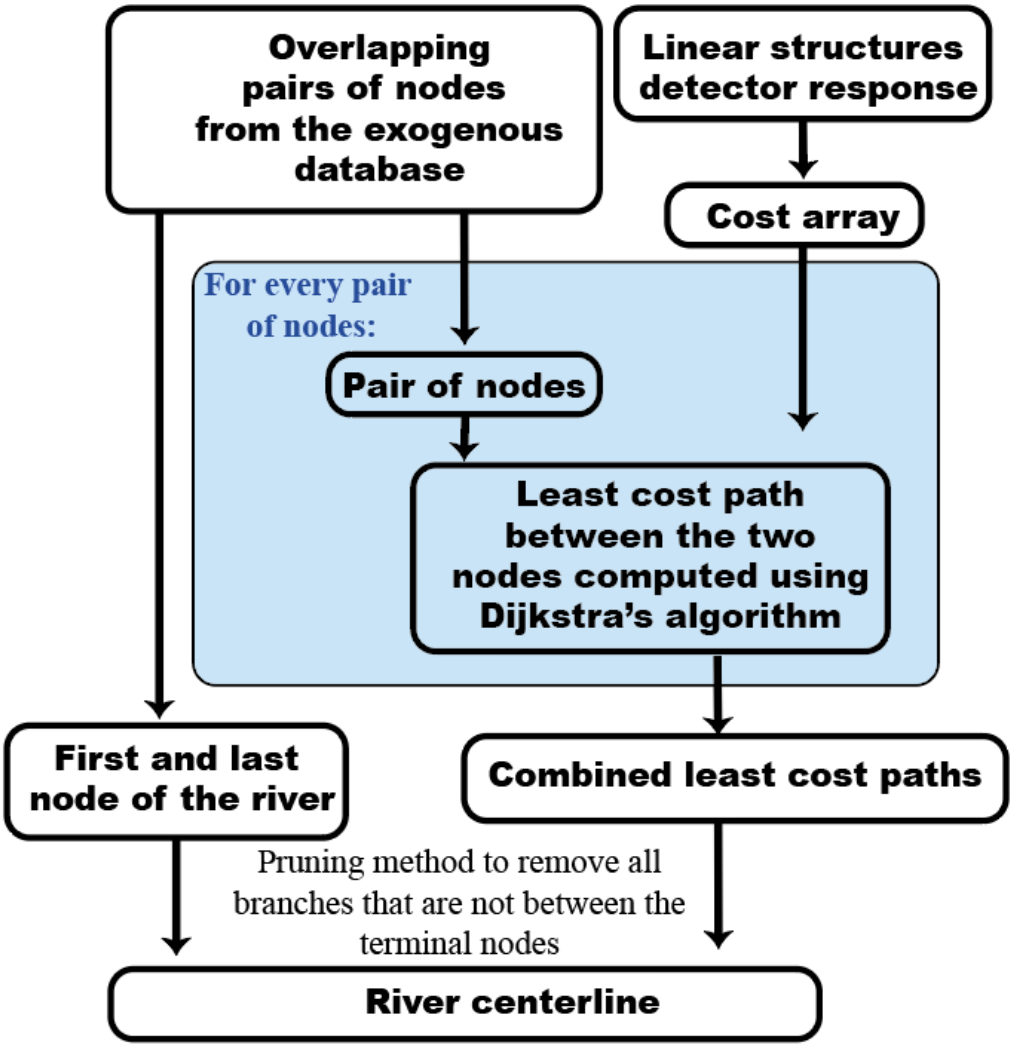


Figure 5.16 – Flowchart describing the second step of the algorithm that uses the previously computed linear structures detector response and nodes from the database to compute the river centerline.

5.2.4 Segmentation of the reach from the centerline by conditional random field

The last step of the proposed method is to get an accurate segmentation of the river reach using the previously estimated centerline and the SAR image. This can be considered as a region growing problem around the estimated centerline taking into account the intensities in the SAR image. Random walk (Grady, 2006) using the centerline as a seed, morphological approaches or graph-cut MRF approaches (Boykov and Jolly, 2001) with hard constraints could be relevant for this problem, but we did not obtain satisfactory results with these.

Instead, we propose an innovative method based on a conditional random field (CRF) (Lafferty et al., 2001). The problem is expressed as the minimization of a global energy function $\mathscr{E}$ that takes both the SAR image and the centerline into account, with an adapted regularization that does not over-penalize narrow rivers. An overview of this method is given by the flowchart in **Figure** 5.20.

The global energy $\mathscr{E}$ which depends on the classification $\boldsymbol{\ell}$ ($\boldsymbol{\ell}(k) = 1$ for water and $\boldsymbol{\ell}(k) = 0$ for land), is the sum of two data terms, a regularization term, and a flux term:

$$\mathscr{E}(\boldsymbol{\ell}, \boldsymbol{I}) = U^I_{\text{data}}(\boldsymbol{\ell}, \boldsymbol{I}) + U^C_{\text{data}}(\boldsymbol{\ell}, C_L) + U_{\text{reg}}(\boldsymbol{\ell}, \boldsymbol{I}) + U_{\text{flux}}(\boldsymbol{\ell}, \boldsymbol{I}). \tag{5.12}$$

The two data terms are U^I_{data} that ensures fidelity with the image intensity $\boldsymbol{I}$ and U^C_{data} that ensures that the centerlines retrieved in the previous step are classified as water. The regularization term U_{reg} is adapted to the segmentation of narrow rivers. Along with this adapted term, we propose a term U_{flux} whose role is to favor a longer water/land contour if this segmentation is in better agreement with the gradients of the SAR image (i.e., to counter-balance the effect of the term U_{reg} that encourages a short contour length).

The image data term U^I_{data} is based on a model that considers two likelihoods: a likelihood that depends on the intensity of the image for the water class and a likelihood that is intensity-invariant for the land class. The likelihood for the water class is based on a gamma distribution (A.2) for the intensity, with two parameters: R_1 for the water reflectivity that is supposed to be homogeneous and L for the number of looks. The reflectivity of water R_1 can be estimated using a debiased geometric mean estimator $\widehat{R_1}$ on the intensity $\boldsymbol{I}$ for every pixel belonging to the centerline. In order to increase robustness, the brightest pixels (for Sentinel-1) that can correspond to bridges or boats can be excluded from the computation of the mean. With these variables, and under the homogeneous water hypothesis, the theoretical distribution of intensity for water is given by:

$$\mathrm{p}(I|R_1) = \frac{L^L I^{L-1}}{\Gamma(L) R_1^L} \exp\left(-L\frac{I}{R_1}\right)., \tag{5.13}$$

where I stands for $\boldsymbol{I}(k)$ for any pixel k. The neg-log-likelihood $\mathscr{L}_1$ for the water class ($\boldsymbol{\ell} = 1$) is then:

$$\mathscr{L}_1(\boldsymbol{I}|R_1) = K(R_1, L) + \frac{L\boldsymbol{I}}{R_1} + (1 - L) \cdot \log(\boldsymbol{I}) \tag{5.14}$$

where $K(R_1, L) = \log(\Gamma(L)) + L \cdot \log(R_1) - L \cdot \log(L)$

For the land class, in the absence of a model for the distribution of the land class, we consider a uniform (homogeneous) likelihood. The constant likelihood value $\mathscr{L}_0$ is chosen so that the data

energy of one well-classified pixel (i.e. its neg-log-likelihood) is equal in expectation for both classes:

$$\mathbb{E}_{\text{True land}}[\mathscr{L}_0(\boldsymbol{I}(k))] = \mathbb{E}_{\text{True water}}[\mathscr{L}_1(\boldsymbol{I}(k), R_1)] \tag{5.15}$$

Provided that the estimator for water reflectivity $\widehat{R_1}$ is accurate enough, the choice of an homogeneous log-likelihood $\mathscr{L}_0 = \mathbb{E}_{\boldsymbol{I}|R_1}[\mathscr{L}_1(\boldsymbol{I}, R_1)]$, with the expected value computed over the water pixels, prevents the classification from being biased towards land. This brings the following expression for $\mathscr{L}_0$:

$$\mathscr{L}_0 = K(R_1, L) + L + (L-1)(\log(\frac{L}{R_1}) - \Psi(L))\,. \tag{5.16}$$

In order to simplify $\mathscr{L}_1$ and $\mathscr{L}_0$, the constant value $K(R_1, L)$ can be subtracted from both neg-log-likelihoods.

For an elementary surface of the image du centered at u, the image data energy is defined by $U^I_{\text{data}}(du) = \boldsymbol{\ell}(u) \cdot \mathscr{L}_1(\boldsymbol{I}, R_1, L)du + (1 - \boldsymbol{\ell}(u)) \cdot \mathscr{L}_0 du$. Another energy term U^C_{data} ensures that the previously determined centerlines are classified as water. It penalizes by a large value of $K_C \cdot du$ the missclassification as land of any elementary surface du that belongs to a centerline ($\boldsymbol{C}_L(du) = 1$). This energy term is given by $U^C_{\text{data}}(du) = K_C \cdot (1 - \boldsymbol{\ell}(du)) \cdot C_L(du)du$.

Finally, a regularization term ensures that the transitions between water and land are compatible with the gradients of the image, by penalizing the transitions that would occur where the gradient magnitude is low, or if the boundaries are not orthogonal to the gradient direction.

We want to minimize over the water boundaries the weighted total variation on the label field $\boldsymbol{\ell}$ that we assume to be continuous and whose spatial gradient at location u is $\|\overrightarrow{\nabla \boldsymbol{\ell}}(u)\|$:

$$U_{\text{reg}}(\boldsymbol{\ell}) = \beta \int_{u \in \mathbb{R}^2} w_{asym}(u) \|\overrightarrow{\nabla \boldsymbol{\ell}}(u)\| du. \tag{5.17}$$

The total variation is weighted with

$$w_{asym}(u) = \exp(-[\overrightarrow{\nabla \boldsymbol{\ell}}(u) \cdot \overrightarrow{\nabla \boldsymbol{I}}(u)]_+ / \lambda). \tag{5.18}$$

This asymmetric weighting w_{asym} favors location of the boundaries that are aligned with the strong gradients of the image. The notation $[x]_+$ returns x if $x > 0$ and 0 otherwise. The variable λ and β are parameters that allow adjusting the regularization and its sensitivity to the gradients.

It can be noted that for sensors with dark rivers on a bright background such as Sentinel-1 or TerraSAR-X, the negative of the gradient $-\overrightarrow{\nabla \boldsymbol{I}}(u)$ should be used instead to segment the rivers.

To prevent transitions from being encouraged by gradient artifacts caused by speckle noise, we use a gradient adapted to SAR images called Gradient by Ratio (GR) proposed by Dellinger et al. (Dellinger et al., 2015), which is an adaptation of ROEWA (Ratio of Exponentially Weighted Average) proposed by Fjørtoft et al. (Fjortoft et al., 1998). It computes at each pixel the gradients in the horizontal and vertical direction, as presented in **Figure** 5.17.

The former regularization term U_{reg} can cause excessive regularization especially in low contrast situations and lead to false positives and false negatives in detection. For example in SWOT images, a bright sand river inner bank in a meander, also called a point bar (visible in **Figure**

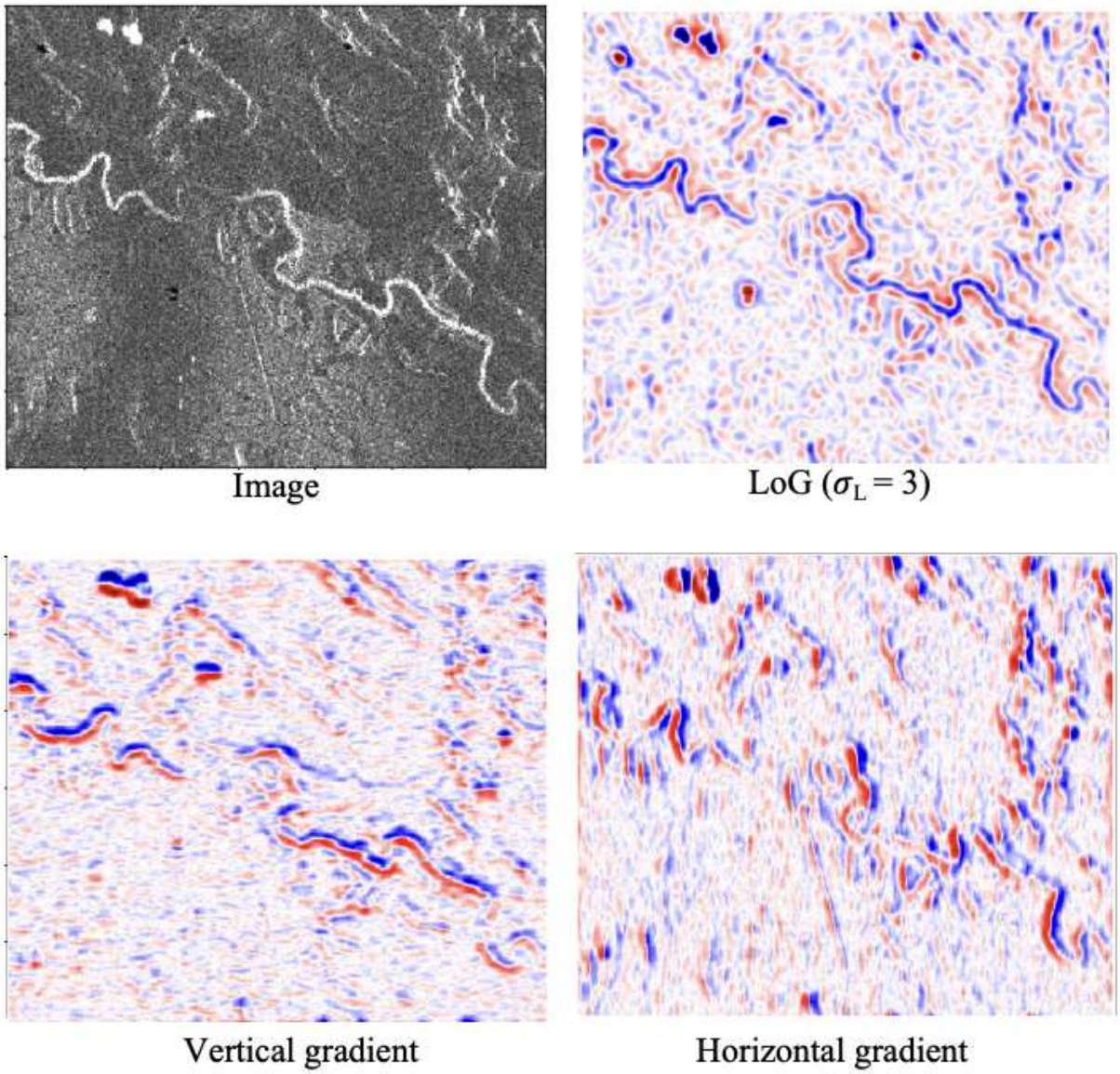


Figure 5.17 – Simulated SWOT image, its Laplacian of Gaussian (LoG) and its gradients. The positive values are displayed in red, the negative values are displayed in blue.The gradients have been computed with ROEWA gradient by ratio approach with a weighting parameter $\alpha = 2.4$ which is a good compromise between smoothing and location for $L = 4$. The LoG have been computed with $\sigma_L = 3$

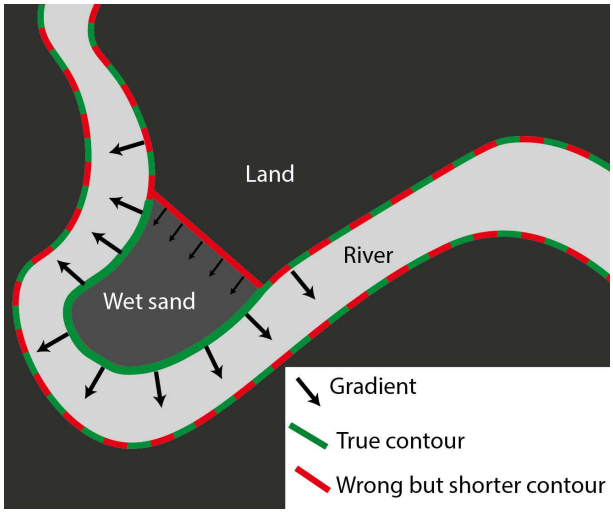


Figure 5.18 – Illustration of a situation in which using solely the total variation regularization may lead to an erroneous contour detection.

5.18) can be erroneously classified as water. Conversely, in the case of a river with an irregular width, the regularization can lead to an incorrect estimation of the width. To cope with these problems that are caused by the regularization that favors shorter water-land boundaries over longer ones despite the weaker gradient, we introduce an additional term that favors longer boundaries co-located with strong gradients.

The boundaries of the river are expected to be located where the gradient of the SAR image is the strongest within a small neighborhood and to be oriented orthogonally to the gradient. Over the boundary $\partial\{\boldsymbol{\ell} = 1\}$ between land ($\boldsymbol{\ell} = 0$) and water ($\boldsymbol{\ell} = 1$), this criterion locally corresponds to maximizing the dot product between the gradient $\overrightarrow{\nabla}\boldsymbol{I}(u)$ and the unit normal vector of the segmentation $\{\boldsymbol{\ell} = 1\}$. Over the whole river, the criterion can be expressed as the outward flux Φ of the gradient through the boundary $\partial\{\boldsymbol{\ell} = 1\}$

$$\begin{aligned}\Phi &= \oint_{u\in\partial\{\boldsymbol{\ell}=1\}} \overrightarrow{\nabla}\boldsymbol{I}(u)\cdot\overrightarrow{n}(u)dl \\ &= \iint_{\{\boldsymbol{\ell}=1\}} \overrightarrow{\nabla}\cdot\overrightarrow{\nabla}\boldsymbol{I}(u)\,\mathrm{d}u\end{aligned} \tag{5.19}$$

where the second line comes from Ostrogradsky's divergence theorem.

Here, the Laplacian of the image can be approximated with a Laplacian of Gaussian (LoG) operator of parameter σ_L

$$\overrightarrow{\nabla}\cdot\overrightarrow{\nabla}\boldsymbol{I} \approx \mathrm{LoG}(\boldsymbol{I}, \sigma_L) \tag{5.20}$$

that can be computed using a convolution. We call the resulting image of the LoG $\mathrm{LoG}_{\boldsymbol{I},\sigma_L}$.

The influence of the flux energy $U_{\text{flux}}(\boldsymbol{\ell})$ can be balanced with a multiplicative parameter η that adjusts its effect:

$$U_{\text{flux}}(\boldsymbol{\ell}) = \int_{u\in\mathbb{R}^2,\boldsymbol{\ell}(u)=1} \eta\cdot\mathrm{LoG}_{\boldsymbol{I},\sigma_L}(u)du. \tag{5.21}$$

The sign of η depends on the sensor: $\eta < 0$ for SWOT (water generally brighter than land) and $\eta > 0$ for Sentinel-1 (land mostly brighter than water).

By combining the four terms: U^I_{data}, U^C_{data}, U_{reg}, U_{flux} of $\mathscr{E}$, we can write the segmentation problem as a minimization problem:

$$\begin{aligned}\arg\min_{\boldsymbol{\ell}} &\int_{u\in\mathbb{R}^2,\boldsymbol{\ell}(u)=1} \mathscr{L}_1(\boldsymbol{I},R_1,L)+\eta\cdot \text{LoG}_{\boldsymbol{y},\sigma_L}(u)du \\ &+\int_{u\in\mathbb{R}^2,\boldsymbol{\ell}(u)=0} \mathscr{L}_0+C_L(u)\cdot K_C du \\ &+\beta\int_{u\in\mathbb{R}^2} w_{asym}(u)\|\overrightarrow{\nabla\boldsymbol{\ell}}(u)\|du.\end{aligned} \tag{5.22}$$

This equation can be discretized as

$$\begin{aligned}\arg\min_{\boldsymbol{\ell}} &\sum_k \boldsymbol{\ell}(k)(\mathscr{L}_1(\boldsymbol{I},R_1,L)+\eta\cdot \text{LoG}_{\boldsymbol{I},\sigma_L}(k) \\ &+(1-\boldsymbol{\ell}(k))(\mathscr{L}_0+C_L(i)) \\ &+\beta\sum_{k\sim k'} w_{asym}(k,k')\cdot|\boldsymbol{\ell}(k')-\boldsymbol{\ell}(k)|\end{aligned} \tag{5.23}$$

with $w_{asym}(k,k')=\exp(-[(\boldsymbol{\ell}(k')-\boldsymbol{\ell}(k))(\boldsymbol{I}(k')-\boldsymbol{I}(k))]_+/\lambda)$, where $k\sim k'$ means that k' is an 8-neighbor of k. In the case of pixels that are 8-neighbors of i but not 4-neighbors, λ is multiplied by $\sqrt{2}$. $(\boldsymbol{I}(k')-\boldsymbol{I}(k))$ is actually approximated using the ROEWA gradient for better robustness to noise.

The relationship between the local gradient value and the energy can be analyzed for a very simple situation by plotting the asymmetric regularization U_{reg} energy and the flux energy U_{flux} as it is done **Figure** 5.19. This figure considers the actual gradient and the actual flux, which are actually approximated respectively by the ROEWA gradient and the integral of the Laplacian of Gaussian. The cost for $\boldsymbol{\ell}(k_1)=1$ and $\boldsymbol{\ell}(k_2)=0$ is high for a null value of the gradient, and even higher if the gradient is in the wrong direction. In contrast, the higher the gradient in the right direction, the lower the cost. For a gradient in the right direction with a magnitude above 0.4, the cost is even negative.

The minimization problem presented in (5.23) can be solved using a minimal cut approach such as the one proposed by (Boykov and Kolmogorov, 2004), with asymmetric edges on a directed graph.

Summary: River segmentation around the centerline

We proposed a new conditional random field model (CRF) that combines 4 terms to obtain a segmentation of the river (label field $\boldsymbol{\ell}$) from its centerline C_L and the SAR intensity image $\boldsymbol{I}$:

- An image data term $U^I_{\text{data}}(\boldsymbol{\ell},\boldsymbol{I})$
- A centerline fidelity term $U^C_{\text{data}}(\boldsymbol{\ell},C_L)$
- A regularization term $U_{\text{reg}}(\boldsymbol{\ell},\boldsymbol{I})$
- A flux term $U_{\text{flux}}(\boldsymbol{\ell},\boldsymbol{I})$

This minimization problem can be solved using a graph cut algorithm.

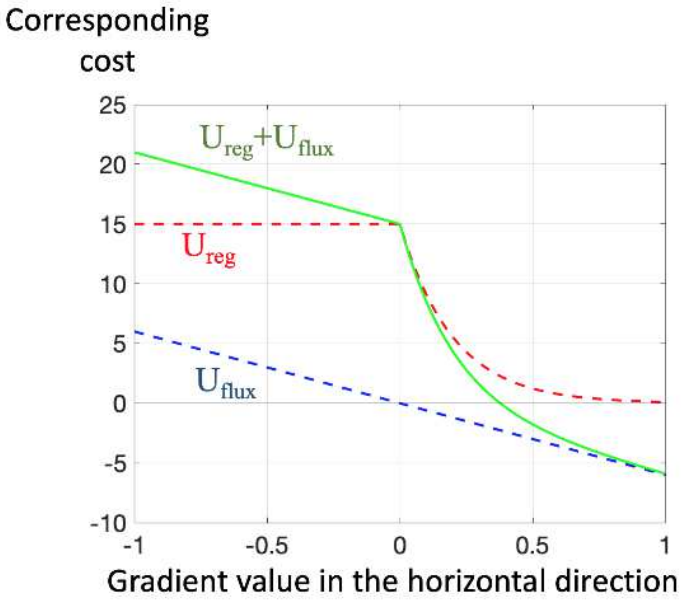


Figure 5.19 – Theoretical regularization energy U_{reg} energy, flux energy U_{flux}, and sum $U_{reg} + U_{flux}$ between a pixels k_1 and its right neighbor k_2 if $\boldsymbol{\ell}(k_1) = 0$ and $\boldsymbol{\ell}(k_2) = 1$ in the case of SWOT images, with the parameter values used for our experiments. The curves would be flipped horizontally if $\boldsymbol{\ell}(k_1) = 1$ and $\boldsymbol{\ell}(k_2) = 0$ or if dealing with Sentinel-1 images.

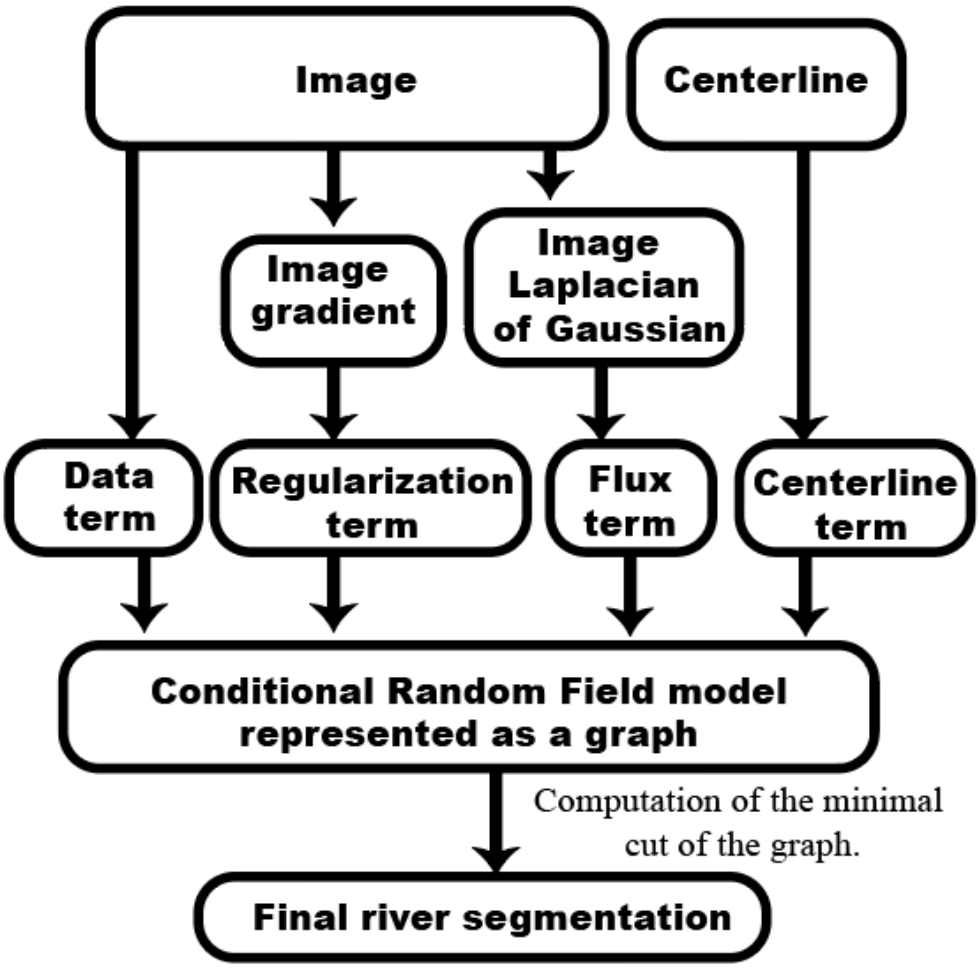


Figure 5.20 – Flowchart describing the third step of the algorithm that uses the previously computed centerline along with the SAR image to detect the river.

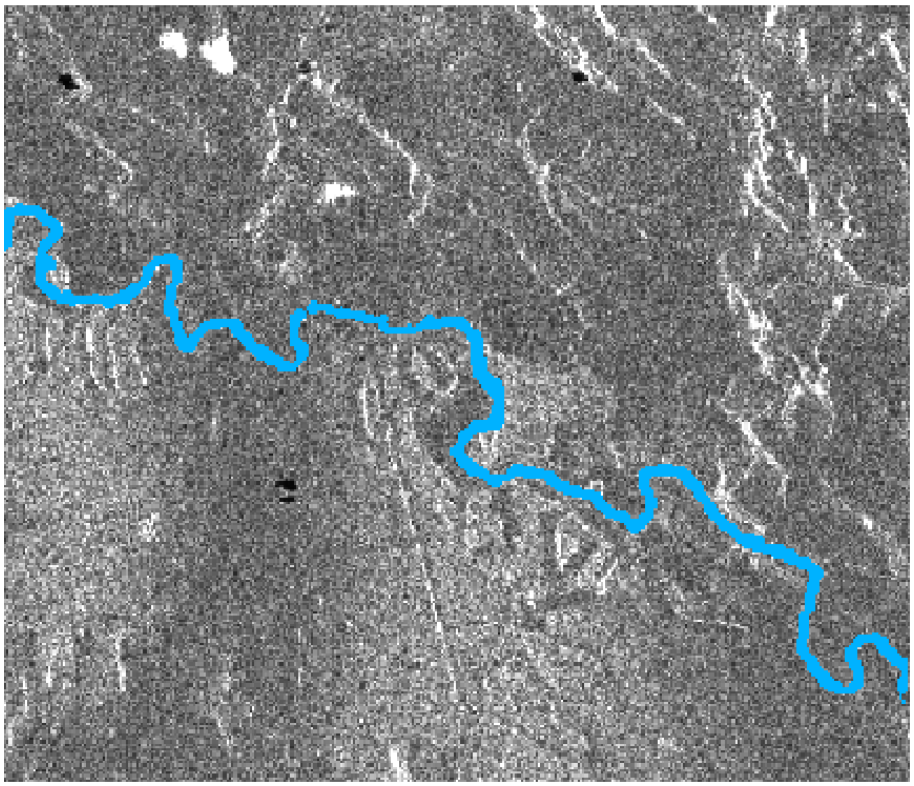


Figure 5.21 – Result of the CRF segmentation for the same SAR image as in Figs. 5.7 and 5.15

.

5.3 Experimental results

In this section, we evaluate the interest and characterize the performances of our method in segmenting small rivers in SAR images using a prior database, both for SWOT and Sentinel-1 images. Even if the images from the experimental dataset have been chosen to be as representative as possible of various situations, the comprehensive calibration of the algorithm on a specific sensor is beyond the scope of our experiments.

The results presented below have been obtained using our published code[2] that uses the PyMaxFlow[3] wrapper to an implementation of Vladimir Kolmogorov's graph cut solver presented in (Boykov and Kolmogorov, 2004).

5.3.1 Dataset

Our method has been tested on Sentinel-1 GRD images and on simulated SWOT HR coherent power images.

Sentinel-1 We used Sentinel-1 GRD images (more specifically Interferometric Wide Ground Range Detected High Definition images) in VV or VH polarization that have been downloaded from a Copernicus Open-Access mirror and cropped around the study area without orthorectification or calibration. Their statistics are presented in section 2.3.

We use 7 images in our Sentinel-1 dataset, presented in Table 5.1, corresponding to various examples of small rivers with different kinds of environments.

These images are associated with a ground truth that has been manually drawn on the SAR image using GIMP software, with the help of Open Street Map and optical images provided by Bing displayed over the SAR images with QGIS software to help to distinguish between actual rivers and other dark linear structures. This ground truth is not binary but classifies the pixels of the images into three classes: Land, Water, Uncertain classification. The Uncertain class corresponds to pixels for which it was not possible to determine whether or not it should belong to the river. We used it for our ground truth in four situations, as illustrated in **Figure** 5.22:

1. Isolated strong reflectors in rivers (most likely boats).
2. Bridges over rivers.
3. Small anabranches (diverging branches of a river, separated by an island, that re-enter the main stream downstream.
4. Flooded areas or small lakes that are only partially connected to a river.

[2]The code used for our experiments and all the images and ground truth for Sentinel-1 images is available : https://gitlab.telecom-paris.fr/ring/guided-river-detection

[3]http://pmneila.github.io/PyMaxflow/

Table 5.1 – Sentinel-1 GRD images used for our experiments

Image #-name	River name	Location	Date	Polarization	River width	Size (pixels)
1 - Des Moines	Racoon	Des Moines, Iowa, USA	2018-08-02	VH	40m - 120m	1313×1750
2 - Sunar	Sunar	Garhakota, Madhya Pradesh, India	2018-06-22	VH	40m - 150m	1026 × 923
3 - Gaoual	Tomine Koumba	Near Gaoual, Guinea	2018-07-15	VH	30m - 130m 30m - 130m	927×1854
4 - Angers	Maine Loire Louet anabranch	Angers, Pays de la Loire, France	2019-12-02	VV	100m - 150m 200m - 1000m 25m-120m	927×1854
5 - Garonne	Garonne	North of Toulouse, France	2020-02-09	VV	80m - 200m	1109×1704
6 - Redon	Oust Vilaine	Redon, Brittany France	2018-07-04	VH	15m - 60m 40m - 160m	618×773
7 - Régina	Arataï Approuague	Régina, French Guiana, France	2017-10-11	VH	25m - 100m 100m - 150m	553×1216

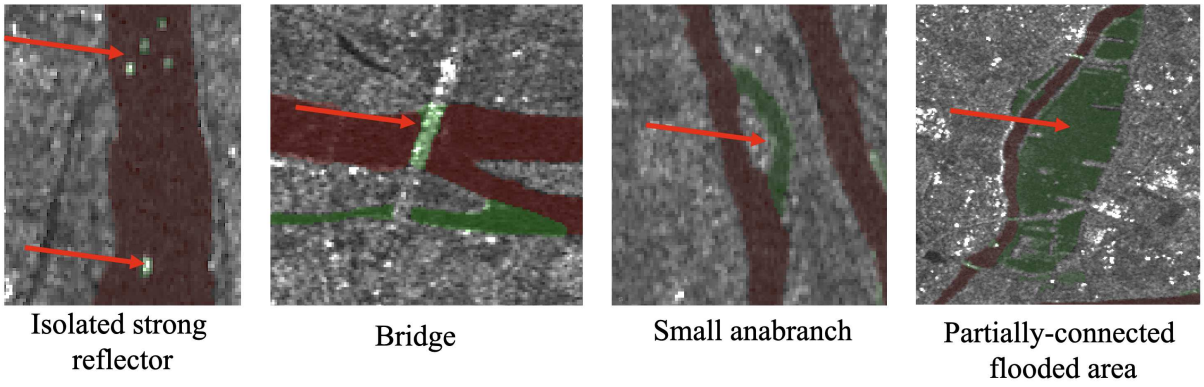


Figure 5.22 – Illustration of situations in which an uncertain classification (green) is used for the ground truth. Red corresponds to a Water classification, and the uncolored image corresponds to a Land classification in the ground truth.

All Sentinel-1 image extracts and associated ground truth are made available in the same repository as our published code.

SWOT Concerning SWOT images, as the SWOT satellite has not yet been launched, all test images have been simulated with the Jet Propulsion Laboratory (JPL) HR science simulator (JPL D-79123, 2014). These images are associated with the water mask that has been used for the simulation as ground truth.

We used three simulated images for our experiments. All images have been simulated considering pessimistic assumptions about the performances of the sensor (worst-case scenario). The

first image has been simulated from Lidar and high-resolution landcover data on the Saline River, Lincoln County, Kansas, USA, and presented in the previous part. This image has been simulated with the so-called dark water phenomenon. Dark water is water with a very low contrast compared to land and is caused by very low water surface roughness at low wind speed. This dark water phenomenon, and numerous bright land structures, make river detection especially difficult on this image. The two other images have been simulated using Lidar data on the Rhône delta, France. Unlike the Saline River image, these two images have been simulated without dark water: the contrast between water and land is more homogeneous. Image 9 corresponds to the downstream Petit Rhône river, whereas image 10 corresponds to the upstream Petit Rhone river and two small channels: Canal Bas-Rhône Languedoc and Canal du Rhône à Sète.

The SWOT images are summarized in Table 5.2. The river widths are here given in pixels and not in meters as the pixel ground range spacing in SWOT depends on the position in the swath.

Table 5.2 – Simulated SWOT images used for our experiments. All images are simulated in worst-case scenario on the sensor performance.

Image # - name	River name	Location	Simulated dark water	River width (pixels)	Size (pixels)
8 - Saline	Saline	Lincoln County, Kansas, USA	Yes	2-5	301×351
9 - Petit Rhône downstream	Petit Rhône	Camargue France	No	3-14	700×800
10 - Petit Rhône Upstream and channels	Petit Rhône Canal Bas-Rhône Languedoc Canal du Rhône à Sète	Camargue France	No	2-8	800×730

5.3.2 Metrics

In order to quantitatively assess the performance of the water detection compared to our ground truth, we use the same six metrics as Lobry et al. (Lobry et al., 2019). These metrics are based on the number of pixels considered as true positives (TP) for adequately classified water, true negatives (TN) for adequately classified land, false negatives (FN) for water classified as land and false positives (FP) for land classified as water.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5.24}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{5.25}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5.26}$$

$$\text{F-score} = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5.27}$$

$$\text{ER} = \frac{\text{FP} + \text{FN}}{\text{TP} + \text{FN}} \tag{5.28}$$

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FN})(\text{FP} + \text{TN})(\text{TP} + \text{FP})(\text{TN} + \text{FN})}} \tag{5.29}$$

The recall is the proportion of actual water pixels that are classified as water. The FPR is the proportion of land pixels that are classified as water. The precision is the proportion of actual water among all the pixels classified as water. The F-score is the harmonic mean of precision and recall and will be our main metrics. ER is the ratio between the number of incorrectly classified pixels and the number of actual water pixels. This metric is similar to the metric of the SWOT mission science requirements (JPL-D-61923-rev-B, 2018), but computed in radar geometry instead of ground geometry. The Matthews correlation coefficient (MCC) (Matthews, 1975) is another metric that takes into account the over-representation of land in the context of river detection.

5.3.3 Implementation and parameters

For each image, we extract the rivers using our method by choosing a very limited number of prior centerline nodes, in order to highlight the robustness of the proposed approach. For single rivers (except for image 8, used as an example in the previous part), we use only two nodes: one for each endpoint. When two rivers are joining in a confluence, we locate one node on the confluence and one node at each endpoint of the two upstream rivers and of the downstream river. In the case of an anabranch (e.g. in Angers image), a node is added in the anabranch in order to prevent its centerline from going through the main stream. The nodes that have been used are displayed on the images.

Table 5.3 – Parameters used for the experiments

	Line detection			Centerline detection	River segmentation			
	Patch size	Scale Range	Angular step	Lin. detector power	Regularization		Flux	
	N	$[S_{min}, S_{max}]$	θ_{step}	N_{pow}	β	λ	σ_L	η
SWOT	9	[1,3]	3°	70	15	0.2	3	6
S1 GRD	9	[1,4]	3°	10	15	0.2	3	6

We used the parameters presented in Table 5.3. These parameters have been chosen empirically by testing multiple values on the SWOT simulated image Saline. We manually increased the maximum scale S_{max} of the detection of the linear structures from 3 to 4 to account for the wider range of river width in our use of Sentinel-1 images and decreased the N_{pow} parameter from 70 to 10 in order to be more robust to dark roads. For both kinds of images, we used L=4 for our experiments.

The results could have been improved by fitting the parameters to the type of image (SWOT, Sentinel-1 VV, Sentinel-1 VH) or even to the environment (urban area, rain-forest, desert...), but

our main goal for these experiments was to show satisfactory performances without fine-tuning of the parameters.

Concerning the optimization of the code we use, we improved the computation of the linear features detection, which is by far the slowest step, by using the fast computational approach proposed in section 5.2.2. Moreover, the convolutions are processed in the Fourier domain and the FFT of the image is computed only once for all the orientations.

5.3.4 Results

Table 5.4 gives the metrics for each image in our dataset. The metrics are computed only for river detection.

Five images are presented in detail below, with their associated detection maps: image 1 (Des Moines) is representative of the results obtained with our method for typical Sentinel-1 images in urban areas, image 2 (Sunar) to present an example where the centerline detection is not successful, and image 9 (Petit Rhône Downstream) as an example for SWOT images. For image 1 (Des Moines) and 2 (Sunar), the centerline detected with a major (more than 1km) discrepancy in the position of the centerline is presented (**Figure** 5.24 and 5.27). All ten images of our dataset and the corresponding segmentation results are presented in the supplementary materials associated with our article (Gasnier et al., 2021b) and for Sentinel-1 images, the results can be reproduced using the published code.

Example 1 Image 1 (Des Moines), displayed in **Figure** 5.23, shows that our method leads to correct detection of the whole river, despite using only two nodes as prior information, and although the river is meandering. The centerline (b) has been correctly classified with the proposed approach based on the response of the linear feature detector. The segmentation of the river from the centerline using our conditional random field approach also gives good results in this example. The river contour is relatively well respected. It can be noted that, despite a reflectivity similar to the reflectivity of the river, the lake (which is not connected to the river) and two large roads (**Figure** 5.23 (a)) are not misclassified as rivers. Thanks to the use of prior information, our approach avoids two typical pitfalls of river detection on SAR images that are lakes close to rivers and highways.

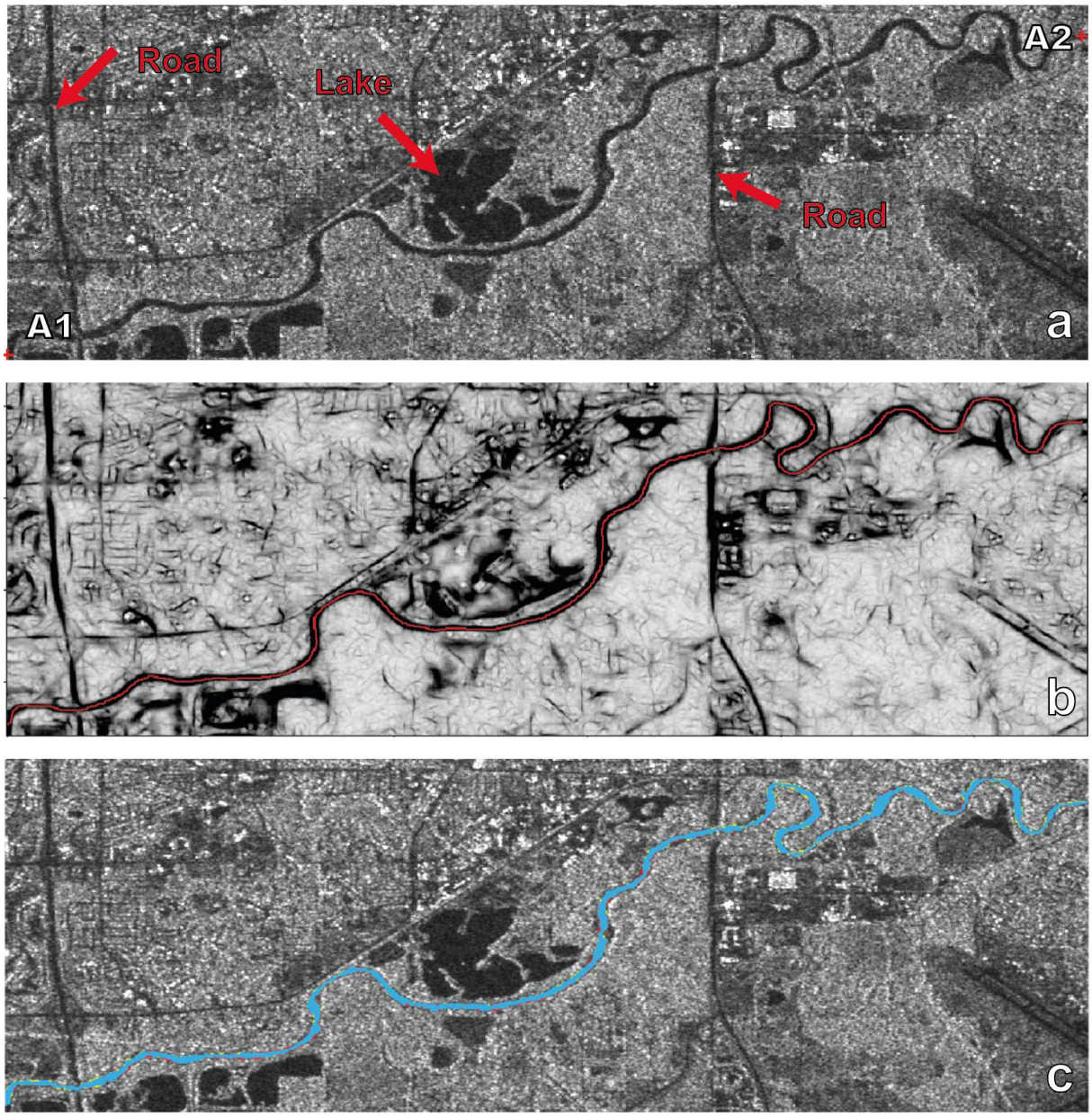


Figure 5.23 – Image 1 (Des Moines): (a) SAR image with annotations, (b) centerline (in red) on the linear features detector and (c) final segmentation. A1 and A2 mark the two nodes used as prior information. The color map of the line detector has been inverted and the centerline has been widened for better visualization. In (c) the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image pixels.

Example 2 Image 2 (Sunar) presented in **Figure** 5.25 illustrates a possible issue with the proposed approach when using insufficient exogeneous information about the location of the river. If a dark linear structure in a river meander in a Sentinel-1 image creates a shorter path between two *a priori* nodes of the centerline and if the actual river is not identifiable, the detected centerline will be incorrect. This leads to false positives on the dark linear structure and false negatives in the part of the river that has been bypassed, such as in **Figure** 5.26. The resulting

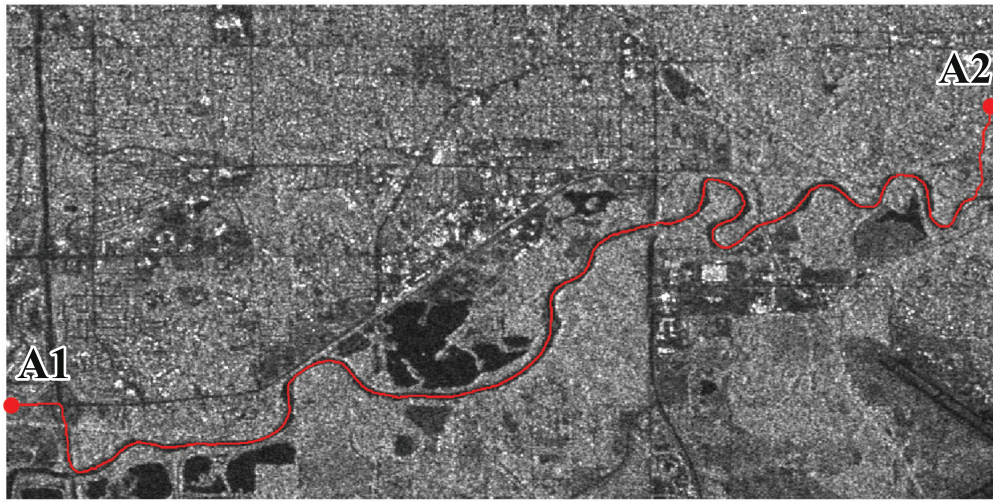


Figure 5.24 – Image 1 (Des Moines): Result with the *a priori* nodes shifted vertically of more than 1 km. The centerline is still correctly detected and the water detection is correct, except for the area between the *a priori* nodes and the river.

classification is erroneous for this part of the river. However, this does not significantly affect the classification of the remaining part of the river, as the estimation of the water reflectivity R_1 is robust enough.

A possible improvement would be to use more centerline nodes as exogenous information and to use a post-processing step to flag as uncertain the river parts where the reflectivity is too high (possibly sand, mud, or flooded vegetation) and remove them if appropriate.

Example 3 Image 9 (Petit Rhone downstream), presented in **Figure** 5.28, illustrates the behavior of the proposed method applied to simulated SWOT HR images. In this example, the river centerline has been correctly detected and the river segmentation is relatively accurate except for some false positives caused by speckle noise, and a very small connected channel that has not been detected. In comparison with the baseline method (Lobry et al., 2019) that only detects a small part of the narrow river, the proposed approach shows an improved detection. Because our approach does not detect other water surfaces, but only rivers that would have been missed by the generic method, both approaches are complementary.

It can be noted that for SWOT images, the bright area corresponding to the river response might be slightly larger than the river itself in the azimuth direction because water is moving and does not necessarily remain coherent during the entire SAR integration time. This issue could be addressed by a morphological post-processing in order to erase such false positive pixels and thereby improve the precision.

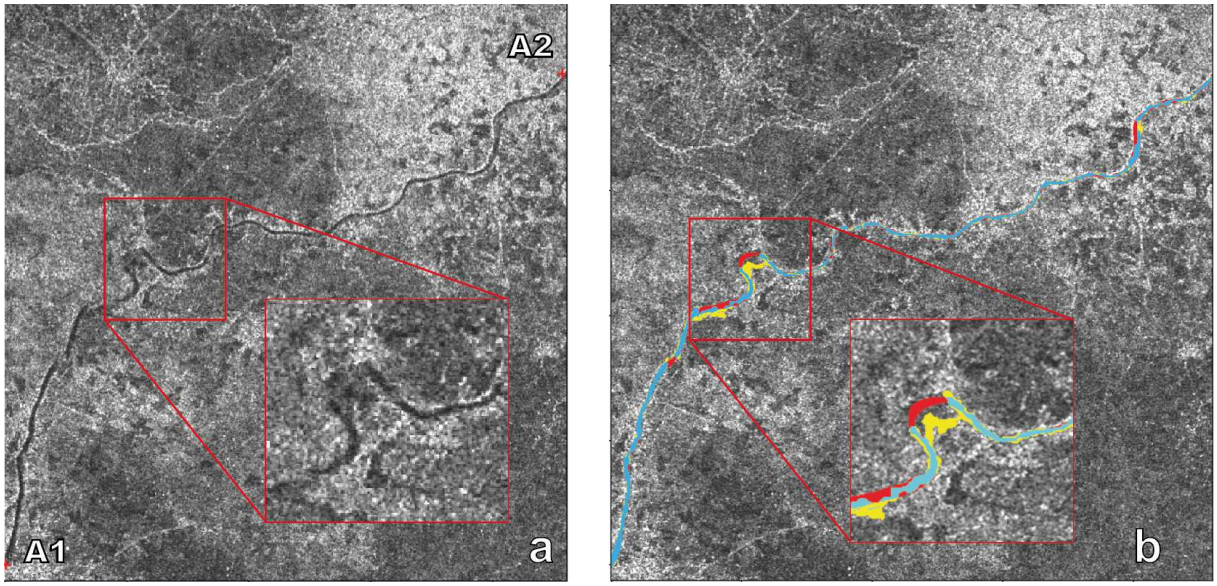


Figure 5.25 – Image 2 (Sunar): (a) SAR image with annotations and (b) final segmentation. A1 and A2 mark the two nodes used as prior information. The close-up squared in red in both images show a meander in which the segmentation is unsuccessful as the centerline bypasses the meander. A1 and A2 mark the two nodes used as prior information. In (b) the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image pixels.

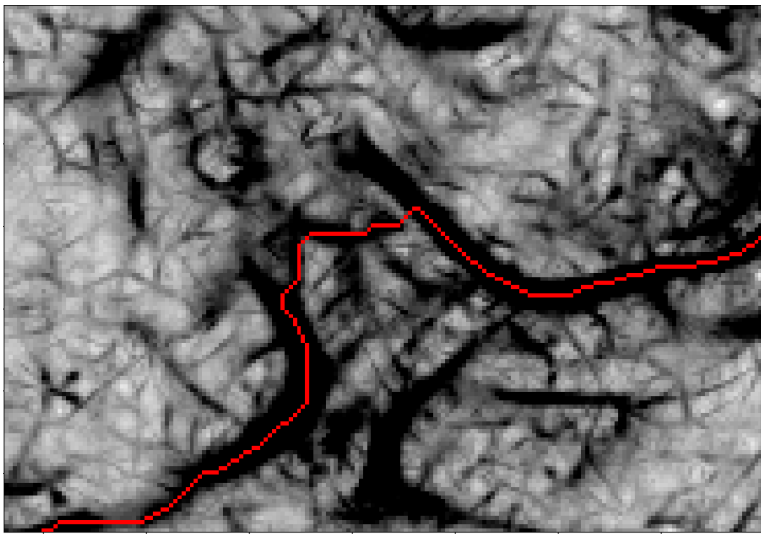


Figure 5.26 – Image 2 (Sunar): Zoom on the red square area in **Figure** 5.25, with the response of the linear structure detector (inverted grayscale) and the detected centerline (red, widened for better visualization).

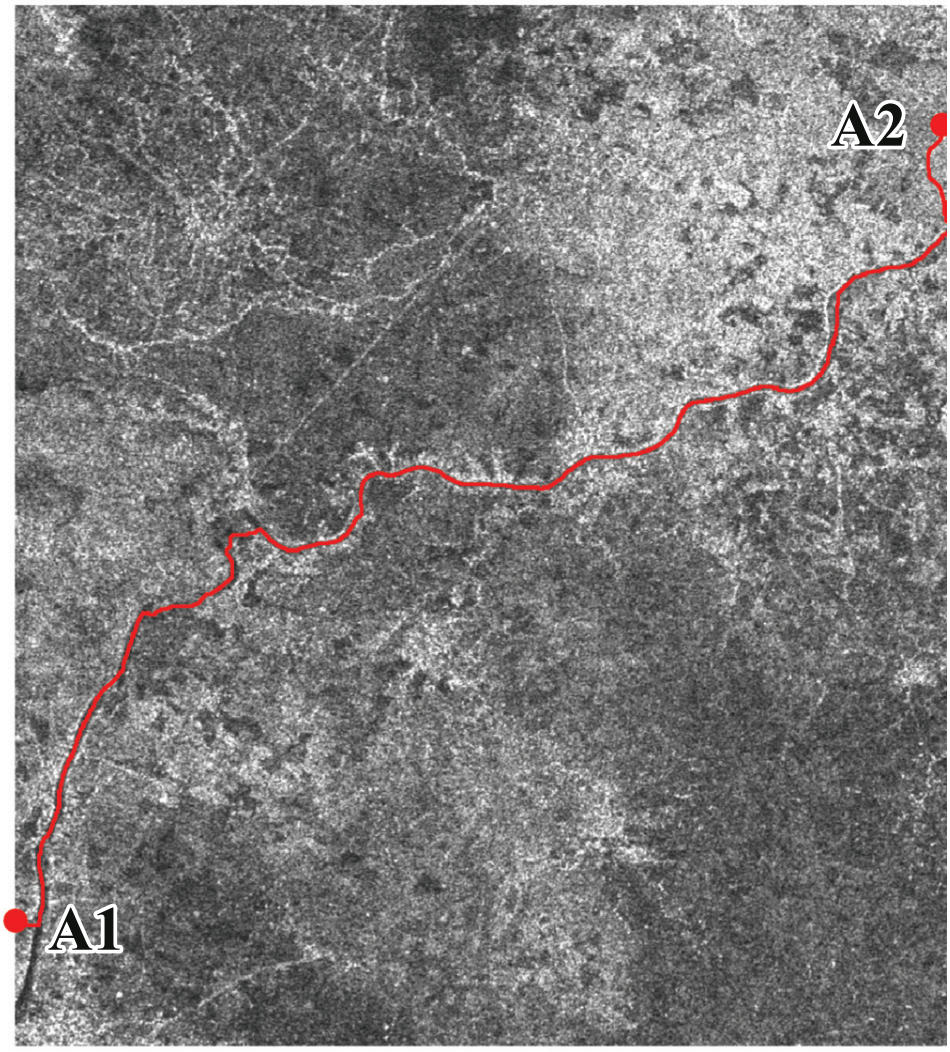


Figure 5.27 – Image 2 (Sunar): Result with the *a priori* nodes shifted vertically of more than 1 km. The centerline is still correctly detected and the water detection is correct, except for the area between the *a priori* nodes and the river.

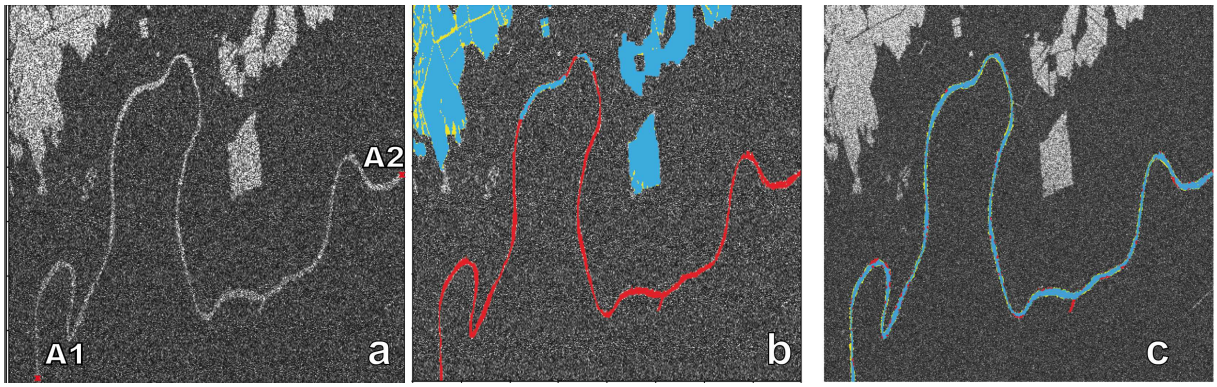


Figure 5.28 – Image 9 (Petit Rhône downstream): (a) SAR image with *a priori* nodes, (b) segmentation with the baseline MRF method, and (c) proposed method segmentation. A1 and A2 mark the two nodes used as prior information. In (b) and (c) the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image pixels.

Table 5.4 – Summary of the metrics for each result

Number	Name (sensor)	Method	Pr (%)	Rec (%)	FPR (%)	**F-Score (%)**	ER (%)	MCC (%)	Exec time (s)
1	Des Moines (S1)	Proposed	92.44	93.35	0.13	**92.89**	14.29	92.78	57.73s
2	Sunar (S1)	Proposed	82.36	81.71	0.15	**82.03**	35.79	81.88	89.32s
3	Gaoual (S1)	Proposed	92.51	89.09	0.12	**90.77**	18.12	90.64	212.96s
4	Angers (S1)	Proposed	98.90	94.04	0.05	**96.40**	7.01	96.28	160.96s
5	Garonne (S1)	Proposed	97.60	82.44	0.02	**89.38**	19.59	89.60	166.01s
6	Redon (S1)	Proposed	90.28	92.70	0.15	**91.47**	17.28	91.35	47.48s
7	Régina (S1)	Proposed	89.33	82.95	0.18	**86.02**	26.96	85.83	62.86s
8	Saline (SWOT)	Proposed	63.24	94.45	1.02	**75.76**	60.45	76.81	10.45s
		Baseline	5.30	87.58	33.87	**10.00**	1576.64	65.92	/
9	Petit Rhône downstream (SWOT)	Proposed	80.71	89.46	0.57	**84.86**	31.92	84.56	47.23s
		Baseline	91.00	9.80	0.03	**17.69**	91.17	9.66	/
10	Petit Rhône upstream and channels (SWOT)	Proposed	73.07	87.45	0.58	**79.62**	44.78	79.55	57.17s
		Baseline	87.32	8.89	0.02	**16.14**	92.40	8.80	/

5.4 Conclusion

In this chapter, a novel river extraction method that corresponds to the contribution 1 presented in section 1.2 is proposed and evaluated. The originality of our approach is that it uses an exogenous river database in order to guide river detection. The proposed technique consists of three phases: first, computing the response of a linear feature detector (corresponding to contribution 1-A), then detecting the centerline using the response and the prior river nodes, and finally segmenting the river around the previously detected centerline using a CRF approach corresponding to contribution 1-B. Experiments performed on both Sentinel-1 and simulated SWOT HR images have shown that our method performs well including in low contrast situations and for very narrow rivers of only a few pixels.

The proposed method has been developed in the context of the SWOT mission to process SWOT HR images that are single-polarization and cannot easily be combined with images from other sensors. This leads us to design a resilient method for river segmentation in such images.

The direct application of the proposed framework has obvious potential for monitoring rivers included in the GRWL database, but it may also be adapted to the detection of rivers unknown to the database. For example, if other hydrological information or a digital elevation model (DEM) indicates that a small tributary is missing from the database, our approach can help to retrieve it by using two inputs: one node in the main river and one node placed further up in the expected tributary. For example, the river databases derived from SRTM such as HydroRIVERS (see section 3.3.2) are much more comprehensive than GRWL but lack its accuracy and its centerline information. These databases could be used as seeds with our guided detection approach to supplement SWORD database.

Other interesting research directions concern the adaptation of the proposed approach to other applications than river monitoring, for example, road extraction in SAR images or coherence images.

Chapter 6

Adaptation of the GrabCut method to SAR images: lake detection from *a priori* polygon

6.1 Single-date GrabCut method for lake detection from *a priori* polygon

6.1.1 Introduction

While the main challenge for water detection in SWOT SAR images with low contrast concerns small rivers, the detection of small lakes can be an issue as well, especially with dark water, and layover effects. Moreover, because the local water reflectivity is not precisely known, especially in SWOT images, assumptions on water and land reflectivities have to be handled with care. Similar issues exist for Sentinel-1 images, in which the detection of small water bodies can also be difficult, with a risk of false detection in case of other dark structures (fields, tarmacs,...). For these reason, we propose a guided detection method that can use the maximum extent polygon from SWOT's *a priori* lake database (see section 3.3.3), similar to the way its river detection counterpart (chapter 5) is guided by the SWORD river database.

To account for possible discrepancies in the position and shape between the projected database polygon and the actual lake in the image, morphological dilations of the database polygon can be used. This ensures that the *a priori* polygon contains the whole lake.

This exogenous information in the form of a bounding shape (not necessarily minimal) that contains the object is very similar to the bounding rectangle used as user input in (Rother et al., 2004)'s GrabCut for interactive segmentation.

This chapter presents an adaptation of this GrabCut method to water detection in SAR images. First, the original GrabCut method is presented in section 6.1.2. Then, the method we propose is presented in 6.1.3 and some results are shown in 6.1.4. An adaptation of this method to the multitemporal segmentation of SAR image time series is proposed in section 6.3.

Both methods rely on prior knowledge about the lake in the form of a bounding shape, but do not require any *a priori* information on the modeling of the water and land classes. On the contrary, they enable the estimation of the distributions of these classes as mixture models.

In section 6.4, we present a segmentation approach that uses these mixture models to detect water in a time series composed of both SAR and optical images.

6.1.2 Original GrabCut method

The segmentation in the original Grab Cut method (Rother et al., 2004) tries to separate a color (RGB) image into two classes: an object class (foreground) $\mathcal{C}_F$ and the background class $\mathcal{C}_B$. Each of these classes $\mathcal{C}$ is modeled as a Gaussian Mixture Model (GMM). Each of these GMMs consists in a set of $n_{\mathcal{C}}$ ($n_{\mathcal{C}} = 5$ in the article) sub-classes $\mathcal{K}$. Each of these sub-classes is associated with a three-dimensional Gaussian model, with a given mean value and covariance matrix, to model color distributions.
First, the classes are initialized using the user-defined bounding box for the object. Pixels inside the box are assigned to $\mathcal{C}_F$ and pixels outside the box are assigned to $\mathcal{C}_B$. Then, a clustering step splits each class into $n_{\mathcal{C}}$ sub-classes. While the clustering method used by (Rother et al., 2004) is not disclosed, (Talbot et al., 2004) propose to use (Orchard and Bouman, 1991)'s binary tree quantization algorithm when dealing with natural RGB images as it is done in (Ruzon and Tomasi, 2000). A simpler approach, used by OpenCV's GrabCut implementation[1], is to use the k-means algorithm for this clustering step.
After this initial stage ($n = 0$), the initial parameters of the two GMMs are known.

Then, three steps are repeated until convergence using the current foreground/background segmentation:

1. Each foreground pixel is assigned to the most likely subclass of the foreground class $\mathcal{K} \in \mathcal{C}_F$. The same is done for the background class.

2. New subclass parameters are estimated from the pixels that belong to them.

3. A graph is built to minimize an energy that combines a data term and a regularization term. The regularization term is based on a CRF model and the data term derives from combinations of the likelihoods for each subclass. These likelihoods rely on the previously estimated subclass parameters. The graph is segmented using the (Boykov and Kolmogorov, 2004) algorithm and the resulting segmentation is used to assign each pixel to the $\mathcal{C}_F$ or $\mathcal{C}_B$ class.

Note that this method could be considered as a kind of "hard labels" Expectation-Minimization (EM) (Dempster et al., 1977; Xu and Jordan, 1996) method. Indeed, as opposed to the original "soft labels" EM clustering approaches, its expectation stage does not attribute for each pixel its probability of belonging to any given class. Instead, the GrabCut assigns them to the class that maximizes this probability. However, the data term for the segmentation step relies on a linear combination of the likelihood of each class, which corresponds to the likelihood in a "soft label" EM.

In (Rother et al., 2004) , this segmentation is followed by a border mating approach that will not be described here.

[1] https://github.com/opencv/opencv/blob/master/modules/imgproc/src/grabcut.cpp

6.1.3 Proposed iterative approach

We propose an adaptation of the method described in the previous paragraph for water detection in SAR images.

6.1.3.1 Overview

Our method is run on log-transformed intensity images and relies on an assumption of fully-developed speckle for the modelization of the sub-classes.

As for the original GrabCut method, the goal is to divide the image into two classes. For our problem, these two classes are defined as follows:

- A water class $\mathcal{C}_W$, which corresponds to water pixels ($\{k, \boldsymbol{\ell}(k) = 1\}$).
- A land class $\mathcal{C}_L$, which corresponds to land (non-water) pixels ($\{k, \boldsymbol{\ell}(k) = 0\}$).

The classes $\mathcal{C}_W$ and $\mathcal{C}_L$ are modeled with Fisher-Tippett Mixture Models (FTMM) and contain respectively $n_{\mathcal{C}_W}$ and $n_{\mathcal{C}_L}$ sub-classes. As land surfaces are more diverse than water surfaces, we can choose $n_{\mathcal{C}_W} < n_{\mathcal{C}_L}$. However, keeping $n_{\mathcal{C}_W} > 1$ can be useful as water surfaces can still be inhomogeneous, for example because of Bragg scattering (see section 2.1.3.2).

A vector $\boldsymbol{V}_{n_{it}}$ is introduced to assign a sub-class $\boldsymbol{V}_{n_{it}}(k) = \mathcal{K}$ to each pixel k at iteration n_{it}.

Each subclass $\mathcal{K}$ is defined by two parameters that are recomputed at each iteration n_{it}:

- A mean value $\boldsymbol{\mu}(\mathcal{K}, n_{it})$, which is the arithmetic mean of the log-reflectivities of the pixels in the subclass.
- A weight $\boldsymbol{\pi}(\mathcal{K}, n_{it})$ which is the proportion of pixels in $\mathcal{C}$ that belongs to $\mathcal{K}$.

$\boldsymbol{\mu}(, n_{it})$ and $\boldsymbol{\pi}(, n_{it})$ are vectors of length $n_{\mathcal{C}_W} + n_{\mathcal{C}_L}$. At each iteration, the parameters are computed as follows:

$$\boldsymbol{\mu}(\mathcal{K}, n_{it}) = \frac{1}{n_{\mathcal{K}}} \sum_{\{k, \boldsymbol{V}_{n_{it}}(k) = \mathcal{K}\}} \boldsymbol{y}(k) \tag{6.1}$$

$$\boldsymbol{\pi}(\mathcal{K}, n_{it}) = \frac{n_{\mathcal{K}}}{n_{\mathcal{C}}} \tag{6.2}$$

where $n_{\mathcal{K}}$ is the number of pixels in the sub-class $\mathcal{K}$ and $n_{\mathcal{C}}$ is the number of pixels in the class $\mathcal{C}$, to which the sub-class $\mathcal{K}$ belongs.

As opposed to the original GrabCut approach, the variance of the subclass is not considered as a class parameter. Indeed, with our fully developed speckle model with an assumption of homogeneous reflectivity inside of a sub-class, the distribution of log-intensities only depends on the reflectivity and on the equivalent number of looks (ENL) L, which is a characteristic of the image and thus the same for all classes.

Likelihood of one sub-class The distribution of the log-intensities $\boldsymbol{y}(k)$ assuming a homogeneous reflectivity $R_{\mathcal{K}}$ for all the pixels $k/\boldsymbol{V}_{n_{it}}(k) = \mathcal{K}$ of the subclass $\mathcal{K}$ is given in section 2.1.3.1. With L the ENL of the image and $x_{\mathcal{K}} = \log(R_{\mathcal{K}})$, the log-reflectivity $x_{\mathcal{K}}$ is related to the arithmetic mean of the log-intensities $\boldsymbol{\mu}(\mathcal{K}, n_{it})$ as follows:

$$\boldsymbol{\mu}(\mathcal{K}, n_{it}) = x_{\mathcal{K}} - \log(L) + \psi(L) \tag{6.3}$$

Hence the likelihood becomes:

$$\mathcal{L}(\boldsymbol{y}(k)|x_{\mathcal{K}}) = \frac{L^L}{\Gamma(L)} e^{L(\boldsymbol{y}(k)-x_{\mathcal{K}})} \exp(-Le^{\boldsymbol{y}(k)-x_{\mathcal{K}}}) \tag{6.4}$$

where $x_{\mathcal{K}} = \boldsymbol{\mu}(\mathcal{K}, n_{it}) + \log(L) - \psi(L)$. Note that $\mathcal{L}$ is a likelihood and not the neg-log-likelihood $\mathscr{L} = -\log(\mathcal{L})$.

Summary of the algorithm The method consists in one initialization step (see section 6.1.3.2) and three subsequent steps that are repeated for a given number of iterations $n_{\max}$:

1. A pixel assignment step (section 6.1.3.3)
2. A parameter learning step (section 6.1.3.4)
3. A segmentation step (section 6.1.3.5)

An overview of the algorithm and of the variable that are used and modified at each step is provided by **Figure** 6.1.

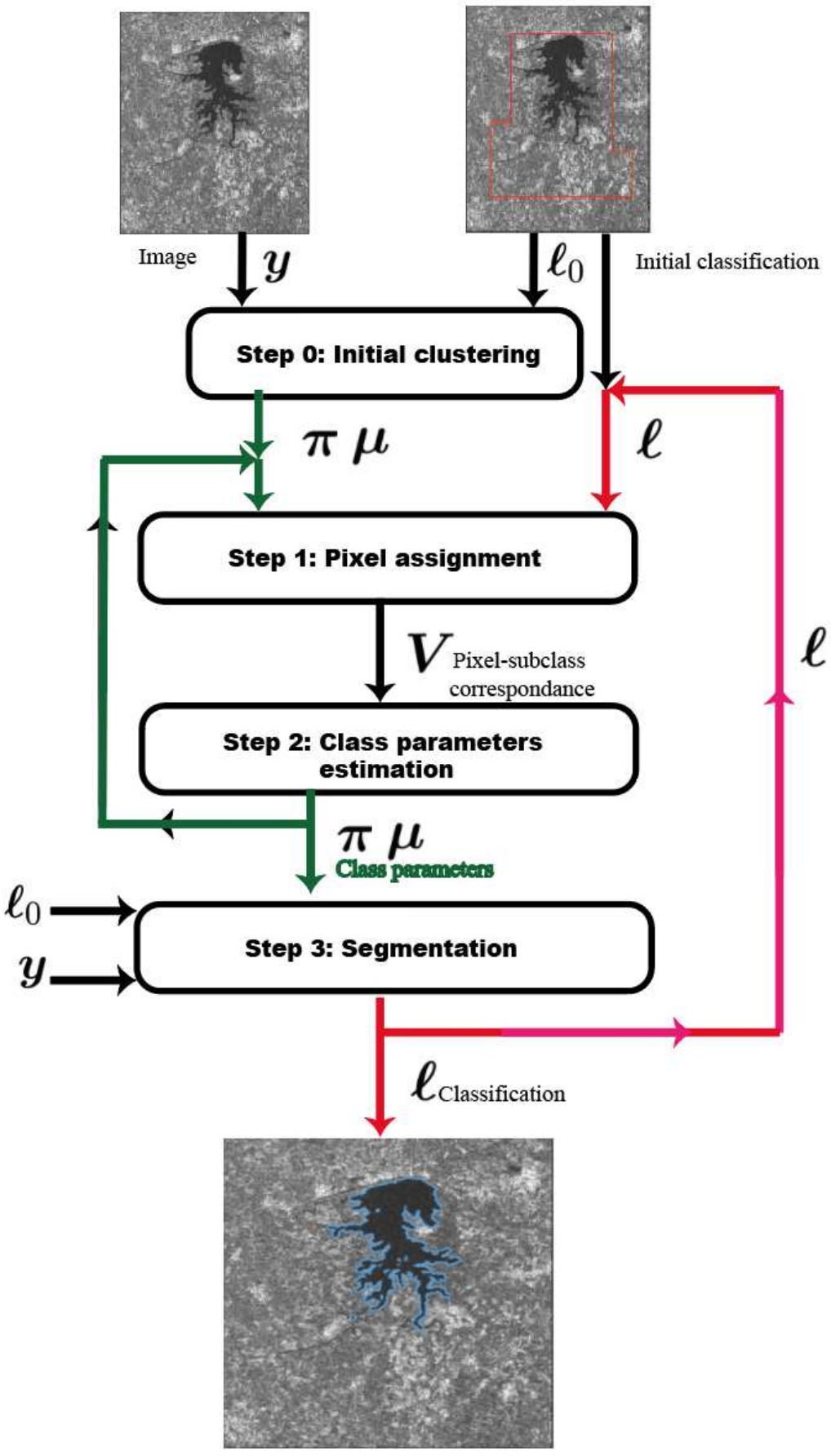


Figure 6.1 – Overview of the proposed algorithm.

.

6.1.3.2 Initialization step

Initial rough segmentation The initialization step starts with the initial rough segmentation given by the *a priori* polygon, such as in **Figure** 6.2. This polygon could be from a morphological dilation of SWOT prior lake database polygon when available. Pixels inside the bounding polygon are given the label $\boldsymbol{\ell}_0 = 1$, while pixels outside the polygon are given the label $\boldsymbol{\ell}_0 = 0$.

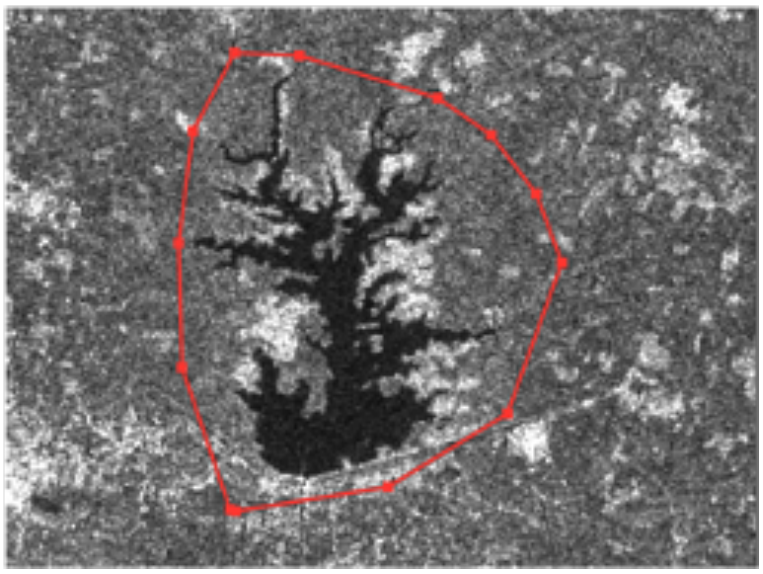


Figure 6.2 – Example of bounding polygon around the Sajnam reservoir (India). In this example, the polygon has been drawn by hand.

This rough segmentation defines two initial sets of pixels: $\{k, \boldsymbol{\ell}_0(k) = 1\}$ and $\{k, \boldsymbol{\ell}_0(k) = 0\}$.

Determining initial sub-classes To be able to perform the next step, the initial class parameters $\boldsymbol{\mu}(\mathcal{K}, 0)$, and $\boldsymbol{\pi}(\mathcal{K}, 0)$ have to be determined. To that end, the most straightforward approach is to use a clustering technique and then compute the parameters for each cluster. In this way, $\{k, \boldsymbol{\ell}_0(k) = 1\}$ and $\{k, \boldsymbol{\ell}_0(k) = 0\}$ have to be clustered into the $n_{\mathcal{C}_W}$ and $n_{\mathcal{C}_L}$ sub-classes. As the distributions of our log-transformed reflectivities are approximately Gaussian (see section 2.1.3.1), a k-means (MacQueen, 1967) clustering algorithm can be appropriate and is used in our implementation. An example of a clustering result on a log-transformed SAR image is presented in **Figure** 6.3. If we were dealing directly with intensity images, their skewed distribution would require the use of clustering algorithms that are specific to gamma-distributed variables (Almhana et al., 2006).

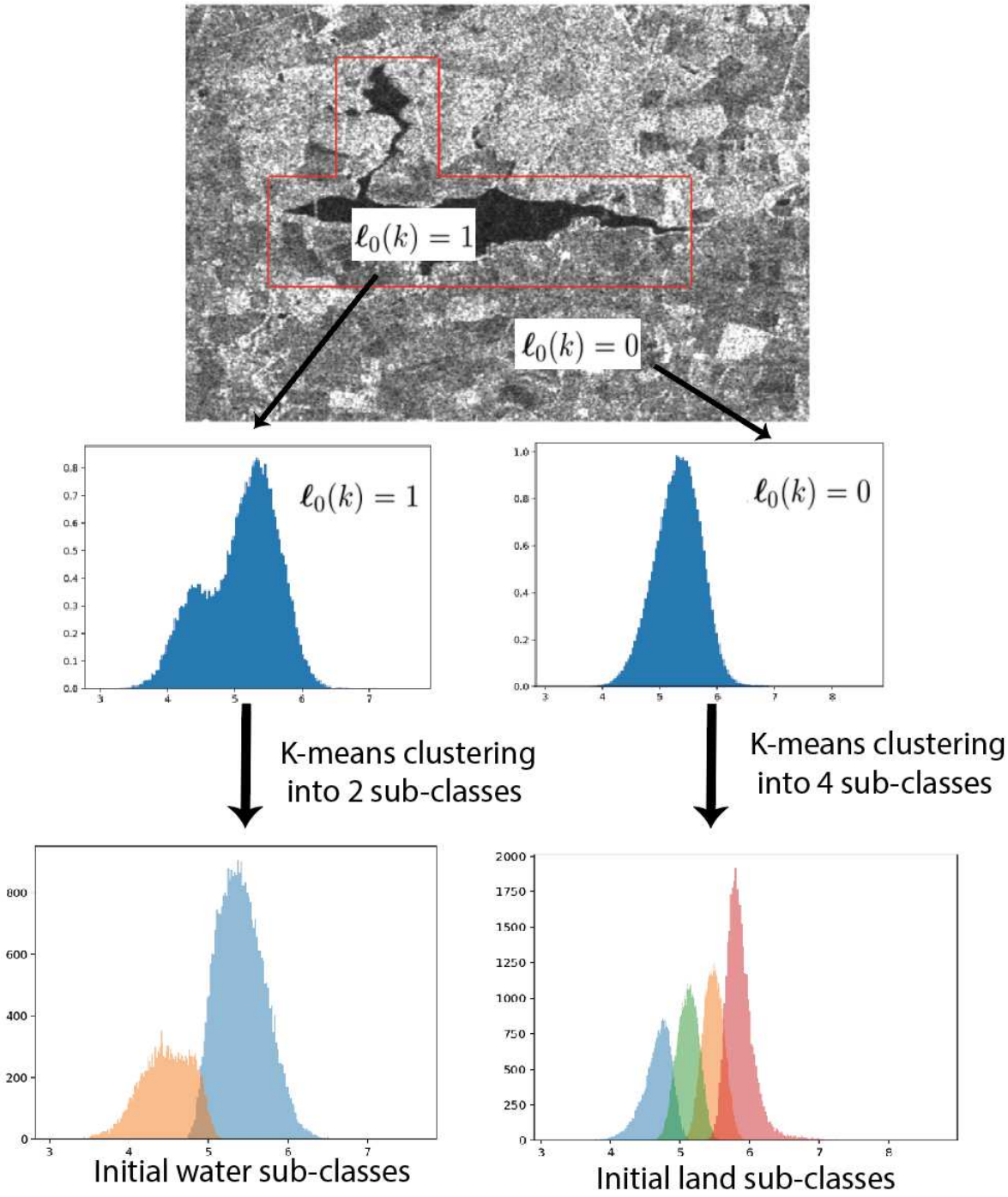


Figure 6.3 – Distribution of the initial water and land pixels and resulting sub-classes after k-means clustering on the log-transformed intensities, with 2 sub-classes for water and 4 sub-classes for land.

.

With more class parameters, another option would be to adapt to log-transformed intensities a Generalized Gamma Mixtures Models (GΓMM), such as the method proposed by (Li et al., 2016), but such approaches would make the model more complex and the improvement that can be expected is limited as the modelization of the empirical distributions is already acceptable.

At the end of this step, the initial parameters for all sub-classes $\boldsymbol{\mu}(\mathcal{K}, 0)$, and $\boldsymbol{\pi}(\mathcal{K}, 0)$ have

been determined.

6.1.3.3 Pixel assignment step

The second step consists in assigning each pixel to the sub-class that maximizes its likelihood among all the sub-classes of the class he belongs to. This means that each pixel k, with $\boldsymbol{\ell}(k) = 0$ (resp. $\boldsymbol{\ell}(k) = 1$) is assigned to the subclass $\mathcal{K} \in \mathcal{C}_k$, where $\mathcal{C}_k = \mathcal{C}_L$ (resp. $\mathcal{C}_k = \mathcal{C}_W$) that maximizes $\mathcal{L}(\boldsymbol{y}(k)|\boldsymbol{V}_{n_{it}}(k) = \mathcal{K})$. The expression of $\mathcal{L}$ is given by equation 6.4 in section 6.1.3.1.

$$\boldsymbol{V}_{n_{it}}(k) = \arg\max_{\mathcal{K}\in\mathcal{C}_k} \mathcal{L}(\boldsymbol{y}(k)|\boldsymbol{x}_{\mathcal{K}}) \tag{6.5}$$

At the end of this step, all pixels k have been assigned to the sub-class they belong to, given their log-reflectivity $\boldsymbol{y}(k)$ and their label $\boldsymbol{\ell}(k)$ (this information is stored in the vector $\boldsymbol{V}_{n_{it}}$).

Note that after the first iteration, a particular sub-class may receive no pixel. Think of a situation in which a "bright water" sub-class (for a conventional SAR image) of the water class contains very bright pixels at the initial iteration, hence a very high mean value $\boldsymbol{\mu}$. These pixels did correspond to the land surrounding the lake that was located inside the bounding polygon. If the subsequent segmentation of the lake is correct, these land pixels will be classified to the land category. Then, the remaining water pixels will be much darker and none of them are likely to be assigned to the "bright water" subclass. This can result in an actual number of subclasses below its initial value.

6.1.3.4 Parameter learning step

The next step consists in updating the parameters of each sub-class. Indeed, except for the first iteration, the sub-classes have changed since the last computation of the parameters.

This is done by simply applying the formulas given in section 6.1.3.1:

$$\boldsymbol{\mu}(\mathcal{K}, n_{it}) = \frac{1}{n_{\mathcal{K}}} \sum_{k/\boldsymbol{V}_{n_{it}}(k)=\mathcal{K}} \boldsymbol{y}(k) \tag{6.6}$$

$$\boldsymbol{\pi}(\mathcal{K}, n_{it}) = \frac{n_{\mathcal{K}}}{n_{\mathcal{C}}} \tag{6.7}$$

where $n_{\mathcal{K}}$ is the number of pixels in the class $\mathcal{K}$.

At the end of this step, the parameters $\boldsymbol{\mu}(\mathcal{K}, n_{it})$ and $\boldsymbol{\pi}(\mathcal{K}, n_{it})$ of the sub-classes are updated.

6.1.3.5 Segmentation step

The goal of the segmentation step is to find a segmentation of the image between the two classes. Each class is modeled by a mixture of Fisher-Tippett distributions defined by the parameters $\boldsymbol{\mu}(\mathcal{K}, n_{it})$ and $\boldsymbol{\pi}(\mathcal{K}, n_{it})$ of its sub-classes.

The method we propose for the segmentation step is similar to the approach used by the original Grab Cut method but with some adaptations, as the characteristics of SAR images are different from RGB natural images. The segmentation is obtained by minimizing a global energy $\mathscr{E}$ that depends on the classification $\boldsymbol{\ell}$ ($\boldsymbol{\ell} = 1$ for water and $\boldsymbol{\ell} = 0$ for land)

$\mathscr{E}$ is the sum of one data term, a CRF-based regularization term, a flux term and a term preventing water detection outside the *a priori* polygon:

$$\mathscr{E}(\boldsymbol{\ell}, \boldsymbol{y}) = U^I_{\text{data}}(\boldsymbol{\ell}, \boldsymbol{y}) + U_{\text{reg}}(\boldsymbol{\ell}, \boldsymbol{y}) + U_{\text{flux}}(\boldsymbol{\ell}, \boldsymbol{y}) + U_{\text{P}}(\boldsymbol{\ell}, \boldsymbol{\ell}_0) \tag{6.8}$$

The data term U^I_{data} ensures fidelity to the log-intensity image $\boldsymbol{y} = \log(\boldsymbol{I})$. The regularization term U_{reg} is derived from a CRF model, using a gradient computed with the ROEWA method (Fjortoft et al., 1998; Dellinger et al., 2015) on the intensities $\boldsymbol{I}$. This regularization term is different from its counterpart of chapter 5. The term U_{flux} favors a high outward flux of the gradient through the water boundary. An U_{P} term prevents the classification of pixels outside the *a priori* polygon as water and is comparable to the U^C_{data} term in chapter 5.

Data term The likelihood of the value $\boldsymbol{y}(k)$ given a sub-class $\mathcal{K}$ is given by equation 6.4 in section 6.1.3.1:

$$\mathcal{L}(\boldsymbol{y}(k)|\boldsymbol{x}_{\mathcal{K}}) = \frac{L^L}{\Gamma(L)} e^{L(\boldsymbol{y}(k) - x_{\mathcal{K}})} \exp(-Le^{\boldsymbol{y}(k) - x_{\mathcal{K}}}) \tag{6.9}$$

where $x_{\mathcal{K}} = \boldsymbol{\mu}(\mathcal{K}, n_{it}) + \log(L) - \psi(L)$.

In our FTMM, the likelihood of a pixel value given a class $\mathcal{C}$ is the weighted average of the likelihoods of all its sub-classes $\mathcal{K} \in \mathcal{C}$.

$$\mathcal{L}(\boldsymbol{y}(k)|k \in \mathcal{C}) = \sum_{\mathcal{K} \in \mathcal{C}} \boldsymbol{\pi}(\mathcal{K}) \mathcal{L}(\boldsymbol{y}(k)|k \in \mathcal{K}) \tag{6.10}$$

In particular, we can write the likelihood for the water and the land classes given a pixel log-intensity $\boldsymbol{y}(k)$ as:

$$\mathcal{L}_W(\boldsymbol{y}(k)) = \sum_{\mathcal{K} \in \mathcal{C}_W} \boldsymbol{\pi}(\mathcal{K}) \frac{L^L}{\Gamma(L)} e^{L(\boldsymbol{y}(k) - x_{\mathcal{K}})} \exp(-Le^{\boldsymbol{y}(k) - x_{\mathcal{K}}}) \tag{6.11}$$

$$\mathcal{L}_L(\boldsymbol{y}(k)) = \sum_{\mathcal{K} \in \mathcal{C}_L} \boldsymbol{\pi}(\mathcal{K}) \frac{L^L}{\Gamma(L)} e^{L(\boldsymbol{y}(k) - x_{\mathcal{K}})} \exp(-Le^{\boldsymbol{y}(k) - x_{\mathcal{K}}}) \tag{6.12}$$

Our data term energy for a pixel k classified as water ($\boldsymbol{\ell}(k) = 1$) is $-\log(\mathcal{L}_W(\boldsymbol{y}(k))$ and $-\log(\mathcal{L}_L(\boldsymbol{y}(k))$ if the pixel is classified as land ($\boldsymbol{\ell}(k) = 1$).

Hence the data term for the whole image is:

$$U^I_{\text{data}}(\boldsymbol{\ell}, \boldsymbol{y}) = -\sum_k \left(\boldsymbol{\ell}(k) \log(\mathcal{L}_W(\boldsymbol{y}(k)) + (1 - \boldsymbol{\ell}(k)) \log(\mathcal{L}_L(\boldsymbol{y}(k))\right) \tag{6.13}$$

Regularization A regularization term ensures that the transitions between water and land are compatible with the gradients of the image, by penalizing the transitions that would occur where the gradient magnitude is low, or if the boundaries are not orthogonal to the gradient direction.

We want to minimize over the water boundaries the weighted total variation (eq 6.14) on the label field $\boldsymbol{\ell}$ that we assume to be continuous and whose gradient at location u is $\|\overrightarrow{\nabla \boldsymbol{\ell}}(u)\|$. The total variation is weighted by a weight $w_{\text{sym}}(u)$ defined in (6.15). This weight is symmetrical unlike its counterpart of chapter 5. Such a symmetrical weight, as it has been used in the original Grabcuts method, is simpler than an asymmetrical weight. In addition to this, it results in an undirected graph. When handling very large data such as remote sensing time series, the computational complexity of exact approaches can be a major issue, and some fast approximate methods, such as those based on electrical flows (Christiano et al., 2011; Lee et al., 2013; Yim, 2015) are available only when dealing with undirected graphs. The regularization energy is then given by equation 6.14:

$$U_{\text{reg}}(\boldsymbol{\ell}) = \beta \int_{u \in \mathbb{R}^2} w_{\text{sym}}(u) \|\overrightarrow{\nabla \boldsymbol{\ell}}(u)\| \cdot \mathrm{d}u \tag{6.14}$$

$$\text{with } w_{\text{sym}}(u) = \exp(-|\overrightarrow{\nabla \boldsymbol{\ell}}(u) \cdot \overrightarrow{\nabla I}(u)|/\lambda). \tag{6.15}$$

This symmetric weighting w_{sym} favors boundaries localizations that are aligned with the edges (strong gradients) of the image and is similar to its counterpart presented in section 5.2. The variables λ and β are parameters that allow adjusting the regularization and its sensitivity to the gradients.

In that way, the penalization for a transition between two neighboring pixels k_1 and k_2 such that $\boldsymbol{\ell}(k_1) \neq \boldsymbol{\ell}(k_2)$ is the same whatever the labeling. On the graph, this means that the two arcs between a given pair of nodes have the same capacity, which is equivalent to an undirected graph.

This regularization energy can be discretized:

$$U_{\text{reg}}(\boldsymbol{\ell}, \boldsymbol{y}) = \beta \sum_{k' \sim k} w_{\text{sym}}(k, k') \cdot |\boldsymbol{\ell}(k') - \boldsymbol{\ell}(k)| \tag{6.16}$$

with $w_{\text{sym}}(k, k') = \exp(-|(\boldsymbol{\ell}(k') - \boldsymbol{\ell}(k))(\boldsymbol{I}(k') - \boldsymbol{I}(k))|/\lambda)$, $k' \sim k$ means that k' is an 8-neighbor of k. In the case of pixels that are 8-neighbors of k but not 4-neighbors, λ is multiplied by $\sqrt{2}$. $(\boldsymbol{I}(k') - \boldsymbol{I}(k))$ is actually approximated using the ROEWA gradient computed on the intensity image for better robustness to noise, as explained in chapter 5.

Flux term In addition to the data term and the regularization term, a flux term is used to favor or penalize the transitions depending on the orientation and magnitude of the gradient. This term is similar to its counterpart in chapter 5. The gradient is expected to be strong on the water boundary $\partial\{\boldsymbol{\ell} = 1\}$ and oriented in the outward direction (inward for SWOT). **Figure** 6.4 illustrates this.

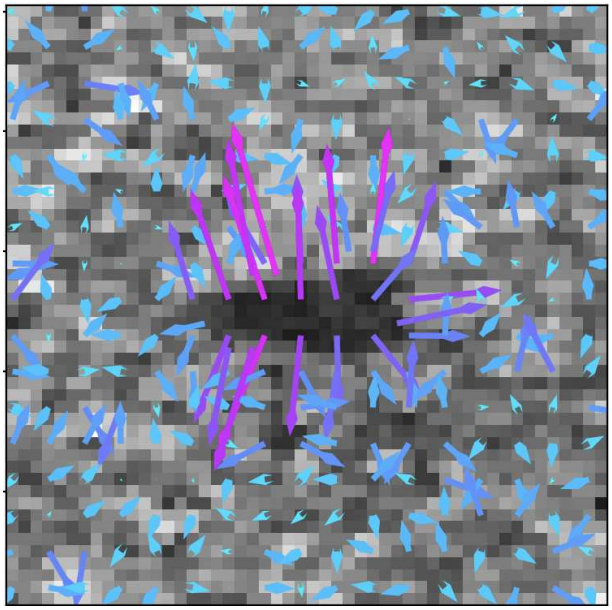


Figure 6.4 – Gradient on a SAR image. The arrows correspond to the local values of the gradient vector on the image of a small lake (crop of the Sentinel-1 GRD VV image of image Québec-12 presented in section 6.1.4). The lake is the dark structure. In this image, the gradient is much stronger on the boundaries of the lake, and oriented outward

.

Over the boundary $\partial\{\boldsymbol{\ell} = 1\}$, this criterion locally corresponds to maximizing the dot product between the gradient $\overrightarrow{\nabla}\boldsymbol{I}(u)$ and the unit outward normal vector of the segmentation $\{\boldsymbol{\ell} = 1\}$. Over each entire water body, the criterion can be expressed as the outward flux Φ of the gradient through the boundary $\partial\{\boldsymbol{\ell} = 1\}$.

This term cannot be directly modeled using graph-cuts, as it would result in negative costs for some transitions and thus fail to meet the sub-modularity condition required by (Kolmogorov and Zabih, 2004), as presented in section 4.1. Some approaches have been proposed by (Kolmogorov and Boykov, 2005), to use flux terms within a graph model, even with fluxes of non-differentiable vector fields. However, for our gradient flux which is differentiable, a much simpler way is to transform this boundary term into a region integral term that can be added pixel-wise to the terminal capacities on the graph as done in chapter 5:

$$\begin{aligned}\Phi &= \oint_{u\in\partial\{\boldsymbol{\ell}=1\}} \overrightarrow{\nabla}\boldsymbol{I}(u)\cdot\overrightarrow{n}(u)dl \\ &= \iint_{\{\boldsymbol{\ell}=1\}} \overrightarrow{\nabla}\cdot\overrightarrow{\nabla}\boldsymbol{I}(u)\,\mathrm{d}u\end{aligned} \tag{6.17}$$

where the second line comes from Ostrogradsky's divergence theorem.

Here, the Laplacian $\overrightarrow{\nabla}\cdot\overrightarrow{\nabla}\boldsymbol{I}(u)$ of the image can be approximated with a Laplacian of Gaussian (LoG) operator of parameter σ_L

$$\overrightarrow{\nabla}\cdot\overrightarrow{\nabla}\boldsymbol{I} \approx \mathrm{LoG}(\boldsymbol{I},\sigma_L) \tag{6.18}$$

that can be computed as the convolution of the image $\boldsymbol{I}$ with a precalculated LoG kernel. We call the resulting LoG image $\mathrm{LoG}_{\boldsymbol{y},\sigma_L}$.

The influence of the flux energy $U_{\text{flux}}(\boldsymbol{\ell})$ can be balanced with a scaling parameter η that adjusts its effect:

$$U_{\text{flux}}(\boldsymbol{\ell}, \boldsymbol{y}) = \int_{u\in\mathbb{R}^2, \boldsymbol{\ell}(u)=1} \eta \cdot \text{LoG}_{\boldsymbol{y},\sigma_L}(u)du, \tag{6.19}$$

which can be discretized as:

$$U_{\text{flux}}(\boldsymbol{\ell}, \boldsymbol{y}) = \eta \sum_k \boldsymbol{\ell}(k)\text{LoG}_{\boldsymbol{y},\sigma_L}(k). \tag{6.20}$$

The sign of η depends on the sensor: $\eta < 0$ for SWOT (water generally brighter than land) and $\eta > 0$ for Sentinel-1 (land mostly brighter than water).

Polygon term A last term U_{P} can be added for improved robustness. It prevents the classification as water of pixels outside of the initial *a priori* polygon (inside this polygon, $\boldsymbol{\ell}_0 = 1$, outside, $\boldsymbol{\ell}_0 = 0$).

$$U_{\text{P}}(\boldsymbol{\ell}, \boldsymbol{\ell}_0) = \sum_{k, \boldsymbol{\ell}(k) > \boldsymbol{\ell}_0(k)} K_C \tag{6.21}$$

Here, K_C is a parameter that is chosen large enough to prevent the minimal cut solution from cutting an edge with capacity K_C, but not too large to prevent numerical issues (overflow). This U_{P} term is similar to the centerline fidelity term U^C_{data} in section 5.2.

Resulting graph and minimization The resulting global energy can be written as follows:

$$\begin{aligned}\mathscr{E}(\boldsymbol{\ell}, \boldsymbol{y}) = \sum_k & -\left(\boldsymbol{\ell}(k)\log(\mathcal{L}_W(\boldsymbol{y}(k)) + \overline{\boldsymbol{\ell}}(k)\log(\mathcal{L}_L(\boldsymbol{y}(k))\right) \\ & +\beta \sum_{k'\sim k} w_{\text{sym}}(k,k') \cdot |\boldsymbol{\ell}(k') - \boldsymbol{\ell}(k)| \\ & +\eta \cdot \text{LoG}_{\boldsymbol{I},\sigma_L}(k) \\ & +\boldsymbol{\ell}(k) \cdot (1 - \boldsymbol{\ell}_0(k)) \cdot K_C\end{aligned} \tag{6.22}$$

where $\overline{\boldsymbol{\ell}}(k) = 1 - \boldsymbol{\ell}(k)$. The global energy is represented as a graph (which is possible as the regularization is sub-modular), and is minimized using the min-cut algorithm proposed in (Boykov and Kolmogorov, 2004).

The relationship between the local gradient value and the energy can be analyzed for a very simple situation by plotting the new symmetric regularization U_{reg} energy and the flux energy U_{flux} as in **Figure** 6.5. This figure considers the actual gradient and flux, which are actually approximated respectively by the ROEWA gradient and the integral of the Laplacian of Gaussian.

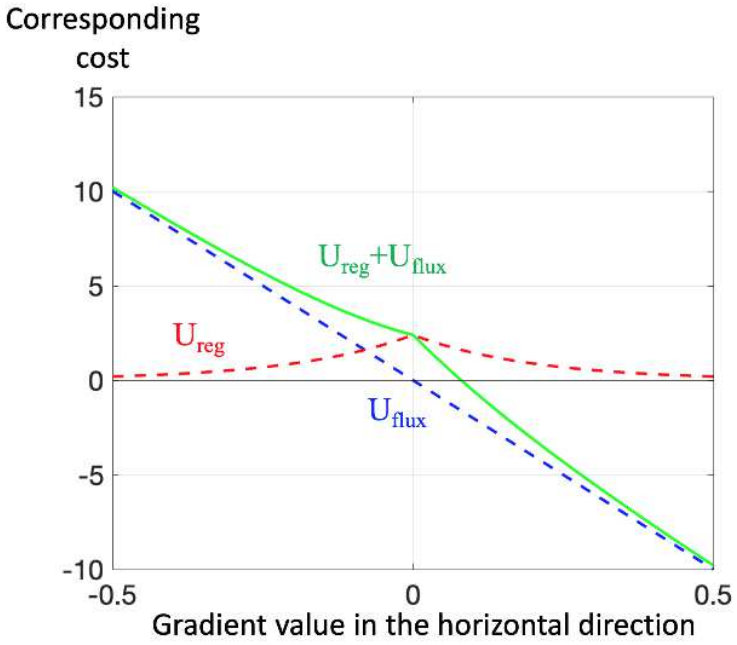


Figure 6.5 – Theoretical regularization energy U_{reg} (red), flux energy U_{flux} (blue), and sum $U_{reg} + U_{flux}$ (green) between a pixels k_1 and its right neighbor k_2 if $\ell(k_1) = 0$ and $\ell(k_2) = 1$, in the case of Sentinel-1 images. The curves would be flipped horizontally for $\ell(k_1) = 1$ and $\ell(k_2) = 0$.

The resulting partition of the image gives the new labels $\ell(k)$ for every pixel in the image.

6.1.4 Experiments

In this section, we evaluate the performances and characterize the interest of our method in segmenting lakes in Sentinel-1 GRD images. We chose the images to be as representative as possible of various situations in terms of environment, shape, and wind condition, but the comprehensive tuning of the parameters and validation of the algorithm are beyond the scope of our experiments.

All the results below have been obtained using our code, which is based on an adaptation of OpenCV's implementation of GrabCuts, and uses (Boykov and Kolmogorov, 2004)'s implementation of the min-cut algorithm.

The metrics used for comparison with the ground truth are the same as in section 5.3 where they are presented.

6.1.4.1 Data

We applied the algorithm on 3 kinds of Sentinel-1 GRD images:

1. Sentinel-1 GRD image with VV polarization, which has a higher contrast between water and flat shores, but is more prone to Bragg resonance.

2. Sentinel-1 GRD image with VH polarization, which is robust to Bragg resonance, but with a lower contrast between water and flat shores.

3. Pixel-wise geometric mean of VV and VH polarized images, which is a compromise between the properties of VV and VH images, as presented in appendix A.

The parameters are the same for all images, except for L that is different for VVVH images (see appendix A).

Table 6.1 – Parameters used for the experiments

Number of sub-classes		Regularization		Flux	
$n_{\mathcal{C}_W}$ (Water)	$n_{\mathcal{C}_L}$ (Land)	β	λ	σ_L	η
2	5	2.4	0.2	4	20

We applied our method on 10 Sentinel-1 images taken from 7 times series, which correspond to 7 lakes with various shape, extent, and environment. Each time series consists of images acquired by the same sensor with the same orbital position and direction (ascending or descending).
The images are presented in table 6.2. The images are named after the time series and the position of the image in the stack. For example, image "Chad-10" is the 10-th image of the time series "Chad". Each of these images is associated with manually defined ground truth. In our ground truth, the waterbodies other than the main lake are classified as "uncertain". In this way, the metric will not penalize their detection as water nor their detection as land.

Table 6.2 – Sentinel-1 GRD images used for our experiments

Image #-name	Lake name	Location	Date	Lake size	Image size (pixels)
1 - Chad-10	unnamed	Chad	2017-01-22	130 m × 80 m	200 × 200
2 - Vioreau-20	Vioreau reservoir & Vioreau pond	Joué-sur-Erdre, Pays-de-la-Loire, France	2017-11-30	3 km × 1 km 1 km × 400 m	386 × 618
3 - Vioreau-40	Vioreau reservoir & Vioreau pond	Joué-sur-Erdre, Pays-de-la-Loire, France	2019-09-09	3 km × 1 km 1 km × 400 m	386 × 618
4 - Sajnam-10	Sajnam reservoir	Chandawali, Uttar Pradesh, India	2017-06-03	4 km × 2 km	769 × 718
5 - Sajnam-30	Sajnam reservoir	Chandawali, Uttar Pradesh, India	2018-02-22	4 km × 2 km	769 × 718
6 - Rougé-20	unnamed	Rougé, Pays-de-la-Loire, France	2018-02-22	200 m × 130 m	150 × 150
7 - Québec-12	unnamed	Québec, Canada	2018-07-22	180 m × 80 m	175 × 270
8 - Der-10	Lake Der	Grand-Est, France	2019-12-17	10 km × 7 km	1316 × 1188
9 - Der-37	Lake Der	Grand-Est, France	2018-03-28	10 km × 7 km	1316 × 1188
10 - Feins-10	Boulet pond	Feins, Brittany, France	2017-11-30	2 km × 2 km	500 × 510

6.1.4.2 Results

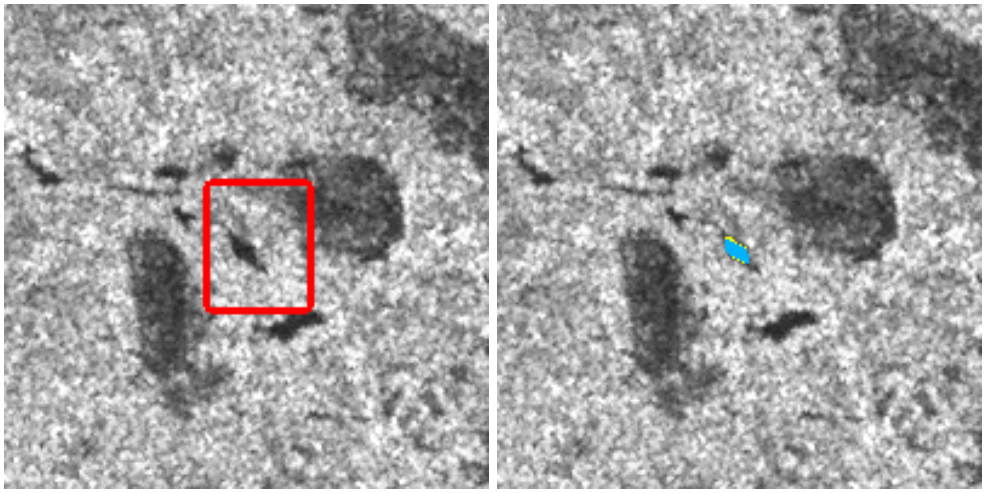
A summary of the metrics for each image and each polarization channel is provided in table 6.1.4. The proposed method is compared with a reference method (REF).

The reference method is a graph cut approach based on an MRF model with an uniform regularization $\beta_{\text{ref}} = 1$ and a quadratic data term (derived from a Gaussian approximation of the distribution of the VVVH combination). The data term is based on the true mean reflectivities of the land class and water class according to the ground truth. A post-processing step removes the water pixels detected outside of the *a priori* polygon. In this way, the reference method takes advantage of both the knowledge of true mean reflectivities of land and water classes and of the *a priori* polygon. This reference method could not be used for actual segmentation as it requires knowledge of the true mean reflectivities of both classes.

The metrics are presented in table 6.3. Note that as the ground truth images have been traced manually, very small differences (below 1%) in a metric between two results may not be considered significative. In our results, VV gives slightly better results than VVVH and much better results than VH in situations without a strong Bragg phenomenon, but behaves poorly when it is present (image Der-37). Overall, VVVH appears as a good compromise between VV and VH.

Table 6.3 – Summary of the metrics for each result

Number	Name	Channel	Pr (%)	Rec (%)	FPR (%)	F-Score (%)	ER (%)	MCC (%)
1	Chad-10	VV	77.78	98.25	0.04	86.82	29.82	87.40
		VH	42.86	94.74	0.18	59.02	131.58	63.65
		VVVH	**83.82**	**100.00**	**0.03**	**91.20**	**19.30**	**91.54**
		REF	46.34	**100.00**	0.17	63.33	115.79	68.02
2	Vioreau-20	VV	**98.66**	93.79	**0.05**	**96.16**	**7.49**	**96.06**
		VH	88.93	97.22	0.44	92.89	14.88	92.72
		VVVH	96.14	95.90	0.14	96.02	7.95	95.87
		REF	93.49	**98.13**	0.25	95.76	8.70	95.63
3	Vioreau-40	VV	**97.37**	93.01	**0.12**	95.14	9.50	94.94
		VH	94.18	92.69	0.28	93.43	13.04	93.12
		VVVH	96.67	92.95	0.15	94.77	10.26	94.54
		REF	96.37	**95.11**	0.17	**95.74**	**8.47**	**95.54**
4	Sajnam-10	VV	**98.30**	**96.15**	**0.24**	**97.22**	**5.51**	**96.83**
		VH	93.50	91.50	0.92	92.49	14.86	91.43
		VVVH	97.95	96.12	0.29	97.03	5.89	96.61
		REF	95.84	95.73	0.60	95.79	8.42	95.18
5	Sajnam-30	VV	**97.78**	94.28	**0.14**	**96.00**	**7.86**	**95.75**
		VH	86.83	96.96	0.99	91.62	17.74	91.18
		VVVH	96.55	95.34	0.23	95.94	8.06	95.67
		REF	92.08	**97.83**	0.56	94.87	10.58	94.57
6	Rougé-20	VV	87.67	**100.00**	0.04	93.43	14.06	93.61
		VH	**92.19**	92.19	**0.02**	92.19	**15.62**	92.17
		VVVH	91.18	96.88	0.03	93.94	12.50	93.96
		REF	88.89	**100.00**	0.04	**94.12**	**12.50**	**94.26**
7	Québec-12	VV	79.49	**100.00**	0.03	88.57	25.81	89.14
		VH	72.09	**100.00**	0.05	83.78	38.71	84.89
		VVVH	72.94	**100.00**	0.05	84.35	37.10	85.38
		REF	**86.11**	**100.00**	**0.02**	**92.54**	**16.13**	**92.79**
8	Der-10	VV	**97.02**	96.63	**0.54**	**96.82**	**6.34**	**96.24**
		VH	75.82	99.03	5.80	85.89	32.55	83.97
		VVVH	91.95	98.01	1.57	94.89	10.56	93.98
		REF	85.67	**99.29**	3.05	91.98	17.32	90.73
9	Der-37	VV	92.79	34.91	**0.92**	50.73	67.80	50.37
		VH	95.18	97.66	1.67	**96.41**	**7.28**	**95.19**
		VVVH	**95.43**	95.22	1.54	95.33	9.34	93.75
		REF	91.13	**98.46**	3.23	94.66	11.12	92.88
10	Feins-10	VV	**95.43**	98.03	**0.11**	**96.71**	**6.66**	**96.65**
		VH	83.06	99.15	0.46	90.40	21.07	90.53
		VVVH	92.77	98.69	0.17	95.64	9.00	95.58
		REF	87.48	**99.43**	0.32	93.07	14.80	93.11

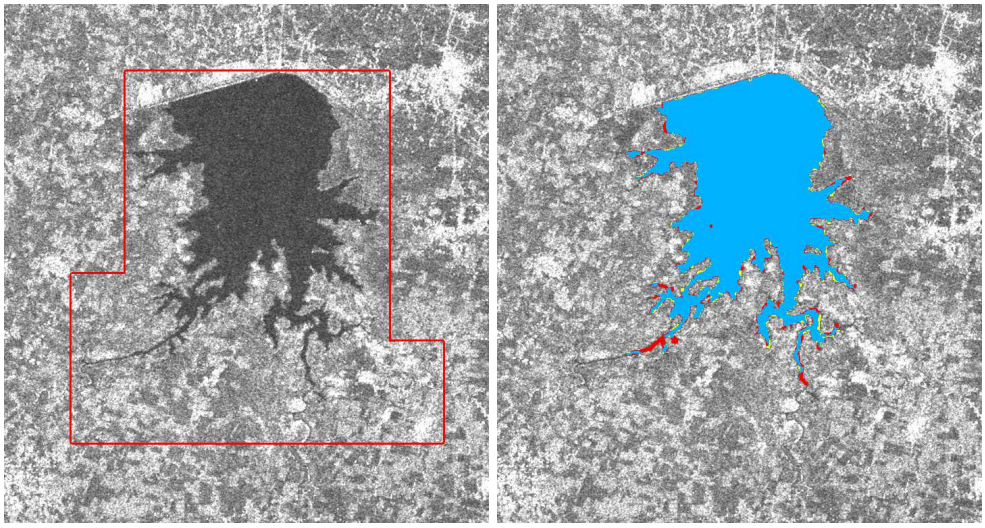


(a) A priori polygon **(b)** Resulting classification error.

Figure 6.6 – Mask and result for image 1: Chad-10 (VVVH). On the classification error image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

Example 1: Chad-10 This image covers a small lake east of Lake Chad, surrounded by dark structures. Note that in this example, the flux term is necessary in order to detect the lake. Our experiments show that running our model without the flux term (i.e. with $\eta = 0$) leads to no water being detected at all for the VV, VH and VVVH images: in these cases, the whole image is classified as land. Indeed, without this term, the detection of water is difficult because the initial polygon contains more bright pixels than dark pixels.

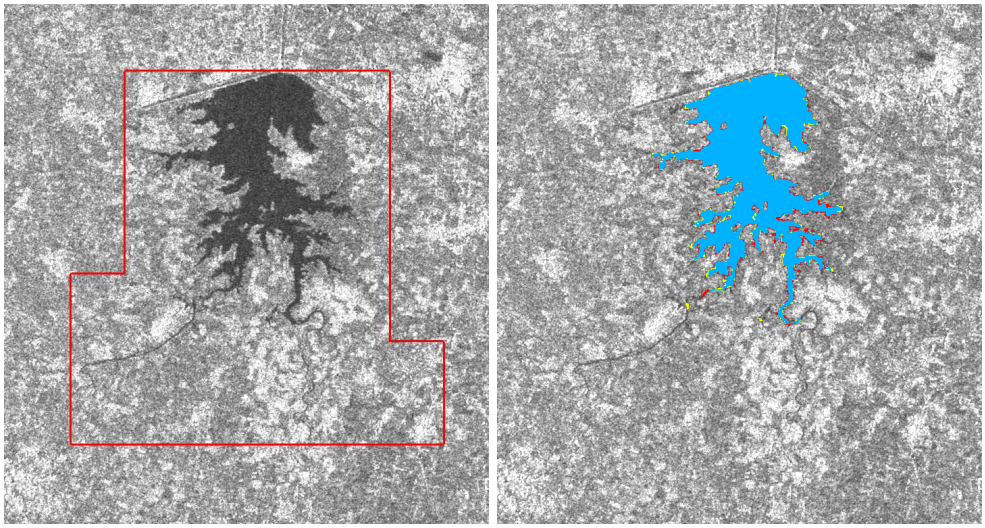
The same can be observed when running the detection with $\eta = 0$ on other small lakes image. No water is detected in the VV, VH and VVVH Québec-12 images without the flux term, nor in VH and VVVH Rougé-20 images.



(a) A priori polygon **(b)** Resulting classification error.

Figure 6.7 – Mask and result for image 1: Sajnam-10 (VVVH). On the classification error image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

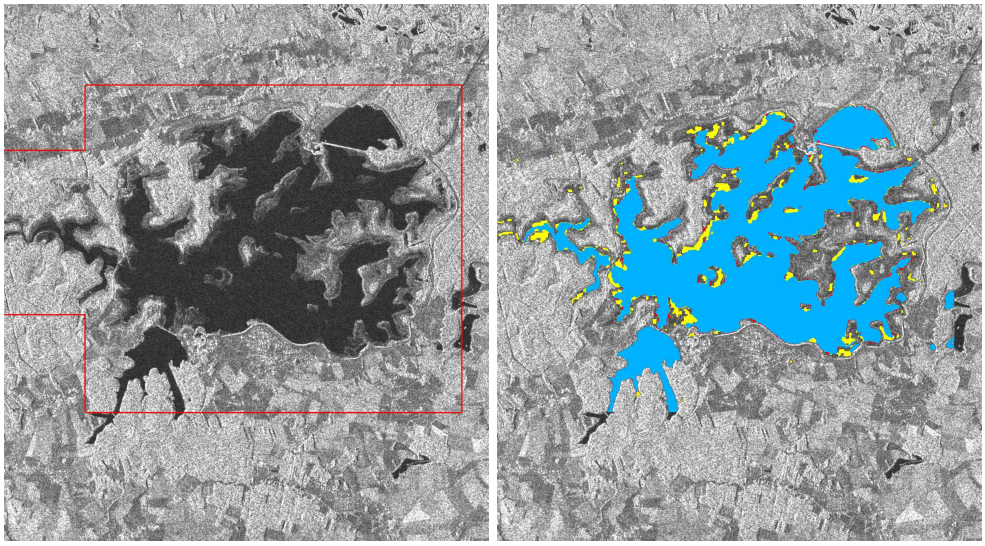
Example 2: Sajnam-10 This image corresponds to Sajnam reservoir on 2017-06-03.



(a) A priori polygon **(b)** Resulting classification error.

Figure 6.8 – Mask and result for image 1: Sajnam30 (VVVH). On the classification error image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

Example 3: Sajnam-30 This image corresponds to Sajnam reservoir on 2018-02-22.

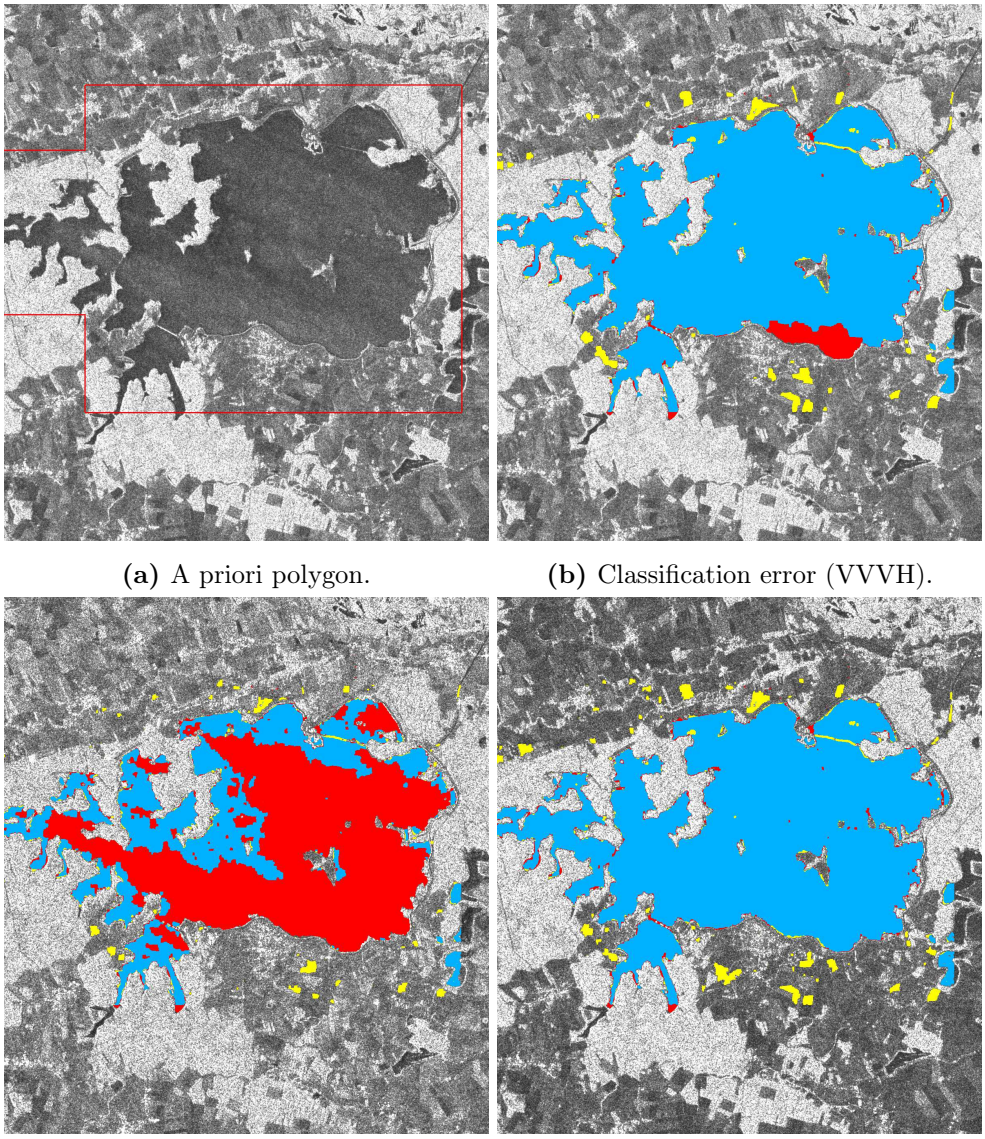


(a) A priori polygon **(b)** Resulting classification error.

Figure 6.9 – Mask and result for image 1: LacDer-10 (VVVH). On the classification errors image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

Example 4: Der-10

Example 4: Der-37 **Figure** 6.10 illustrates the effects of Bragg scattering on the water detection results using VVVH, VV and VH images. The a priori polygon is displayed on the VVVH image.



(a) A priori polygon. **(b)** Classification error (VVVH).

(c) Classification error (VV). **(d)** Classification error (VH).

Figure 6.10 – Mask and result for image 1: LacDer-37. On the classification errors image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

6.1.5 Conclusion

We proposed a water extraction method for SAR images that only relies on an *a priori* polygon, but does not require *a priori* knowledge about the reflectivity distributions other than the fact that water is generally darker than land (brighter for SWOT). This last information is integrated into the flux energy.

This method gives good detection results in most cases, but can still fail for very difficult situations (bright water due to Bragg phenomenon, transient dark fields).

The computing speed could be improved by using a minimal cut approach that takes into account the similarity between the graphs built for each iteration, such as the one proposed by (Kohli and Torr, 2005), or by using approximate approaches.

Summary: Lake extraction on SAR image with GrabCut

In order to detect water in SAR images without any prior knowledge about water and land distributions, we propose an adaptation of the Grab Cut approach that takes an *a priori* polygon as input. Our method models both classes as Fisher-Tippett mixture models (FTMMs) and alternatively runs three steps after an initialization:

1. A pixel assignment step (section 6.1.3.3)
2. A parameter learning step (section 6.1.3.4)
3. A segmentation step (section 6.1.3.5)

The segmentation step has been adapted to our problem. We minimize a global energy that is the sum of multiple terms:

1. A data term, that depends on the parameters of the FTMMs
2. A symmetric regularization term
3. A flux term, that favors water boundaries with a large outward gradient flux. This term is not used in the original GrabCut method and is needed in order to be able to detect small lakes.
4. A polygon term, that prevents pixels outside of the bounding polygon from being classified as water.

   This method uses a bounding polygon as an input, but does not require any prior information on the water and land reflectivities, other than that water is darker (brighter for SWOT) than land.

6.2 Multitemporal and multi-sensor adaptations of the method

6.2.1 Background on multi-temporal and multi-sensor methods

While multitemporal approaches for SAR images have been widely used to address problems such as crop classification (see (Skriver et al., 2011) or (Van Tricht et al., 2018), which uses machine learning (random forests)) or land cover mapping (Waske and Braun, 2009). Indeed, crop growth as well as plowing and harvesting operations result in seasonal changes of the reflectivity of the field. Contrariwise, there is generally no such characteristic seasonal change for water and land surfaces.

Most multitemporal approaches for water detection have been limited to specific applications such as flood detection, considered as a kind of change detection (see for example (Martinis et al., 2011), or (Landuyt et al., 2019) for an assessment of existing approaches) or to process the images separately without using actual joint segmentation (e.g. (Uddin et al., 2019)). A simple yet efficient approach for temporal regularization proposed in (Peña-Luque et al., 2021) consist in first computing water masks separately for all the dates, then regularizing these water masks, for example using a temporal sliding window.

The information from the phase of repeat pass single look complex images can be used as well. There are specific kind of multitemporal SAR water detection approaches based on temporal coherency between the dates: temporal coherence between the complex backscattered signals is stronger for land than for water (see for example (Brisco et al., 2017; Canisius et al., 2019), that uses coherence for separating water and vegetation in RADARSAT-2 SAR images).

In addition to these multitemporal methods that combine images from the same (or identical) sensor at multiple dates, multi-sensor approaches combines images taken from different sensors, potentially at different dates. In particular, the combination of optical and SAR images is promising.

To this end, approaches that combine SAR and optical images have been proposed. For instance, (Irwin et al., 2017) propose a pixel-based decision tree that combines SAR, optical and airborne LiDAR images. (Rambour et al., 2020) combine SAR Sentinel-1 and optical Sentinel-2 within a deep learning approach for flood detection.

One theoretical framework that has been used for the fusion of information from different sensor is Dempster and Shafer's evidence theory (Dempster, 1967; Shafer, 1976). Its use for data fusion in remote sensing has been studied by (Le Hegarat-Mascle et al., 1997) for unsupervised classification and by (Ouled Sghaier et al., 2018) for flood mapping based on time series SAR images.

In the following, we propose a new approach that jointly detects water surfaces on all the images of the time series, using a temporal regularization to improve the detection.

6.3 2D+T GrabCut of SAR images with temporal regularization for lake detection within an *a priori* mask

6.3.1 Introduction

This section introduces a multitemporal extension of the GrabCut segmentation approaches presented in section 6.1.3. It features a temporal regularization that aims at preventing false positive detection in case of transient dark land (or transient bright structures for SWOT) and false negative detection in case of Bragg resonance (or dark water in SWOT).

The underlying assumption behind this temporal regularization is that there is a temporal regularity in the water surface extent. This is generally true as changes in the water surface extent are slow compared to the revisit time of the sensor, and only concern limited areas of the water surface. However, more brutal changes can happen, especially in the case of artificial dams and reservoirs, and flloding events. For this reason, the temporal regularization has to be robust to rapid temporal changes. This can be done using a CRF regularization that takes the reflectivity changes into account.

Here, the input data is a stack of T SAR images, sorted in chronological order or, if available from external information, sorted by water level. $\boldsymbol{I}(k,t)$ is the intensity of the pixel (k) at date t, and $\boldsymbol{y}(k,t)$ is its logarithm. The goal is to obtain one water label map $\boldsymbol{\ell}(.,t)$ for each image in the temporal stack. Note that $\boldsymbol{I}(\cdot,t)$ corresponds to the t-th image of the stack.

6.3.2 Temporal adaptation of the Grabcut approach

The 2D+T method we present here for the segmentation of a SAR time series is a multitemporal extension of the 2D method that is presented in section 6.1.3. Likewise, water and land are here modeled with one Fisher-Tippett Mixture Model (FTMM) each for the entire time series, but the clustering operations are applied on each temporal pixel (k,t) separately.

The initialization is done using an initial labeling $\boldsymbol{\ell}_0(k,t) = \boldsymbol{\ell}_0(k)\forall t$ defined from an *a priori* polygon.

After an initialization step that is identical to its counterpart in section 6.1.3, 3 subsequent steps are repeated for a given number of iterations $n_{\max}$:

1. A pixel assignment step that is identical to its counterpart in section 6.1.3.3.
2. A parameter learning step that is identical to its counterpart in section 6.1.3.4.
3. A segmentation step, that is adapted from the approach presented 6.1.3.5, but features a new temporal regularization term in addition to the spatial regularization term.

The new segmentation step is presented in the next section.

6.3.2.1 Segmentation with a temporal regularization

The segmentation step consists in minimizing a global energy $\mathscr{E}_{\mathrm{MT}}(\boldsymbol{\ell},\boldsymbol{y})$ that depends on the 2D+T label field $\boldsymbol{\ell}(k,t)$. This global energy is equal to the sum of the energies $\mathscr{E}(\boldsymbol{\ell}(\cdot,t),\boldsymbol{y}(\cdot,t)$

for every image in the stack plus a temporal regularization term between every pair of consecutive images:

$$\mathscr{E}_{\mathrm{MT}}(\boldsymbol{\ell}, \boldsymbol{y}) = \sum_{t=1}^{T} \mathscr{E}(\boldsymbol{\ell}(\cdot, t), \boldsymbol{y}(\cdot, t)) + \sum_{t=1}^{T-1} U_{\mathrm{TR}}(\boldsymbol{y}(\cdot, t), \boldsymbol{y}(\cdot, t+1)). \tag{6.23}$$

where $\mathscr{E}(\boldsymbol{\ell}(\cdot, t), \boldsymbol{y}(\cdot, t))$ corresponds to the global energy for one image, as defined in equation 6.22.

Temporal regularization The temporal regularization $U_{\mathrm{TR}}(\boldsymbol{y}(\cdot, t), \boldsymbol{y}(\cdot, t+1))$ between the images at date t and $t+1$ is the sum of the local regularization energies for every pixels. This regularization energy in (k,t) depends on the difference between the log-intensities $\boldsymbol{y}(k, t)$ and $\boldsymbol{y}(k, t+1)$:

$$U_{\mathrm{TR}}((\cdot, t), (\cdot, t+1)) = \beta_T \sum_{k} |\boldsymbol{\ell}(k, t) - \boldsymbol{\ell}(k, t+1)| \exp(-|\frac{\boldsymbol{y}(k, t) - \boldsymbol{y}(k, t+1)|}{\lambda_T}) \tag{6.24}$$

Here, β_T is a tuning parameter for the regularization. The temporal change $\boldsymbol{y}(k, t) - \boldsymbol{y}(k, t+1)$ can be regularized, for example, using spatial smoothing by convolving with a Gaussian kernel of parameter σ_T, to limit the consequences of the speckle noise, or a better denoising method.

The resulting graph is a stack of T layers of non-terminal nodes. Each is of the shape of the image and similar to the single-layer graph in the single-date situation of the previous chapter. A theoretical assessment of the properties of such a multilayer model and its relationship with majority voting with naive Bayes segmentation can be found in (Lermé et al., 2020).

6.3.3 Experiments

We tested our method on the seven time series from which the images used in the previous section have been taken. This allows a comparison between the 2D approach of the previous chapter and the 2D+T approach presented here.

Our code for the 2D+T Grabcut derives from the code from the 2D Grabcut presented in section 6.1.3 and uses elements of code from the implementation of 3D Grabcut developed by (Yoruk et al., 2018) for medical applications.

We used the same parameter values for our experiments on the 2D+T Grabcut as for their 2D counterpart, and the values for the extra parameters β_T and λ_T have been set empirically. Note that β_T and λ_T cannot be directly compared to their spatial counterparts as the computations of the spatial and of the temporal gradients are different.

Table 6.4 – Parameters used for the experiments, with temporal regularization

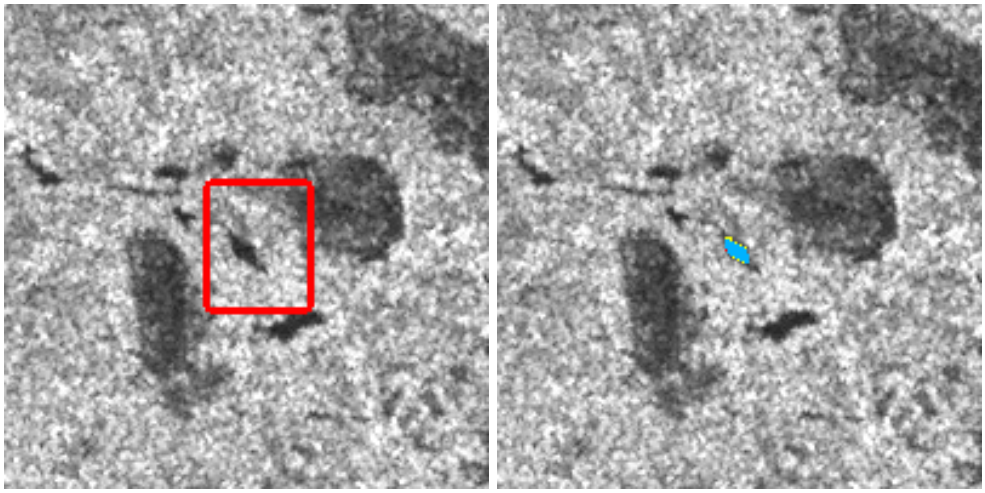
Number of sub-classes		Regularization		Temporal regularization		Flux	
$n_{\mathcal{C}_W}$	$n_{\mathcal{C}_L}$	β	λ	β_T	λ_T	σ_L	η
2	5	2.4	0.2	12	2.5	4	20

6.3.3.1 Results

The results are given in table 6.5. For each image, the line "VVVH 1T" recalls the result obtained using the single date method presented in the previous section.

Table 6.5 – Summary of the metrics using the multitemporal approach

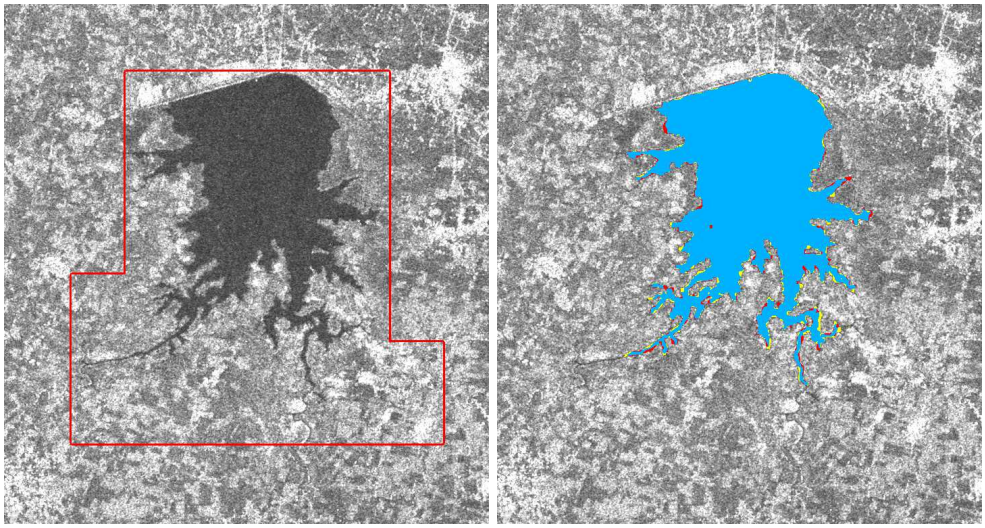
Number	Name	Channel	Pr (%)	Rec (%)	FPR (%)	F-Score (%)	ER (%)	MCC (%)
1	Chad-10	VV	75.68	98.25	0.05	85.50	33.33	86.20
		VH	72.97	94.74	0.05	82.44	40.35	83.12
		VVVH	**84.85**	98.25	**0.03**	91.06	**19.30**	91.29
		VVVH 1T	83.82	**100.00**	**0.03**	**91.20**	**19.30**	**91.54**
2	Vioreau-20	VV	**97.76**	92.00	**0.10**	94.79	10.11	94.60
		VH	88.23	93.60	0.60	90.83	18.89	90.42
		VVVH	96.09	93.07	0.18	94.55	10.72	94.31
		VVVH 1T	96.14	**95.90**	0.14	**96.02**	**7.95**	**95.87**
3	Vioreau-40	VV	**98.82**	92.08	**0.05**	95.33	9.02	95.18
		VH	98.18	92.43	0.08	95.21	9.29	95.04
		VVVH	97.43	**93.54**	0.12	**95.44**	**8.93**	**95.25**
		VVVH 1T	6.67	92.95	0.15	94.77	10.26	94.54
4	Sajnam-10	VV	97.81	96.78	0.31	**97.29**	**5.39**	**96.91**
		VH	86.83	**98.71**	2.16	92.39	16.26	91.46
		VVVH	97.08	97.45	0.42	97.26	5.48	96.87
		VVVH 1T	**97.95**	96.12	**0.29**	97.03	5.89	96.61
5	Sajnam-30	VV	96.52	95.68	**0.23**	**96.10**	**7.77**	**95.84**
		VH	86.06	**97.72**	1.06	91.52	18.11	91.12
		VVVH	95.46	96.52	0.31	95.99	8.07	95.72
		VVVH 1T	**96.55**	95.34	**0.23**	95.94	8.06	95.67
6	Rougé-20	VV	86.49	**100.00**	0.04	92.75	15.62	92.98
		VH	89.71	95.31	**0.03**	92.42	15.62	92.44
		VVVH	**91.43**	**100.00**	**0.03**	**95.52**	**9.38**	**95.61**
		VVVH 1T	91.18	96.88	**0.03**	93.94	12.50	93.96
7	Québec-12	VV	82.67	**100.00**	**0.03**	**90.51**	**20.97**	**90.91**
		VH	**83.33**	96.77	**0.03**	89.55	22.58	89.79
		VVVH	80.52	**100.00**	**0.03**	89.21	24.19	89.72
		VVVH 1T	72.94	**100.00**	0.05	84.35	37.10	85.38
8	Der-10	VV	**98.47**	95.42	**0.27**	**96.92**	**6.06**	**96.39**
		VH	76.73	**98.99**	5.51	86.45	31.03	84.58
		VVVH	94.19	97.49	1.10	95.81	8.53	95.05
		VVVH 1T	91.95	98.01	1.57	94.89	10.56	93.98
9	Der-37	VV	93.39	10.68	**0.25**	19.17	90.07	27.07
		VH	93.36	97.98	2.36	**95.61**	**8.99**	**94.13**
		VVVH	**96.68**	49.64	0.58	65.60	52.06	63.46
		VVVH 1T	91.13	**98.46**	3.23	94.66	11.12	92.88
9	Feins-10	VV	**94.64**	98.49	**0.13**	**96.53**	**7.09**	**96.47**
		VH	84.42	**99.34**	0.41	91.28	18.99	91.38
		VVVH	92.10	98.99	0.19	95.42	9.50	95.38
		VVVH 1T	92.77	98.69	0.17	95.64	9.00	95.58



(a) A priori polygon **(b)** Resulting classification error.

Figure 6.11 – Mask and result for image 1: Chad-10 (VVVH). On the classification errors image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

Example 1: Chad-10 This image corresponds to a small lake east of Lake Chad, surrounded by dark structures.

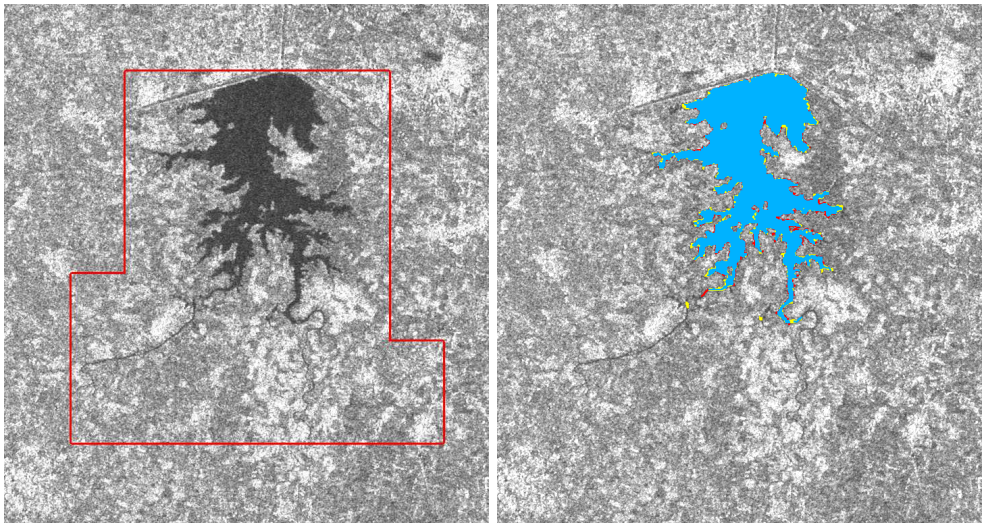


(a) A priori polygon **(b)** Resulting classification error.

Figure 6.12 – Mask and result for image 1: Sajnam-10 (VVVH). On the classification errors image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

Example 2: Sajnam-10 This image corresponds to Sajnam reservoir on 2017-06-03.

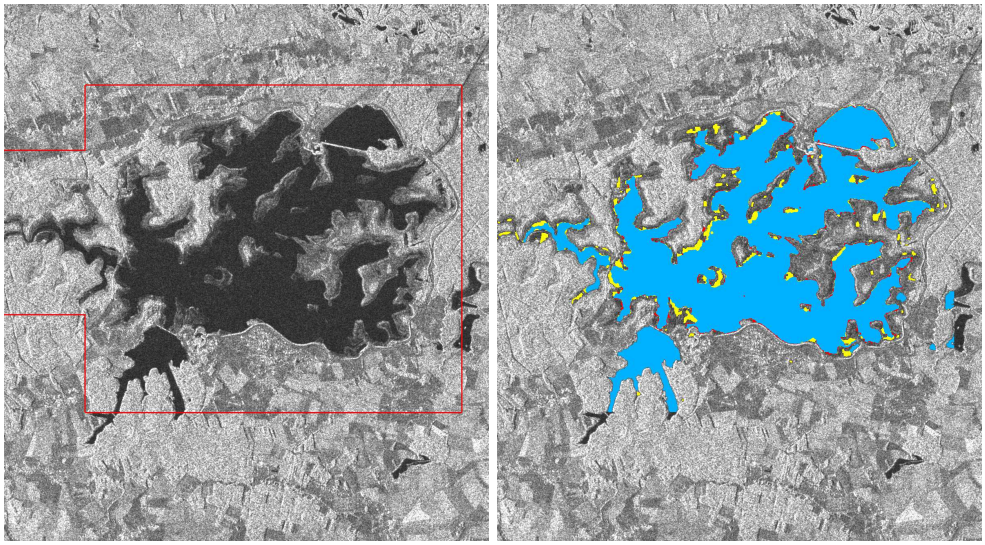
Example 3: Sajnam-30 This image corresponds to Sajnam reservoir on 2018-02-22.



(a) A priori polygon **(b)** Resulting classification error.

Figure 6.13 – Mask and result for image 1: Sajnam30 (VVVH). On the classification errors image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

Example 4: Der-10 For this multitemporal result, just like for Der-37, we could not process the whole time series at once because of technical limitations related to available memory. Instead, we processed a smaller time series of 7 images, with the image on which the evaluation is done plus the three preceding and three following dates.

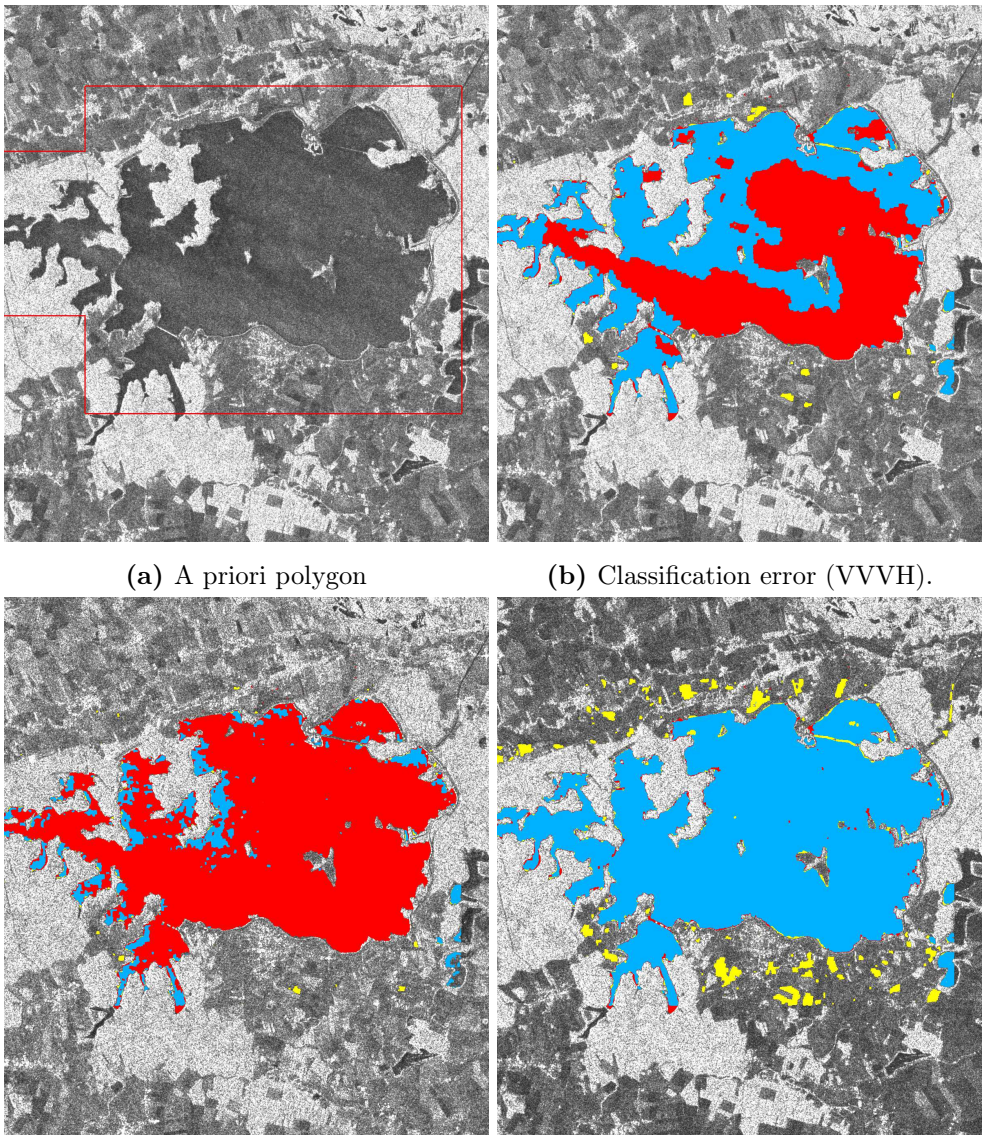


(a) A priori polygon **(b)** Resulting classification error.

Figure 6.14 – Mask and result for image 1: LacDer-10 (VVVH). On the classification errors image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

Example 4: Der-37 For this multitemporal result, as well as for Der-10, we could not process the whole time series at once because of technical limitations related to available memory.

Instead, we processed a smaller time series of 7 images, with the image on which the evaluation is done plus the three preceding and three following dates.



(a) A priori polygon **(b)** Classification error (VVVH).

(c) Classification error (VV). **(d)** Classification error (VH).

Figure 6.15 – Mask and result for image 1: LacDer-37. On the classification errors image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

6.3.3.2 Discussion

Our results show that the 2D+T approach can lead to slightly better segmentation results than its 2D counterpart, but as the parameter estimation is done for the whole time series at each iteration, the resulting FTMMs may lead to a low likelihood for water reflectivities that are very different from the average, such as in image Der-37.
Overall, the temporal regularization nevertheless improves the detection of contours between water and land while preserving the temporal evolution of the shape of the lake (see for example the water area differences between the Sajnam images or the Der images).

While this approach does not rely on *a priori* knowledge about the reflectivities of the water and land classes, it can be noted that the final components of the FTMMs are quite similar from one time series to another, at least for the Sentinel-1 images we used.

In the next section, we propose an approach that uses these FTMM values as input to extract water surfaces on a time series of Sentinel-1 and Sentinel-2 images.

6.4 Joint 2D+T segmentation of SAR and optical images

6.4.1 Introduction

This section presents a simple approach that aims at extracting water surfaces on combined time series of Sentinel-1 and Sentinel-2 images. Unlike the method described in previous section, there is no guiding geometric information: no *a priori* polygon is used, but approximate parameters for the mixtures are assumed to be known *a priori*, for example using the methods presented in the previous sections of this chapter.

When available, the use of optical images allow for more accurate water detection and have been generally preferred to SAR images in water monitoring applications for that reason. However, they lack the temporal regularity of SAR images as their usability is affected by the cloud coverage.

In this chapter, we propose to combine SAR images and optical images in the same temporal stack and to extract the water surfaces from the resulting time series.

First, in sections 6.4.1.1 and 6.4.1.2, we introduce Sentinel-2 images and the indexes (combination of spectral bands) that are usually used for water detection. Then the proposed method is presented in 6.4.2. Finally, we test our method on a combined Sentinel-1 and Sentinel-2 time series in 6.4.3.

6.4.1.1 Sentinel-2 optical images

Sentinel-2 is a constellation of polar-orbiting optical satellites, with currently two operational satellites. This results in a theoretical revisit time of 5 days at the equator and 2-3 days at mid-latitudes. However, the images can only be used in the absence of clouds, which leads to a lack of information for periods with high cloud coverage, and the time between two usable images can be much longer. Each Sentinel-2 image product consists of 13 images, each corresponding to a given spectral band:

- 4 spectral bands with 10m spatial resolution: Blue, Green, Red, and Near InfraRed (NIR).
- 6 spectral bands with 20m spatial resolution, including two Short Wave InfraRed bands (SWIR).
- 3 spectral bands with 60m spatial resolution.

The most relevant spectral bands for water detection are the following, with the SNR given for the reference radiance of the spectral band:

Band number	Band name	Resolution (m)	Centralwavelength (nm)	Bandwidth (nm)	SNR
B02	Blue	10	490	65	154
B03	Green	10	560	35	168
B04	Red	10	665	30	142
B08	NIR	10	842	115	172
B11	SWIR-1	20	1610	90	100
B11	SWIR-2	20	2190	180	100

6.4.1.2 Combination of spectral bands for water detection in Sentinel-2 images

Unlike SAR images, Sentinel-2 optical images have a very high SNR that enables the direct detection of water without the need for a denoising preprocessing step or regularization approaches. Multiple methods based only on the values of individual pixels have been proposed for optical images (Sentinel-2 or other sensors, such as Landsat). Most are based on indexes, which are pixel-wise combinations of the values of several spectral bands, and rely on thresholding or clustering methods for the classification (see for example (Cordeiro et al., 2021)). An extensive comparison of water detection approaches and indexes for optical images can be found in (Yang et al., 2020).

The main indexes for detecting water in optical images are presented below. Note that two different expressions have been proposed for the NDWI:

Name	Proposed by	Formula
NDWI	(McFeeters, 1996)	$(\text{Green} - \text{NIR})/(\text{Green} + \text{NIR})$
NDWI	(Rogers and Kearney, 2004)	$(\text{Red} - \text{SWIR1})/(\text{Red} + \text{SWIR1})$
MNDWI	(Xu, 2006)	$(\text{Green} - \text{SWIR1})/(\text{Green} + \text{SWIR1})$
MBWI	(Wang et al., 2018)	$3 \cdot \text{Green} - \text{Red} - \text{NIR} - \text{SWIR1} - \text{SWIR2}$

6.4.2 Proposed method

We propose a method that operates on a combined 2D+T time series of Sentinel-1 SAR images and Sentinel-2 optical images. For each modality, only a single channel is used: either a measured channel or a synthetic combination (e.g. VV.VH product for SAR images or NDWI for optical images). The stack contains both SAR and cloud-free optical images, and all the images are stacked in chronological order or, if available from external data, by water level. A vector $\boldsymbol{S}$ associates the date with the corresponding sensor: $\boldsymbol{S}(t) = 0$ if the t-th image is an optical image, $\boldsymbol{S}(t) = 1$ if it is a SAR image. We call $\boldsymbol{I}_{\text{OR}}$ the resulting stack of images. $\boldsymbol{I}_{\text{OR}}(\cdot, \cdot, t)$ is a log-transformed SAR intensity image if $\boldsymbol{S}(t) = 1$ and an optical image (e.g. NDWI index) if $\boldsymbol{S}(t) = 0$.

The proposed method returns the 2D+T field of labels $\boldsymbol{\ell}$ that minimizes a global energy $\mathscr{E}_{\text{OR}}(\boldsymbol{\ell}, \boldsymbol{I}_{\text{OR}}, \boldsymbol{S})$.

This method is comparable to the segmentation step of the method presented in the previous chapter, even though the data and the energy terms involved are different.

This global energy is defined as the sum of a data term, a spatial regularization term, and a temporal regularization term:

$$\mathscr{E}_{\mathrm{OR}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}) = U_{\mathrm{Data}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}, \boldsymbol{S}) + U_{\mathrm{regS}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}, \boldsymbol{S}) + U_{\mathrm{Flux}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}, \boldsymbol{S}) + U_{\mathrm{regT}}(\boldsymbol{\ell}). \tag{6.25}$$

6.4.2.1 Data term

The data term for the whole stack of images is the sum of the data terms for every image in the stack. The data term for an image depends on its sensor.

SAR images For the log-transformed SAR images ($\boldsymbol{S}(t) = 1$), a Fisher-Tippett Mixture Model (FTMM), similar to its counterpart of section 6.1.3.5 is used. The likelihood presented equation 6.12 for a log-transformed intensity $\boldsymbol{y}(k,t)$ and a mixture of Fisher-Tippett $\mathcal{C}_W$ or $\mathcal{C}_L$ is recalled below:

$$\mathcal{L}_W(\boldsymbol{y}(k,t)) = \sum_{\mathcal{K} \in \mathcal{C}_W} \boldsymbol{\pi}(\mathcal{K}) \frac{L^L}{\Gamma(L)} e^{L(\boldsymbol{y}(k,t) - x_{\mathcal{K}})} \exp(-L e^{\boldsymbol{y}(k,t) - x_{\mathcal{K}}}) \tag{6.26}$$

$$\mathcal{L}_L(\boldsymbol{y}(k,t)) = \sum_{\mathcal{K} \in \mathcal{C}_L} \boldsymbol{\pi}(\mathcal{K}) \frac{L^L}{\Gamma(L)} e^{L(\boldsymbol{y}(k,t) - x_{\mathcal{K}})} \exp(-L e^{\boldsymbol{y}(k,t) - x_{\mathcal{K}}}) \tag{6.27}$$

$$\tag{6.28}$$

Considering one FTMM $\mathcal{C}_W$ for water and one FTMM $\mathcal{C}_W$ for land, the data term for a log-transformed SAR image ($\boldsymbol{I}_{\mathrm{OR}}(\cdot,t)$, with $\boldsymbol{S}(t) = 1$) is:

$$\begin{aligned} U_{\mathrm{Data}}(\boldsymbol{I}_{\mathrm{OR}}(\cdot,\cdot,t), \boldsymbol{\ell}(\cdot,t), \boldsymbol{S}(t) = 1) = -\sum_k & \boldsymbol{\ell}(k,t) \cdot \log(\mathcal{L}_W(\boldsymbol{I}_{\mathrm{OR}}(k,t)) \\ & + (1 - \boldsymbol{\ell}(k,t)) \cdot \log(\mathcal{L}_L(\boldsymbol{I}_{\mathrm{OR}}(k,t)) \end{aligned} \tag{6.29}$$

Here, the two FTMMs (one for water and one for land) are known *a priori*, for example by running the method presented in the previous chapter on various kinds of images.

Optical images For optical images ($\boldsymbol{S}(t) = 0$), the image is either a Sentinel-2 spectral channel or a composite index. The water class and the land class are associated with a mean value (μ_W and μ_L). The data term for one image is then given by the following equation:

$$\begin{aligned} U_{\mathrm{Data}}(\boldsymbol{I}_{\mathrm{OR}}(\cdot,\cdot,t), \boldsymbol{\ell}(\cdot,\cdot,t), \boldsymbol{S}(t) = 0) = \omega \sum_k & \boldsymbol{\ell}(k,t) \cdot (\boldsymbol{I}_{\mathrm{OR}}(k,t) - \mu_W)^2 \\ & + (1 - \boldsymbol{\ell}(k,t)) \cdot (\boldsymbol{I}_{\mathrm{OR}}(k,t) - \mu_L)^2, \end{aligned} \tag{6.30}$$

where ω is a tuning hyper-parameter.

Global data term The global data term for the time series $\boldsymbol{I}_{\mathrm{OR}}$ is the sum for every date t of the corresponding data term.

$$U_{\mathrm{Data}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}, \boldsymbol{S}) = \sum_{t=1}^{T} U_{\mathrm{Data}}(\boldsymbol{I}_{\mathrm{OR}}(\cdot,t), \boldsymbol{\ell}(\cdot,t), \boldsymbol{S}(t)) \tag{6.31}$$

6.4.2.2 Regularization

There are two regularization terms: one spatial regularization term between one pixel and its neighbors at the same date, and one temporal regularization term between one pixel (k,t) and the pixels at the same position in the preceding and following images $(k,t-1)$ and $(k,t+1)$.

Spatial regularization The spatial regularization depends on the kind of image: for optical images, no regularization is used while for SAR images, a CRF-based regularization identical to its counterparts in the beginning of the chapter is used. This regularization term is presented in equation 6.22 and recalled below:

$$U_{\text{regS}}(\boldsymbol{\ell}(\cdot,t),\boldsymbol{I}_{\text{OR}}(\cdot,t)) = \beta \sum_{(k',t)\sim(k,t)} w_{\text{sym}}((k,t),(k',t)) \cdot |\boldsymbol{\ell}(k,t) - \boldsymbol{\ell}(k',t)| \tag{6.32}$$

with the same notation and the same expression for w_{sym} as in section 6.1.3.5

The global spatial regularization term is then given by:

$$U_{\text{regS}}(\boldsymbol{\ell},\boldsymbol{I}_{\text{OR}},\boldsymbol{S}) = \beta \sum_{t=1}^{T} \boldsymbol{S}(t) U_{\text{regS}}(\boldsymbol{\ell}(\cdot,t),\boldsymbol{I}_{\text{OR}}(\cdot,t)) \tag{6.33}$$

Flux term We add a spatial flux term $U_{\text{flux}}(\boldsymbol{\ell}(\cdot,t),\boldsymbol{I}_{\text{OR}}(\cdot,t)$ identical to the flux term presented in section 6.1.3.5 for every SAR image of the series.

The global flux term boils down to:

$$U_{\text{flux}}(\boldsymbol{\ell},\boldsymbol{I}_{\text{OR}},\boldsymbol{S}) = \eta \sum_{t=1}^{T} \boldsymbol{S}(t) \sum_{k} \boldsymbol{\ell}(k,t) \text{LoG}_{\boldsymbol{I}_{\text{OR}},\sigma_L}(k,t). \tag{6.34}$$

6.4.2.3 Temporal regularization

The temporal regularization term penalizes any change in the classification of every pixel between the dates t and $t+1$. Unlike its counterparts in the previous chapters, the regularization does not depend on the data, and is the same between two consecutive images, regardless of the sensor. The cost for two pixels at the same spatial location in two consecutive images being classified differently is β_T.

The temporal regularization is given by:

$$U_{\text{regT}}(\boldsymbol{I}_{\text{OR}}(\cdot,t),\boldsymbol{I}_{\text{OR}}(\cdot,t+1)) = \beta_T \sum_{t=1}^{T-1} \sum_{k} |\boldsymbol{\ell}(k,t) - \boldsymbol{\ell}(k,t+1)| \tag{6.35}$$

6.4.2.4 Global energy

The global energy $\mathscr{E}_{\text{OR}}(\boldsymbol{\ell},\boldsymbol{I}_{\text{OR}})$ presented below is mapped to the cost of a cut of the graph. In this way, the optimal $\boldsymbol{\ell}$, which corresponds to the partition of the pixels nodes in a minimal cut can be computed using a graph-cut method.

$$\mathscr{E}_{\text{OR}}(\boldsymbol{\ell},\boldsymbol{I}_{\text{OR}}) = U_{\text{Data}}(\boldsymbol{\ell},\boldsymbol{I}_{\text{OR}},\boldsymbol{S}) + U_{\text{regS}}(\boldsymbol{\ell},\boldsymbol{I}_{\text{OR}},\boldsymbol{S}) + U_{\text{regT}}(\boldsymbol{\ell}) + U_{\text{flux}}(\boldsymbol{\ell},\boldsymbol{y}) \tag{6.36}$$

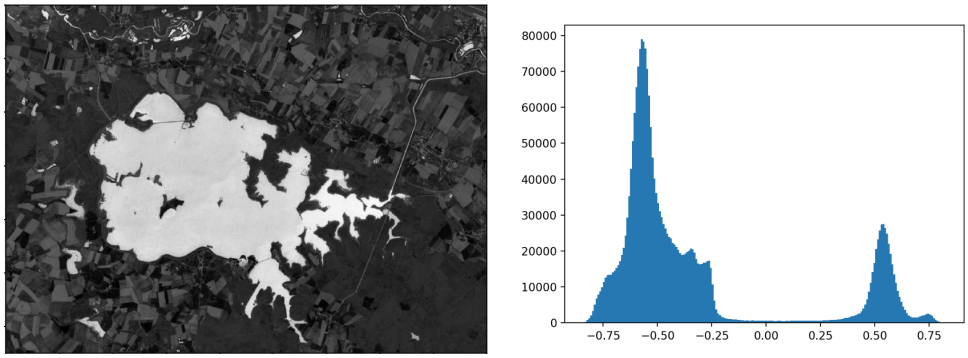
6.4.3 Experiments

6.4.3.1 Data

We ran our method on a time series that combines two kinds of images:

1. Sentinel-2 NDWI (McFeeters, 1996), with a spatial resolution of 10m.
2. Sentinel-1 VV-VH combination (see appendix A for more details).

Sentinel-1 images have been registered on the Sentinel-2 grid using S1Tiling[2]. For this reason, the geometry and the dynamic of these image is different from the SAR images used in the previous sections. The optical images have been prepared by manually removing all cloudy images from the stack and the NDWI index .**Figure** 6.16 has been computed from the green and NIR channels.



(a) NDWI image **(b)** NDWI histogram

Figure 6.16 – NDWI index for the optical Sentinel-2 image on 2018-01-02 and its histogram

6.4.3.2 Results

Figure 6.17 presents the classification results for the Sentinel-1 image on 2018-02-25 in which water detection was difficult because of a high water reflectivity caused by Bragg resonance. Thanks to the information of the optical images and to the regularization, the segmentation result is more accurate.

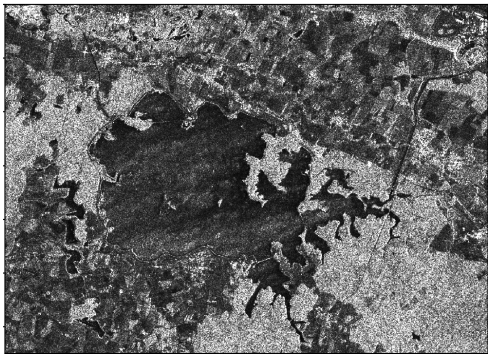
Note that for this example, the GRD SAR images have been calibrated and orthorectified, while the previous experiments have been run on SAR GRD images without these preprocessing steps.

Some results are presented in Table 6.6 that compares the results obtained with the multi-temporal SAR and optical method presented above (Opt+SAR) with the same method using a stack of SAR images only (SAR) and the same method using only one SAR image (SAR 1T).

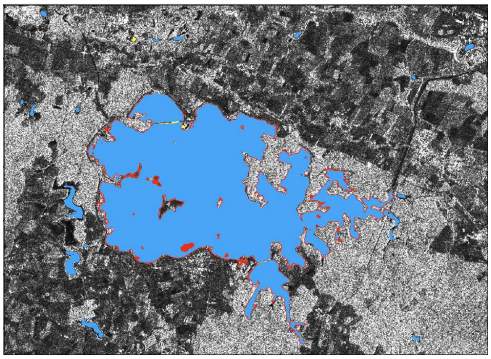
[2] https://github.com/CNES/S1Tiling

Table 6.6 – Summary of the metrics for each result

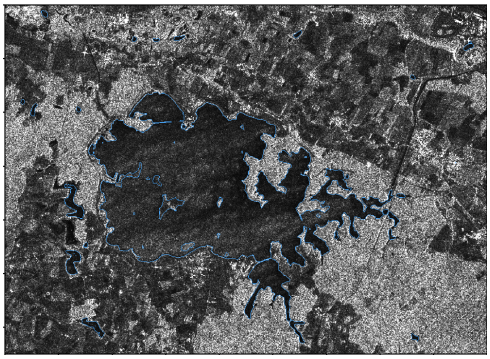
Number	Name	Channel	Pr (%)	Rec (%)	FPR (%)	F-Score (%)	ER (%)	MCC (%)
1	Der-2018-02-25	Opt + SAR	99.81	93.34	0.04	96.47	6.84	95.78
		SAR	99.94	81.92	0.01	90.04	18.13	88.74
		SAR 1T	94.44	0.00	0.00	0.01	100.00	0.56



(a) Calibrated, orthorectified GRD image on 2018-02-25



(b) Resulting classification error (stack of SAR and optical images)



(c) Detected water contours

Figure 6.17 – Mask and result for image 1: Der-2018-02-25. On the classification error image, the true positives are displayed in blue, the false positives in yellow, and the false negatives in red. True negatives are displayed as the actual SAR image amplitudes.

6.4.4 Conclusion

We proposed an extension of the segmentation step of our 2D+T GrabCut method (presented in section 6.3) to combined stacks of SAR and optical images. The terms of the model have been adapted to the characteristics of each kind of image. Our preliminary experiments delivered promising results, which opens the way for a joint SAR and optical multitemporal GrabCut method.

In combination with the previous sections of this chapter, it corresponds to contribution 2 presented in section 1.2.

Summary: Joint multitemporal segmentation for optical and SAR images

We propose a multitemporal water detection method for time series that combines SAR images (VVVH combination, as presented in A) with optical images (NDVI combination). The method is based on the minimization of a global energy $\mathscr{E}_{\mathrm{OR}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}})$ that combines multiple terms:

1. A data term $U_{\mathrm{Data}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}, \boldsymbol{S})$
2. A spatial regularization term $U_{\mathrm{Sreg}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}, \boldsymbol{S})$
3. A flux term $U_{\mathrm{Flux}}(\boldsymbol{\ell}, \boldsymbol{I}_{\mathrm{OR}}, \boldsymbol{S})$
4. A temporal regularization term $U_{\mathrm{Treg}}(\boldsymbol{\ell})$

The expression for the three first terms depends on the kind of image at the considered date. For SAR images, an expression similar to the one in the previous sections is used. For optical images, no regularization nor flux term is needed and a quadratic data term is used.

Chapter 7

Denoising of the temporal geometric mean

This chapter is based on work that has already been published in an IEEE Geoscience and Remote Sensing Letters article:

- N. Gasnier, L. Denis and F. Tupin, "On the Use and Denoising of the Temporal Geometric Mean for SAR Time Series," in IEEE Geoscience and Remote Sensing Letters, 2021, doi: 10.1109/LGRS.2021.3051936 (Gasnier et al., 2021d)

Some figures and substantial parts of the text of this chapter have been taken from this article.

7.1 Introduction

The previous chapter presented an approach utilizing multitemporal information to improve the water segmentation by jointly segmenting all the images of the stack. This chapter introduces an other way to take advantage of the temporal redundancy of information, by denoising the temporal geometric mean and using it for further processing. This denoised geometric mean can be used within a denoising-by-ratio framework or for other applications such as change detection.

As mentioned in section 3.1, SAR image denoising can be the first step for several water detection approaches. In particular, approaches based on multitemporal regularization such as RABASAR ((Zhao et al., 2018)) achieve an interesting denoising performance by exploiting the temporal regularity in the data. In this chapter, we propose an extension of this multi-temporal filtering based on the temporal geometric mean instead of the arithmetic mean.

After introducing RABASAR's multitemporal denoising in section 7.1.1, we compare the properties of the arithmetic and the geometric mean in section 7.2. Then, we propose an adaptation of the MuLoG method (Deledalle et al., 2017a) for the denoising of the temporal geometric mean in section 7.3. Finally, in section 7.4, we present some experiments on the use of the temporal geometric mean in the RABASAR framework along with its use for change detection.

7.1.1 RABASAR multitemporal denoising

The RABASAR method (Zhao et al., 2019) summarized in **Figure** 7.1 relies on a "super-image" with very little residual speckle noise and consists in five steps:

1. Temporal multilooking: the temporal mean is computed from a time series of T images. The original approach (Zhao et al., 2018) uses the temporal arithmetic mean. The multilooked image has a reduced level of speckle compared to the original individual images, but the speckle is still visible unless the number of images is very high (T>100) and the scene very stable over time.

2. The temporal multilooked image is denoised using a suitable denoiser. For example, (Zhao et al., 2018) uses the variational framework MuLoG (Deledalle et al., 2017a), presented in section 4.2.2, with a BM3D (Block Matching and 3D filtering) (Dabov et al., 2007) denoiser and a data term corresponding to the arithmetic mean of a stack of T images with fully-developed speckle. The denoised multilooked image is called the **super-image**.

3. The **ratio image** is computed by dividing pixel-wise the image to denoise by the super-image.

4. The ratio image is denoised using the MuLoG method with a data term that takes into account the speckle distribution of the ratio image

5. The denoised image is obtained by multiplying the denoised ratio image by the super-image.

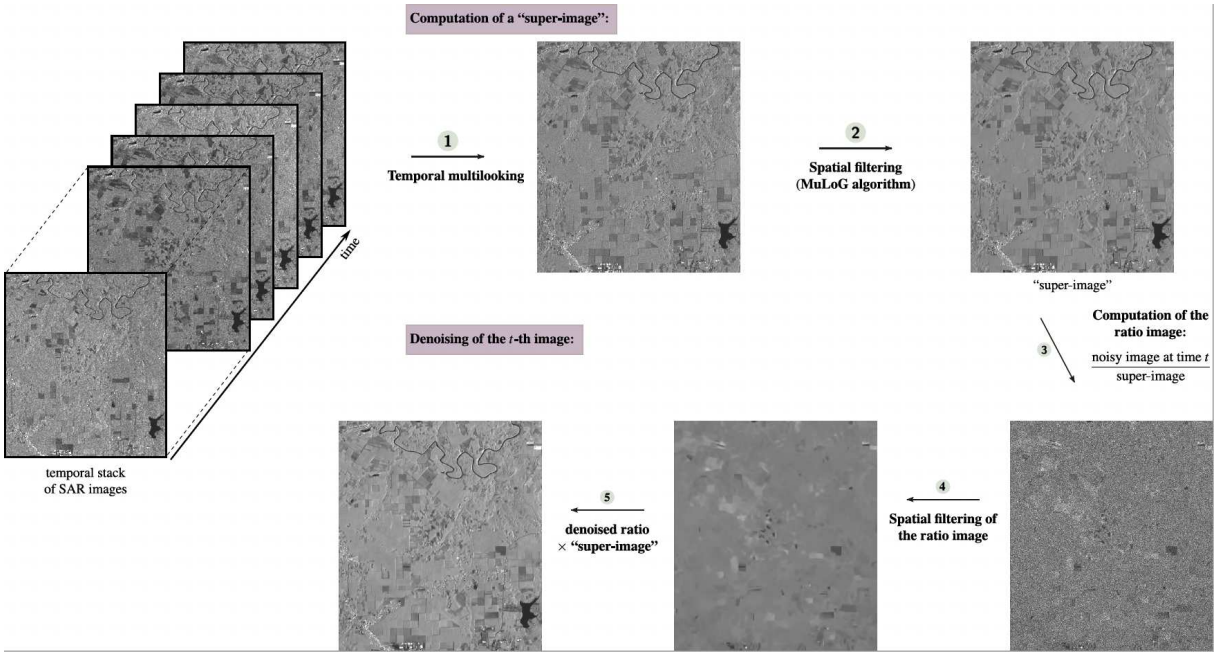


Figure 7.1 – Schematic summary of RABASAR method presented in (Zhao et al., 2018).

This denoising approach gives good results both in terms of speckle suppression, in terms of preservation of small details and generally also features that are specific to that date (even if they are not present in all or most of the original images in the time series). However, it comes with some limitations:

1. The BM3D denoiser that is used does not handle spatially correlated noise. A preliminary decorrelation is necessary. In our experiments, we simply used undersampled images.

2. The denoising assumes a fully developed speckle. For strong scatterers or Bragg phenomenon (see 2.1.3.4 and 2.1.3.2), the denoising result might have a slight bias.

3. When using the temporal arithmetic mean in the presence of transient strong scatterers (such as boats on a river), the super-image can be strongly affected, with very bright spots, which result in "ghost structures" in the final result (see 7.2).

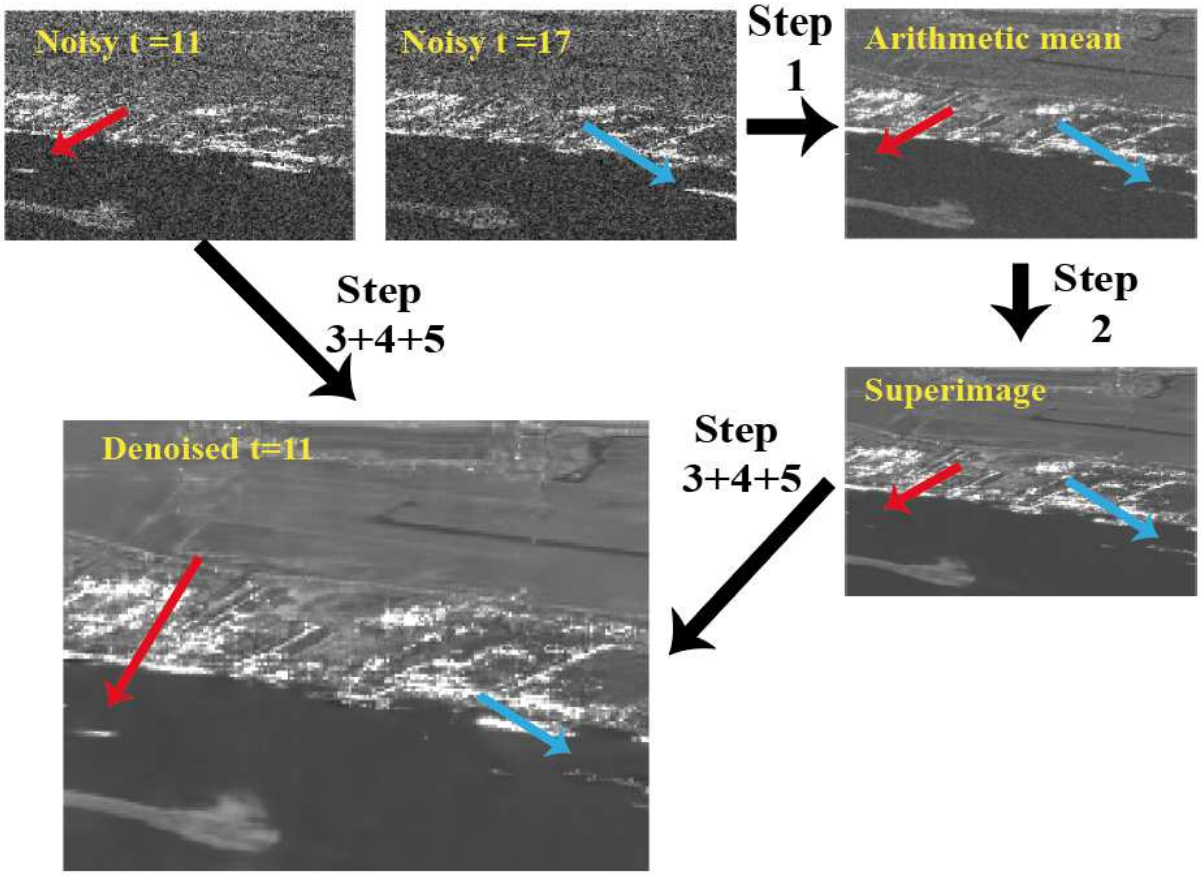


Figure 7.2 – Illustration of the issue with strong transients scatterers when using a denoising by ratio strategy with RABASAR. On the image, one boat is present at $t = 11$ (red arrow) and an other boat is present at $t = 17$ (blue arrow). The boats are not present in any other image in the 20 images stack, but can clearly be seen on the arithmetic mean image and, after denoising, in the super-image. As a consequence, a structure corresponding to the boat present at $t = 17$ appears in the denoised image $t = 11$ (blue arrow) even though it is not in the original image. We call this phenomenon a "ghost structure".

This latter issue with transient strong scatterers can be addressed in two ways:

1. By using a super-image that is specific to the denoised image and that excludes from the computation of the temporal mean in each pixel the dates of the temporal stack that are

too different from the image to denoise. This approach is used in (Zhao et al., 2018) but requires the computation of a super-image for every image that is denoised.

2. By using a super-image that is robust to these strong transients scatterers.

For the second strategy, besides approaches detecting and removing these strong scatterers, a good candidate for such a super-image is the denoised geometric mean. Indeed, as we will demonstrate in the next section, the geometric mean is much more robust to transient strong scatterers, along with other advantages.

Summary: RABASAR multitemporal denoising

The RABASAR multitemporal denoising-by-ratio framework gives good denoising results by exploiting the temporal regularity of the images through the use of a super-image. However, the arithmetic mean proposed in the original method for the computation of the super-image might not be the best option in case of temporal variations of the reflectivity. In these cases, the temporal geometric mean could be more appropriate

7.2 Statistics of the temporal geometric mean of SAR intensities

In this section, we study the statistics of the geometric mean of SAR images to motivate its use in the processing of SAR time series. The temporal arithmetic mean (temporal multi-looking) has long been used for this purpose (Nieuwenhuis and Schotten, 1992). The use of other kinds of averaging procedures such as Hölder or Lehmer means has been studied in (Quin et al., 2014). Among these means, the geometric mean stands out for having particularly interesting properties. In particular, the geometric mean can be combined with the arithmetic mean within a likelihood ratio test to obtain a simple yet effective change detector (Lombardo and Oliver, 2001). More broadly, the multiplicative approaches demonstrated their usefulness in the processing of long SAR time series, such as in (Atto et al., 2016).

7.2.1 Statistics of a SAR image

As presented in section 2.1.3.1, the logarithmically transformed intensity image with fully developed speckle follows a Fisher-Tippett distribution (Hua Xie et al., 2002) defined by the following expression, where $y = \log(I)$ and $x = \log(R)$:

$$\mathrm{p}(y|x) = \frac{L^L}{\Gamma(L)} e^{L(y-x)} \exp(-Le^{y-x}). \tag{7.1}$$

The expectation and the variance of y are:

$$\mathbb{E}[y] = x - \log(L) + \psi(L) \tag{7.2}$$

$$\mathrm{Var}[y] = \psi'(L) \tag{7.3}$$

$$\tag{7.4}$$

where $\psi(.)$ is the digamma function and $\psi'(.)$ is the trigamma function, see for example (Olver et al., 2010).

7.2.2 Statistics of the geometric mean

The geometric mean in the pixel (i,j) of T intensity values $\boldsymbol{I}(i,j,t)$ is defined by:

$$\boldsymbol{I_G}(i,j) = \sqrt[T]{\prod_{t=1}^{T} \boldsymbol{I}(i,j,t)} = \exp\left(\frac{1}{T}\sum_{t=1}^{T} \log \boldsymbol{I}(i,j,t)\right), \tag{7.5}$$

it corresponds to computing the exponential of an arithmetic mean of the log-transformed intensities.

If the speckle is completely decorrelated from one image to another, and if the reflectivity remains constant ($\forall t, \boldsymbol{R}(i,j,t) = \boldsymbol{R}(i,j)$), it is possible to express the probability density function of the geometric mean $\boldsymbol{I_G}(i,j)$ using Meijer functions (Nicolas and Tupin, 2016). Using the notations standardized in (Olver et al., 2010) for Meijer functions, the pdf is given by:

$$\mathrm{p}(I_G|R) = T\left(\frac{L}{R\cdot\Gamma(L)}\right)^T I_G^{T-1} \times \bar{G}_{0,T}^{T,0}\left(\frac{L^T I_G^T}{R^T}; \underbrace{L-1,\ldots,L-1}_{\mathrm{T}}, \overset{\cdot,\cdot}{\cdot}\right). \tag{7.6}$$

This expression can hardly be used for numerical computations as the evaluation of Meijer functions with numerous parameters is very slow. We have therefore proposed an alternative to evaluate numerically the pdf.

This evaluation can be done for $\boldsymbol{y_G}(i,j) = \log(\boldsymbol{I_G}(i,j))$ as it is the sum for all dates t of the $\boldsymbol{y}(i,j,t) = \log(\boldsymbol{I}(i,j,t))$ divided by T: its pdf is thus given by the convolution of all the pdf of the $\boldsymbol{y}(i,j,t)/T$. Under a constant reflectivity hypothesis, the pdf is the same for all dates t: $\mathrm{p}(\boldsymbol{y}(i,j,t)|\boldsymbol{x}(i,j,t)) = \mathrm{p}(\boldsymbol{y}(i,j)|\boldsymbol{x}(i,j))$, with $\boldsymbol{x}(i,j)) = \log(\boldsymbol{R}(i,j))$, as presented in section 2.1.3.1.

$$\mathrm{p}(y_G|x) = \underbrace{\mathrm{p}(y\cdot T|x) * \mathrm{p}(y\cdot T|x) * \cdots * \mathrm{p}(y\cdot T|x)}_{T} \tag{7.7}$$

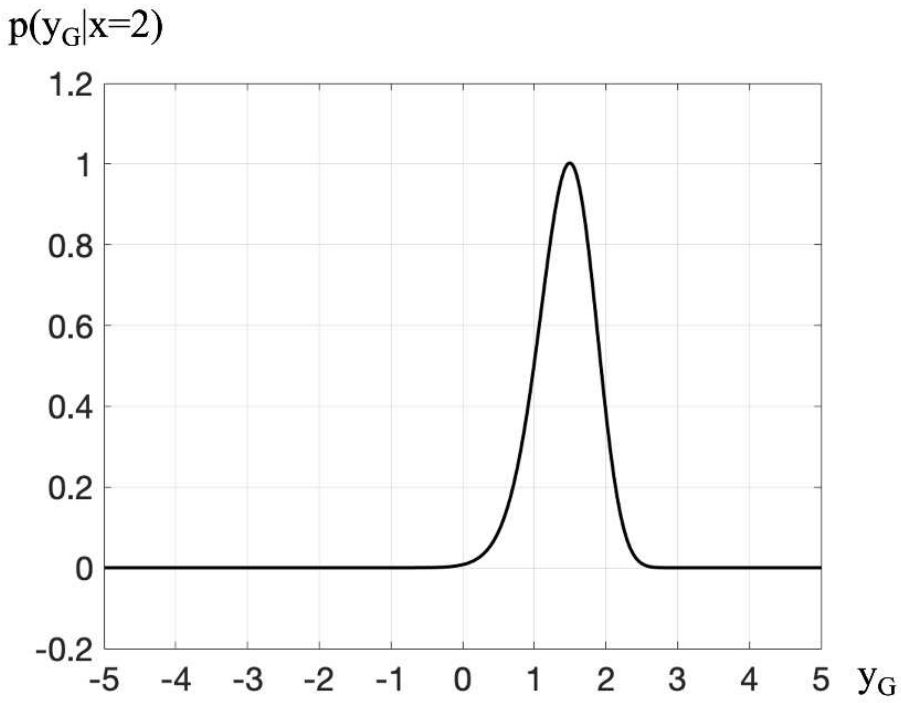


Figure 7.3 – Numerical computation of the distribution of the logarithm of the temporal geometric mean for a log-reflectivity of $x = 2$ and $T = 10$ single-look images

The geometric mean is affected by a bias that can be computed and compensated for (Quin, 2014):

$$\mathbb{E}[\boldsymbol{I_G}] = \frac{R}{L}\left(\frac{\Gamma(L)}{\Gamma(\frac{L.T+1}{T})}\right)^{-T} \tag{7.8}$$

In the following, $\widetilde{\boldsymbol{I}}_{\boldsymbol{G}}$ is the debiased geometric mean estimator obtained by dividing $\boldsymbol{I_G}$ by the bias $B_G(T, L) = \frac{1}{L}\left(\frac{\Gamma(L)}{\Gamma(\frac{L.T+1}{T})}\right)^{-T}$, L being the original number of looks of each date (here $L = 1$ for single-look images).

7.2.3 Comparison between geometric and arithmetic means

The comparison between the geometric mean and the arithmetic mean estimators of the reflectivity performed in this section shows that while the arithmetic mean estimator is preferable when there is no change in the underlying scene, the geometric mean estimator behaves better as soon as there are significant changes of the reflectivity in at least one image of the time series.

Four situations are considered:

1. No change
2. Fluctuations around a mean value (i.e., temporal texture)
3. Transient temporal changes
4. Permanent temporal changes

7.2.3.1 Situation without changes in the reflectivity, i.e., $\boldsymbol{R}$ is constant over time

Arithmetic mean When the reflectivity remains constant ($\forall t, \boldsymbol{R}(i,j,t) = \boldsymbol{R}(i,j)$), the arithmetic mean estimator has the smallest standard deviation and corresponds to the maximum likelihood estimator. The arithmetic mean of T intensities, assuming a constant reflectivity and no speckle correlation, follows a gamma distribution where the number of looks L is multiplied by T (L is thus replaced by LT in the expression of its pdf (equation A.2 for the intensity images and equation 2.7 for their log). The standard deviation σ_A of the arithmetic mean of T intensity values is:

$$\sigma_{\mathrm{A}} = \frac{R}{\sqrt{TL}}\,. \tag{7.9}$$

Geometric mean The standard deviation σ_{G} of the geometric mean estimator $\widetilde{\boldsymbol{I}_{\boldsymbol{G}}}$ can be computed through the first and the second moments of the distribution (Quin et al., 2014):

$$\sigma_{\mathrm{G}} = RL\left(\frac{\Gamma(L)}{\Gamma(\frac{L.T+1}{T})}\right)^{T}\left[\frac{\Gamma(\frac{TL+2}{T})^{T}}{\Gamma(L)^{T}} - \frac{\Gamma(\frac{TL+1}{T})^{2T}}{\Gamma(L)^{2T}}\right]^{1/2}. \tag{7.10}$$

Both estimators are consistent (σ_{A} and σ_{G} tend to zero for large values of T). The arithmetic mean is a more efficient estimator than the geometric mean. When T is large, we obtain:

$$\lim_{T\to\infty}\frac{\sigma_{\mathrm{G}}}{\sigma_{\mathrm{A}}} = \sqrt{L\cdot\Psi(1,L)}\,. \tag{7.11}$$

This ratio tends to 1 when L is large, as shown in **Figure** 7.4. It is maximal for $L = 1$ where it is equal to $\pi/\sqrt{6} \approx 1.28$.

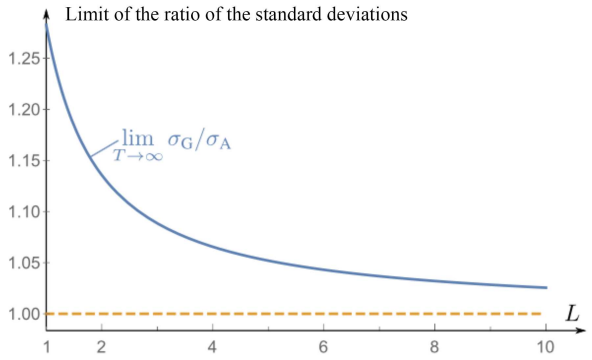


Figure 7.4 – Ratio $\sigma_{\mathrm{G}}/\sigma_{\mathrm{A}}$, for large values of T, as a function of the number of looks L.

7.2.3.2 Situation with fluctuations of the reflectivity (temporal texture)

In this paragraph we consider the case of intra-class fluctuations, inducing a temporal texture.

Although it is a well-known result that the geometric mean is more robust to strong outliers than the arithmetic mean, this paragraph shows that it is also less affected by temporal texture.

Texture models (Oliver, 1993) have long been used to describe fluctuating reflectivities in speckle. To study the impact of these fluctuations, we considered the following situation: a Gaussian distribution for the temporal evolution of the soil moisture is assumed. As there is a linear relationship between the log of the reflectivity and the moisture for a given soil (Quesney et al., 2000; Bousbih et al., 2017), we assumed a Gaussian temporal distribution of

the soil moisture and thus modeled the homogeneous reflectivities with a log-normal temporal distribution with parameters μ_R and σ_R:

$$\mathrm{p}(\boldsymbol{R}(i,j,t)|\mu_R,\sigma_R) = \frac{1}{\boldsymbol{R}(i,j,t)\sigma_R\sqrt{2\pi}} \exp\left(-\frac{(\ln \boldsymbol{R}(i,j,t) - \mu_R)^2}{2\sigma_R^2}\right). \tag{7.12}$$

Under this assumption, we can compare in numerical simulations the spatial variance of the arithmetic and geometric means for various levels of variance σ_R of the reflectivity distribution and different levels of temporal correlation. Fig. 7.5 shows that the coefficient of variation (i.e. the ratio between the standard deviation and the mean) of the geometric mean remains constant when the temporal fluctuations σ_R increase. In contrast, the coefficient of variation of the arithmetic mean rises with the temporal fluctuations. This behavior is confirmed for all levels of temporal correlation of the speckle considered. As soon as the temporal fluctuations are non-negligible (e.g., a standard deviation σ_R that exceeds 0.37 in the conditions of our numerical experiments: absence of temporal correlations and stack of T=12 dates), the geometric mean offers a better signal-to-noise ratio (i.e., a smaller coefficient of variation) compared to the arithmetic mean. This behavior can be easily explained: the arithmetic mean is heavily influenced by the large variance of the intensities corresponding to the largest radiometries.
On the contrary, the spatial variance of the geometric mean is proportional to its mean value. This way, the coefficient of variation of the geometric mean does not depend on the value of σ_R, hence the constant value for the green curve in **Figure** 7.5.

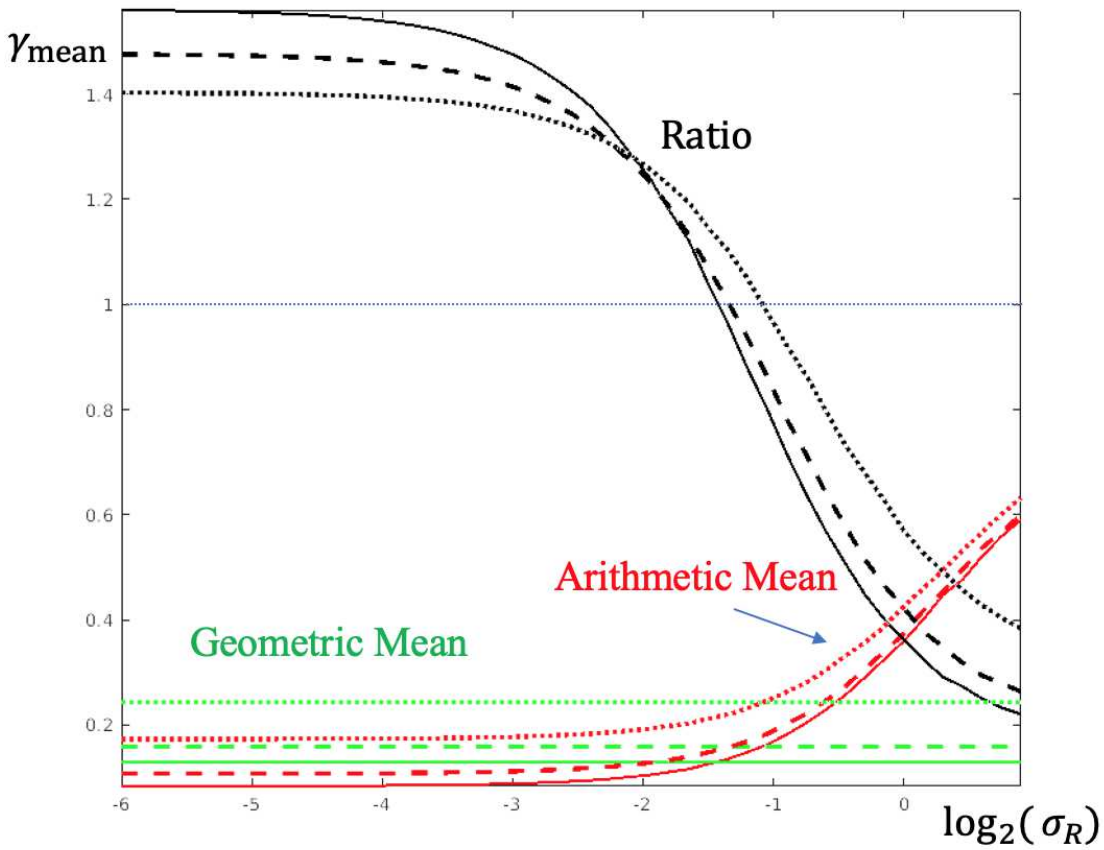


Figure 7.5 – Coefficient of variation of the arithmetic mean estimator γ_{A} (red) and debiased geometric mean estimator γ_{G} (green) for $\sigma_R \in [2^{-6}, 1]$ and T=12. The ratio $\gamma_{\mathrm{G}}/\gamma_{\mathrm{A}}$ is shown in black. Dotted and dashed lines correspond to simulations with temporally correlated speckle (correlations between successive images are 0.62 and 0.37, respectively).

The difference is also visible in real images. In **Figure** 7.6, both the arithmetic mean image and the debiased geometric mean image are computed for a time series of Sentinel-1 SAR images over an area of rice fields, where the underlying reflectivity changes over time. The remaining fluctuations of the speckle noise are stronger in the arithmetic mean image (Fig.7.6a) than in the geometric mean image (Fig.7.6b).

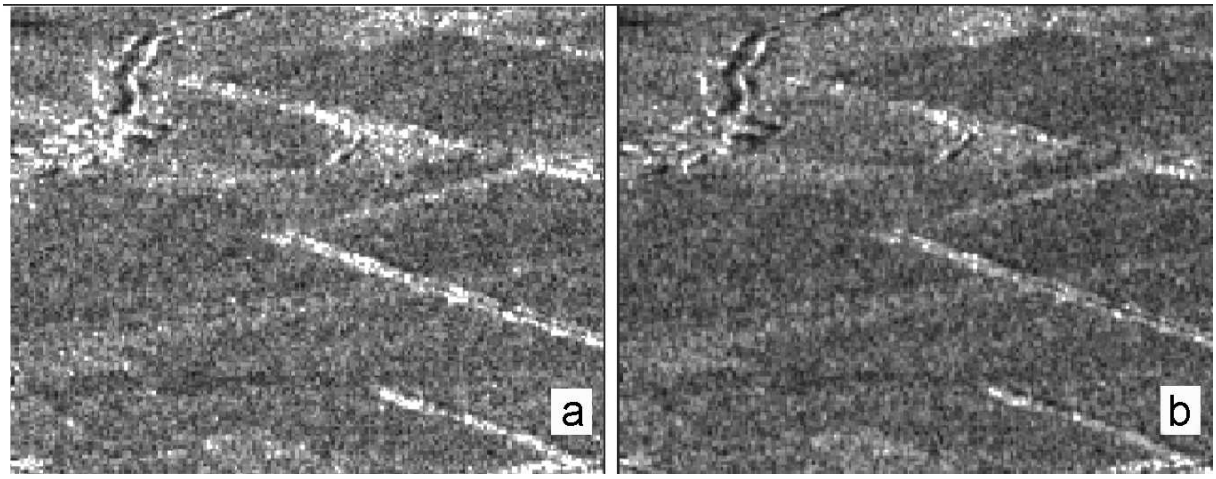


Figure 7.6 – Comparison of the arithmetic mean image (left) and geometric mean image (right) on a time series of Sentinel-1 SLC images. The fluctuations caused by the remaining noise are stronger in the arithmetic mean image.

Beyond our example on soil moisture, other phenomenons can cause temporal fluctuations of the reflectivities of both lands and water surfaces, including for SWOT images. Water surface reflectivity depends on wind conditions and the reflectivity of fields is subject to seasonal changes

In the case of temporal changes, the mean reflectivity obtained with the arithmetic or geometric means does not coincide with the actual reflectivities of the time series, but still provides useful geometrical information (e.g., border of fields, forests, roads, lakes or rivers).

7.2.3.3 Situation with transient temporal changes

Bright transient changes of the reflectivity are often seen in SAR time series and can be caused by vehicles, boats, or by temporary constructions. For instance, when there is a boat visible at one date, it produces strong echoes as illustrated in **Figure** 7.2. If we model the reflectivity change by a multiplication by a factor of $K >> 1$ at this date, the geometric and arithmetic means are modified as follows:

- the geometric mean estimator is multiplied by $K^{\frac{1}{T}}$,
- the arithmetic mean estimator is multiplied by $1 + \frac{K-1}{T}$.

When $T > 1$, $\lim_{K\to\infty} K^{\frac{1}{T}}/(1 + \frac{K-1}{T}) = 0$, which indicates that the geometric mean is more robust to the presence of strong scatterers at a single date: the impact of these scatterers in the geometric mean image is much smaller. In contrast, the arithmetic mean is less sensitive to the dark counterpart of these transient changes

7.2.3.4 Situation with permanent changes

If the change is present in a large number of images, neither the arithmetic mean nor the geometric mean are good estimators of the scene. Indeed, in this situation where two classes are successively present in the time series, a single estimate cannot capture both classes. In this case, the

geometric mean will bias towards the dark class, while the arithmetic mean will bias toward the bright class.

In conclusion, the geometric mean has many advantages compared to the arithmetic mean, being more adapted for homogeneous classes with temporal texture and transient situations.

Summary: Statistics of the temporal geometric mean

We are not aware of any closed form for the distribution of the temporal geometric mean, but we can estimate numerically the distribution of its logarithm with T convolutions of Fisher-Tippett distributions.

While the arithmetic mean is a better estimator of the underlying reflectivity if the reflectivity is constant over time, the geometric mean is more robust to bright temporal outliers and more stable in the case of fluctuating reflectivity.

7.3 Denoising method

The temporal averaging reduces speckle fluctuations in the images obtained by the geometrical and arithmetic means. To further reduce the fluctuations, an additional denoising step is beneficial. In this section, we extend the MuLoG framework (Deledalle et al., 2017b) to denoise images obtained with the geometric mean. MuLoG has been developed for SAR images with gamma-distributed intensities, it should thus be adapted to account for Meijer-distributed variables.

A denoised image (i.e., an image of estimated reflectivities) is obtained with MuLoG by minimizing the following cost function:

$$\widehat{\boldsymbol{x}}_{\boldsymbol{G}} = \underset{\boldsymbol{x}_{\boldsymbol{G}} \in \mathbb{R}^n}{\arg\min} [-\log(\mathrm{p}(\boldsymbol{y}_{\boldsymbol{G}}|\boldsymbol{x}_{\boldsymbol{G}})) + f_{\mathrm{reg}}(\boldsymbol{x}_{\boldsymbol{G}})] \tag{7.13}$$

where $\widehat{\boldsymbol{x}}_{\boldsymbol{G}}$ is the restored image, in log domain ($\boldsymbol{x}(i,j)$ is the log of the "true reflectivity"[1] of the geometric mean at pixel (i,j)), $\boldsymbol{y}_{\mathrm{G}}(i,j)$ is the log of the geometric mean image ($\boldsymbol{y}_{\mathrm{G}}(i,j)$ corresponds to the value of $\log(\widetilde{\boldsymbol{I}}_{\boldsymbol{G}})$ at pixel (i,j)). The term $\log(\mathrm{p}(\boldsymbol{y}_{\boldsymbol{G}}|\boldsymbol{x}_{\boldsymbol{G}}))$ is the log-likelihood and the regularization function f_{reg} ensures that the estimated image $\widehat{\boldsymbol{x}}_{\boldsymbol{G}}$ has a satisfying regularity (f_{reg} can be the Total Variation (TV) or patch-based regularization like BM3D (Dabov et al., 2007) (Deledalle et al., 2017b)).

The problem (7.13) is solved by a few iterations of the ADMM (Alternating Direction Method of Multipliers) algorithm (Chan et al., 2017), i.e., by alternating a Gaussian denoising step given in equation (7.14) below and the non-linear correction defined by equation (7.16) to account for

[1] Because the temporal geometric mean does not follow a Gamma distribution, its denoised version does not really correspond to any underlying reflectivity of the noisy image.

the non-Gaussianity of speckle fluctuations in images of the geometric mean:

$$\widehat{\boldsymbol{z}} \leftarrow \underset{\boldsymbol{z}\in\mathbb{R}^n}{\arg\min} \ \frac{\beta}{2}||\boldsymbol{z} - \widehat{\boldsymbol{x}_{\boldsymbol{G}}} + \widehat{\boldsymbol{d}}||^2 + f_{\text{reg}}(\boldsymbol{z}), \tag{7.14}$$

$$\widehat{\boldsymbol{d}} \leftarrow \widehat{\boldsymbol{d}} + \widehat{\boldsymbol{z}} - \widehat{\boldsymbol{x}_{\boldsymbol{G}}}, \tag{7.15}$$

$$\widehat{\boldsymbol{x}_{\boldsymbol{G}}} \leftarrow \underset{\boldsymbol{x}_{\boldsymbol{G}}\in\mathbb{R}^n}{\arg\min} \ \frac{\beta}{2}||\widehat{\boldsymbol{z}} - \boldsymbol{x}_{\boldsymbol{G}} + \widehat{\boldsymbol{d}}||^2 - \log \mathrm{p}(\boldsymbol{y}_{\boldsymbol{G}}|\boldsymbol{x}_{\boldsymbol{G}}), \tag{7.16}$$

where $\beta_v > 0$ is a parameter that acts on the speed of convergence and $\widehat{\boldsymbol{d}_v}$ is a variable of the same size than the image that is initialized with $\widehat{\boldsymbol{d}_v}(i,j) = 0 \ \forall (i,j)$.

The minimization (7.16) can be solved with Newton's method by using the following formula for all pixels (i,j):

$$\widehat{\boldsymbol{x}_{\boldsymbol{G}}}(i,j) \leftarrow \widehat{\boldsymbol{x}_{\boldsymbol{G}}}(i,j) - \frac{\widehat{\boldsymbol{x}_{\boldsymbol{G}}}(i,j) - \widehat{\boldsymbol{z}}(i,j) - \widehat{\boldsymbol{d}_v}(i,j) + \frac{D_1(\boldsymbol{x}_{\boldsymbol{G}}(i,j),\boldsymbol{y}_{\boldsymbol{G}}(i,j))}{\beta}}{|1 + \frac{D_2(\boldsymbol{x}_{\boldsymbol{G}}(i,j),\boldsymbol{y}_{\boldsymbol{G}}(i,j))}{\beta}|} \tag{7.17}$$

with $D_1(\boldsymbol{x}_{\boldsymbol{G}}(i,j), \boldsymbol{y}_{\boldsymbol{G}}(i,j))$ and $D_2(\boldsymbol{x}_{\boldsymbol{G}}(i,j), \boldsymbol{y}_{\boldsymbol{G}}(i,j))$ the first and second derivatives of the log-likelihood:

- $D_1(\boldsymbol{x}_{\boldsymbol{G}}(i,j), \boldsymbol{y}_{\boldsymbol{G}}(i,j)) = -\partial \log \mathrm{p}(\boldsymbol{y}_{\boldsymbol{G}}(i)|\boldsymbol{y}_{\boldsymbol{G}}(i))/\partial \boldsymbol{x}_{\boldsymbol{G}}(i,j)$ and
- $D_2(\boldsymbol{x}_{\boldsymbol{G}}(i,j), \boldsymbol{y}_{\boldsymbol{G}}(i,j)) = -\partial^2 \log \mathrm{p}(\boldsymbol{y}_{\boldsymbol{G}}(i)|\boldsymbol{x}_{\boldsymbol{G}}(i))/\partial \boldsymbol{x}_{\boldsymbol{G}}(i,j)^2$.

The likelihood of the geometric mean has been defined in the intensity domain using Meijer functions in equation (7.6). In the log-domain, it can be defined as the iterated convolution product of Fisher-Tippett distributions. In the absence of closed-form expressions for the result of these convolution products, it is necessary to evaluate them numerically, as well as their derivatives. We computed the convolutions between the T Fisher-Tippett distributions as multiplications in the Fourier domain and then obtained the derivatives D_1 and D_2 by finite differences. These values, which depend only on L, T and $\boldsymbol{y}_{\boldsymbol{G}}(i,j) \cdot T - \boldsymbol{x}_{\boldsymbol{G}}(i,j)$ can be precomputed and stored in two one-dimensional tables to speed up the restoration process.

Summary: Denoising method

We proposed an adaptation of the MuLoG framework based on the use of numerically pre-tabulated values for the data fidelity term, as no closed-form is available for the likelihood.

7.4 Experiments

In **Figure** 7.7 both the arithmetic (c) and the geometric (d) temporal mean show an obvious improvement in terms of noise level compared to the individual images from the time series (a and b). Nevertheless, there is still a significant level of noise in these images. As presented in section 7.2.3.1, the noise level in water areas (stable reflectivity) is stronger with the geometric mean than with the arithmetic mean. However, in both denoised images (e) and (f), the remaining noise has been very strongly reduced.

Concerning temporary strong scatterers, some boats that are only present at one time of the time series such as the one on the left of the 11th image of the time series (a), are clearly visible in the noisy (c), and denoised arithmetic mean (e), but not in the noisy (d) nor denoised geometric mean (f).

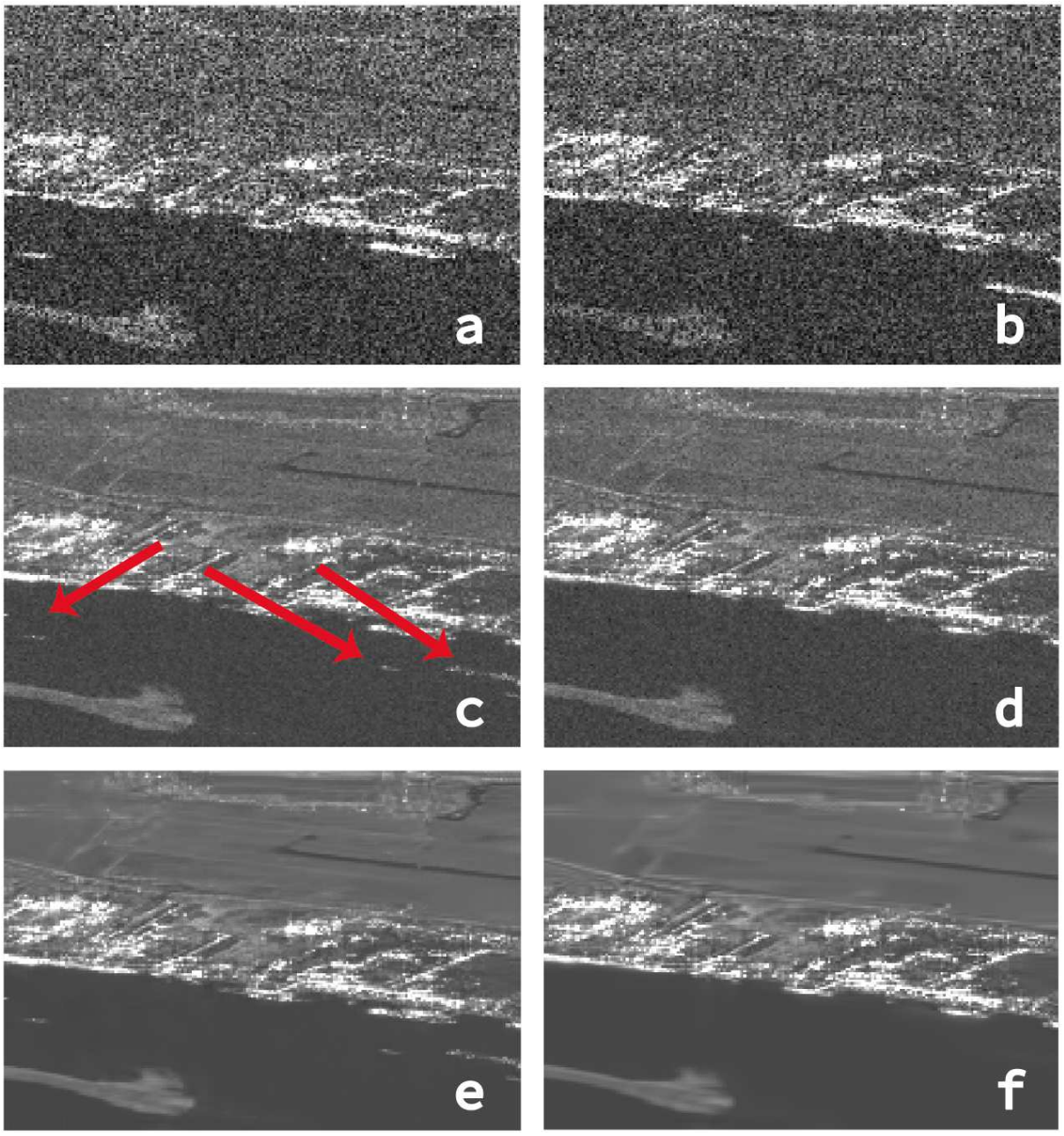


Figure 7.7 – Geometric vs. arithmetic mean on a time series of 20 Sentinel-1 images: (a) and (b) two images from the time series corresponding to dates t=11 and t=17, (c) arithmetic mean and (d) debiased geometric mean, (e) denoised arithmetic mean and (f) denoised geometric mean.

7.5 Application to change detection

Comparisons of the arithmetic and geometric temporal means can be used to perform change detection. Consider for example the MIMOSA method proposed by (Quin et al., 2014). The use of denoised mean images improves detection methods based on these comparisons. A much simpler way to detect changes is with the ratio between the arithmetic mean and the geometric mean. Because of the residual speckle fluctuations in the mean images, when the total number of dates is moderate, this ratio image is noisy which leads to false alarms and non-detections. **Figure** 7.8 illustrates, in the same Sentinel-1 SAR time series as in **Figure** 7.5, the improvement of the ratio image brought by denoising.

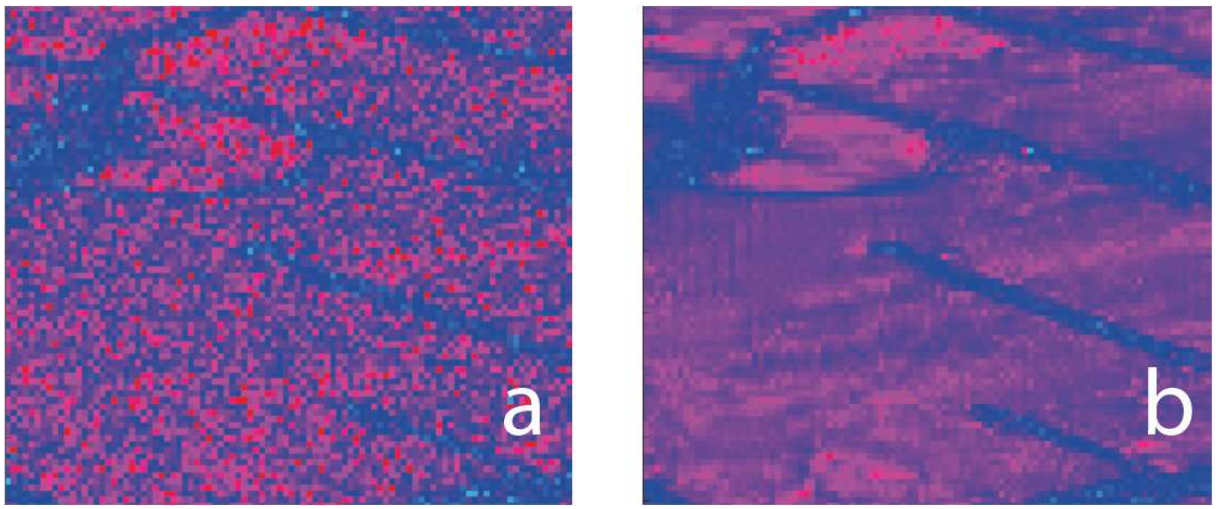


Figure 7.8 – Improved change detection with denoised arithmetic and geometrical means: (a) changes identified by the ratio arithmetic mean / geometric mean, (b) ratio of the denoised mean images. Stable areas are shown in blue, changing areas in red.

7.6 Application to ratio-based denoising of single SAR images within a time series

RABASAR (Zhao et al., 2019), introduced in section 7.1.1 is a speckle reduction method for time series. It uses the arithmetic mean to produce a so-called "super-image" and to form a ratio-image where most of the spatial variability of the reflectivity is compensated for. As discussed in Section 7.2, in the presence of an intra-class temporal texture the geometrical mean is less impacted by speckle fluctuations. It is also more robust to bright scatterers appearing only on a few dates. In these contexts, the geometrical mean leads to a better super-image and improved multi-temporal filtering results. **Figure** 7.9 shows how the restoration of an image of the time series illustrated in **Figure** 7.7 is improved when the denoised geometrical mean is used as the super-image in RABASAR: ghost structures due to transient bright scatterers (boats visible only at a few dates) are suppressed in **Figure** 7.9(b) in the areas indicated by red circles. While a possible workaround to the presence of bright targets at only a few dates could consist of creating a different super-image for each image of the stack (by selecting only the dates that are sufficiently similar, as done in the original RABASAR framework (Zhao et al., 2019)), this

latter approach involves a significant increase of the computational load (a super-image must be re-created for each date) and does not offer improvement of the signal to noise ratio in areas with a temporal texture.

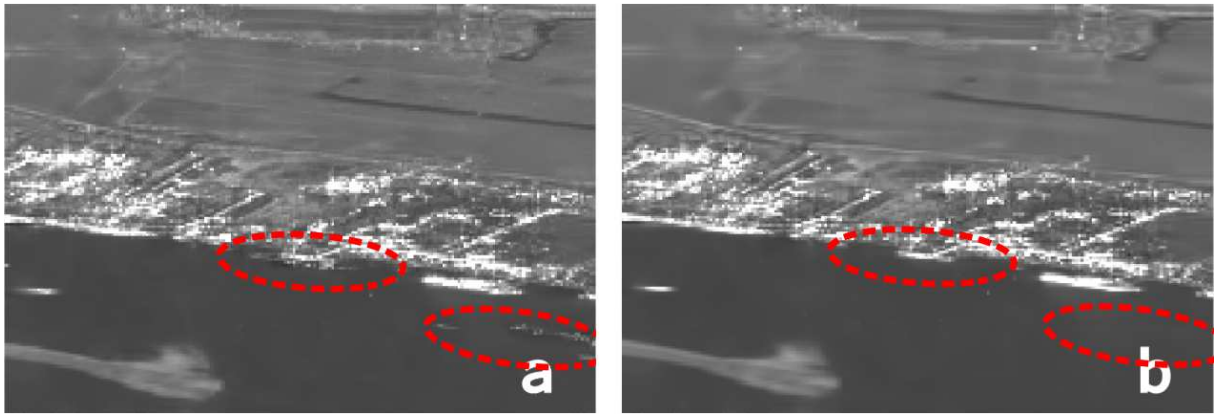


Figure 7.9 – Improved temporal filtering with a super-image obtained by denoising the geometric mean: the speckled image at the date $t = 11$ (shown in Fig.7.7(a)) is restored by RABASAR using a super-image obtained from (a) the arithmetic mean, or (b) the geometric mean. Note the reduction of artifacts with the geometric mean in the areas indicated by red circles.

Summary: Experiments and results

Our experiments show that the increased robustness of the geometric mean compared to the arithmetic mean is able to address the "ghost structure" issue of the original RABASAR framework (see 7.1.1). The denoising of the geometric mean can be useful for other applications such as change detection.

7.7 Conclusion

We have shown the benefits of using the geometric mean as a representative super-image for multi-temporal SAR data stacks and a modified approach to further reduce the speckle on this geometric-mean image. Due to the non-linear combination of speckled images in a geometrical mean, the denoising process must be carefully adapted to account for the statistical distribution of speckle in the geometrical mean image. The geometrical mean may be preferred over the arithmetic mean for several reasons: improved robustness to the occasional presence of bright scatterers (e.g., boats) and an improved signal-to-noise ratio in areas with temporally fluctuating reflectivities (e.g., vegetation). Denoised geometric images can be interesting for instance to obtain a temporal summary of a multi-temporal stack of SAR images for visualization purposes.

The ratio of denoised arithmetic and geometric images can also indicate changes occurring in the time series. The denoising step offers a notable improvement of the quality of the simple change detection map. Our method to denoise geometric mean images has also been applied to the multi-temporal filtering algorithm RABASAR and shown to effectively reduce the "ghost structures" appearing at the location of strong scatterers that were visible only at some other

dates. As our numerical approach allowed for the denoising of any noise distribution, even without any analytic expression for its distribution, it could be used as well for other kinds of noise models.

However, the limitations of the embedded BM3D denoiser limited the practical application of this method, especially, its incapacity to handle spatial correlation of the noise. In order to exploit its full potential, a better embedded denoiser will be needed. For that purpose, denoising methods based on deep learning approaches that can handle correlated noise such as SAR2DAR (Dalsasso et al., 2021) or Speckle2Void (Molini et al., 2021) are very promising. Meanwhile, we used a single-date denoising method for GRD images derived from SAR2SAR to assess the opportunity of using a denoising step for our narrow river detection method. This preliminary work has shown promising results and has been presented at the 2021 IGARSS conference (Gasnier et al., 2021a)

This chapter corresponds to the contribution 3 presented in section 1.2

Chapter 8

Conclusion and perspectives

Conclusion

The main objective of this PhD work was to provide alternative methods to detect water in difficult situations, where traditional methods that rely only on the information contained within the SAR image may not be sufficient, because of low water/land contrast, low SNR, or confusing structures. To this end, we explored three main strategies:

- Use of exogenous information, that could be taken from external databases, to guide the detection.
- Combination of multitemporal and multi-sensor information to take advantage of the temporal redundancy and the complementarity of SAR and optical images, when available.
- Denoising of SAR images as a preprocessing step, before the actual water detection.

We proposed two different guided approaches, for rivers and lakes, that are adapted to their specific shape and to the kind of guiding information that is available from the corresponding databases. The first approach is based on *a priori* centerlines that can be extracted from the SWORD database and consists of three steps, including a new linear structure detector and a new CRF model that we proposed. This first guided approach showed good results on all the images we tested and would be the most straightforward way to adapt to the SWOT processing chain if needed, as the river database is already used and therefore directly available.
The second guided approach concerns lakes and is based on the use of a bounding polygon taken from the SWOT lake database. This kind of prior information prompted us to adapt the GrabCut method to our task. We did that by replacing the Gaussian mixture models used for RGB images with Fisher-Tippett mixture models (FTMMs) for our log-transformed intensity images. This approach provided good results in most situations.

We also adapted this GrabCut method to multitemporal stacks of SAR images, using a temporal regularization term and the same FTMMs for all the dates. This multitemporal adaptation leads to some improvements in the localization of boundaries, but at the cost of poorer robustness to Bragg phenomenon, hence an ambivalent outcome. This robustness issue could be addressed in later work by using a different pair of FTMMs for each date.

For all these approaches, we found that using a flux term and, for the detection of rivers, an asymmetric regularization term significantly improved their ability to detect water. Indeed, this allows to fully exploit the fact that water is darker than the surrounding land in conventional SAR images, and brighter for SWOT. As the reflectivity distributions for land and water are generally not known *a priori*, this relative information is valuable, but was untapped by previous approaches.

Because of the wealth of information contained in optical images, we also proposed a combined detection approach for stacks of both SAR and optical images, exploiting prior statistical models for water and land distributions in both kinds of images. While our tests for this method were not as extensive as for the others, the first results look promising. As the inputs and outputs of this method correspond to those of the segmentation step of any GrabCut approach, plugging it into a GrabCut approach could be relevant for multitemporal and multi-sensor water detection guided by an *a priori* polygon.

Finally, we considered the use of a denoising step before the actual detection. Promising preliminary results using a denoiser based on a GRD adaptation of the SAR2SAR method (Dalsasso et al., 2021), based on deep learning, and our narrow river detection framework presented in chapter 5 were presented at the 2021 IGARSS conference (Gasnier et al., 2021a). Our work on the temporal geometric mean, presented in chapter 7, showed its potential as a basis for denoising-by-ratio methods that could be used for such denoising steps. However, the technical limitations of the embedded denoiser, for instance with spatial correlation, limited its practical applications.

Beyond the SWOT mission, our methods can have a major interest for the detection of water with other SAR sensors, such as Sentinel-1, that we used in several of our tests. Other applications of our work can also be considered. They are straightforward for the denoising-by-ratio methods based on the temporal geometric mean, but more specific methods could see other uses. For example, our narrow river detection framework may be useful to detect roads in SAR images or SAR coherence images.

Outlook

Once the actual SWOT data are available, the calibration/validation process of the baseline water detection method will determine if the methods we proposed are needed as a complement to detect waterbodies that would have been missed otherwise. In this case, the most straightforward methods to use are those that combine single-date SWOT images with information from SWOT databases, as it would require only limited modifications of the global SWOT processing chain. Nonetheless, some further tuning of our methods, based on the characteristics of the actual SWOT images, would still be necessary. In addition, an optimization of the implementation of our method would be beneficial, in particular for the linear structure detector (section 5.2.2), whose computation is still relatively slow despite our optimized algorithm and could greatly benefit from a parallelized implementation. Likewise, the minimal cut computation for large multitemporal graphs is relatively slow and requires a significant amount of memory. The use of an alternative, more efficient exact or approximate minimal cut solver could improve

this.

In addition to the SWOT mission, the proposed methods should also be considered for situations in which traditional methods based only on the information of one single SAR image fail to accurately extract water surfaces. In particular, such issues can be expected in the context of the proposed WiSA mission, which aims at acquiring measurements similar to SWOT, but with less complex instrument. Several applications on Sentinel-1 data, whose high revisit frequency is a valuable asset for water monitoring, could also be considered, possibly in combination with its optical counterpart Sentinel-2 or even with SWOT data. Such combination of information from several sensors is planned in future data hubs dedicated to hydrology and water detection methods for Sentinel 1 and 2 are being developed for that purpose. Beyond hydrology, other applications could benefit from the methods we proposed. For example, or narrow river detection method can be used to address the road discontinuity issues encountered by land use and land cover classification methods

Concerning denoising preprocessing for water detection, it could be worthwhile to adapt the detection methods we proposed for noisy images to denoised images. This preliminary denoising could be based on the temporal geometric mean (using denoising-by-ratio) or use single-date denoising methods, as we did for the detection of narrow rivers (Gasnier et al., 2021a).

Beyond denoising-by-ratio applications, water detection can also be done directly on the denoised temporal geometric mean, for example, to complete the current SWORD river database with rivers that are missing there, but that are included in the HydroRIVERS database (Linke et al., 2019). Indeed, the geometric temporal mean favors dark structures, even if mixed with bright outliers. In this way, seasonal rivers or rivers covered with ice and snow in the cold season are clearly visible while they would be faded in the temporal arithmetic mean image.

Concerning the use of auxiliary data, we limited the scope of our work to those that are available within the SWOT river and lake databases, but other kinds of auxiliary data can be beneficial to water detection. In particular, information on the water level of lakes, or an accurate, high-resolution digital surface model, should improve water detection if used within a multitemporal framework and are relatively straightforward to implement as an additional energy term in a graph.

Appendices

Appendix A

Combination of VV and VH polarization for Sentinel-1 images

Sentinel-1 SAR images consist of two channels: the co-polarized channel VV and the cross-polarized channel VH (see section 2.1.1.1). These two channels can be processed jointly through polarimetric approaches. However, the joint processing of the two images is not straightforward for methods that were designed to process single-channel images (such as the methods we propose in this thesis).

For methods that requires single-channel images, three options can be used:

1. Process only the co-polarized VV channel.
2. Process only the cross-polarized VH channel.
3. Process a combination of VV and VH channel.

Several possibilities exist to create a combination of VV and VH channel. A simple yet efficient option for water detection have been proposed by (Nunziata et al., 2016). It consists in computing the pixel-wise geometric mean of VV and VH intensities $\boldsymbol{I}_{VV}$ and $\boldsymbol{I}_{VH}$, which is equivalent to multiplying pixel-wise the VV and VH amplitude images. (Nunziata et al., 2016) used it for coastal line segmentation and in (Ferrentino et al., 2020) for segmenting lakes. In the following, we call this product VVVH product and the corresponding notation is $\boldsymbol{I}_{VVVH}$:

$$I_{VVVH}(x,y) = \sqrt{I_{VV}(x,y)I_{VH}(x,y)} = A_{VV}(x,y)A_{VH}(x,y). \tag{A.1}$$

The logarithm of $\boldsymbol{I}_{VVVH}$ is $\boldsymbol{y}_{VVVH} = \frac{\boldsymbol{y}_{VV}+\boldsymbol{y}_{VH}}{2}$

A.1 Comparison of VV, VH and VVVH for water detection

Following the experiments of (Nunziata et al., 2016), (Ferrentino et al., 2020) and (Ferrentino et al., 2017), we compared the performances of VV, VH and VVVH for water detection. For our experiments, we used 6 images associated with a manually defined ground truth, corresponding to the largest lakes in the test data presented in 6.1.4.1 to compare the ability of VV, VH and VVVH at separating water from land using a thresholding operation on noisy image. For each

image, we plotted the receiver operating characteristic (ROC) curve which is the false positive rate versus the true positive rate for every possible threshold value. **Figure** A.1 shows that VVVH is generally better than VV and VH at separating water from land. Indeed, VVVH (orange) is better than both VV and VH in the two Vioreau images and in images Sajnam 30 and LacDer 10. In image Sajnam 10, VV is slightly better than VVVH and much better than VH, because wet land can appear very dark in VH and cause confusions with water. On the contrary, in image LacDer 37, in which a very strong Bragg phenomenon affects mainly the VV channel, the VV image under-performs, as bright water is more likely to be confused with land.

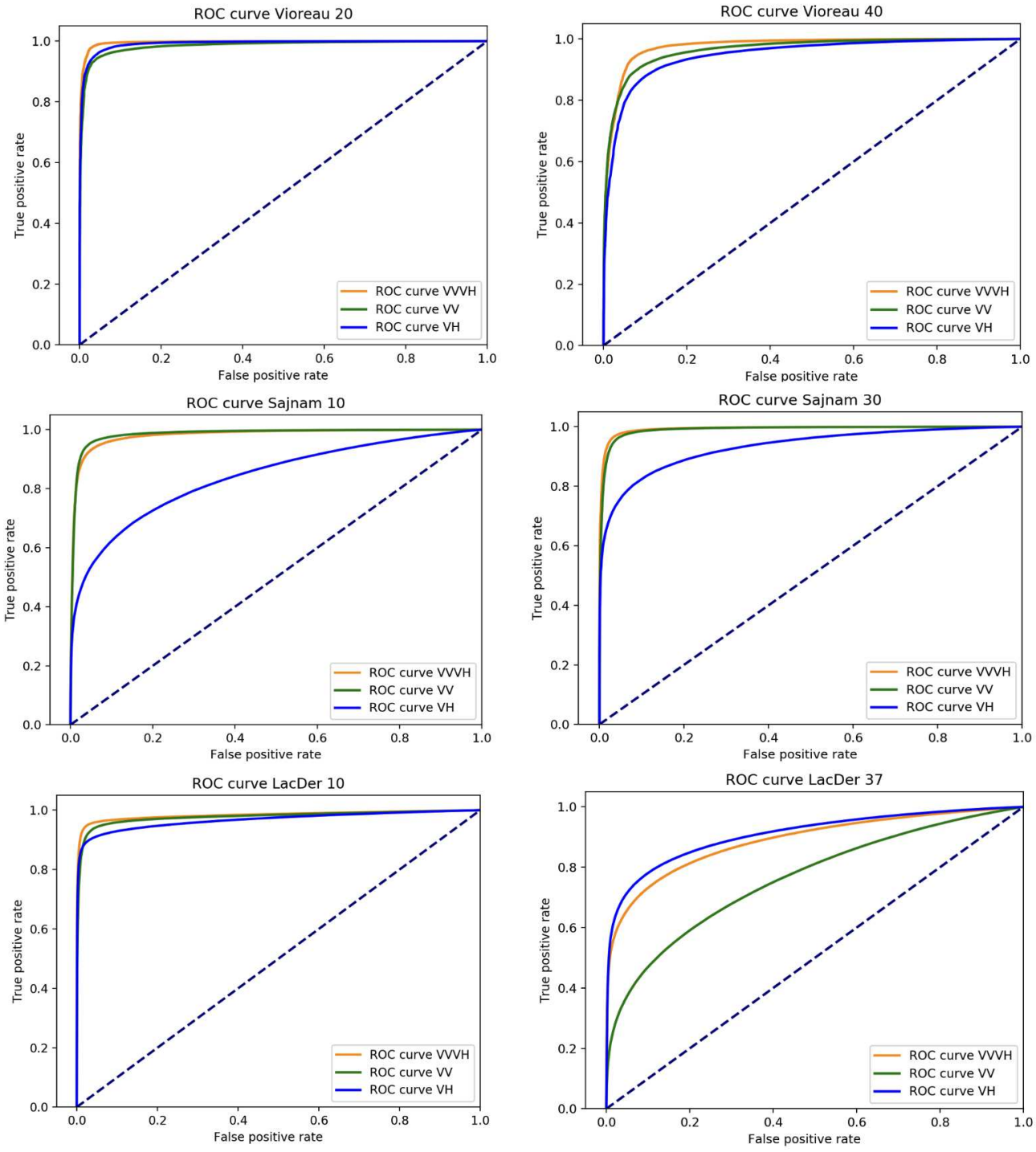


Figure A.1 – ROC curves comparing the performances of thresholding on VV (Green), VH (Blue) and VVVH (Orange) images for water detection (compared to a ground truth). Each figure corresponds to an image. Each row corresponds to a different lake, left and right figures correspond to different dates.

A.2 Modeling the VVVH combination

In the following, we will consider a co-polarized intensity channel $\boldsymbol{I}_{VV}$ and a cross-polarized channel $\boldsymbol{I}_{VH}$. Each channel follows a Gamma distribution (see equation A.2, recalled below, in section 2.1.3.1).

$$\mathrm{p}(I|R) = \frac{L^L I^{L-1}}{\Gamma(L)R^L} \exp\left(-L\frac{I}{R}\right) \tag{A.2}$$

According to (Nunziata et al., 2016), co- and cross-polarized channels (VV and VH) can be assumed to be uncorrelated.This assumption is derived from theoretical considerations on symmetry proprieties of the reflected pulses (Nghiem et al., 1992; Nunziata et al., 2012) on simple situations that meet Goodman's requirements, but according to (Nunziata et al., 2016), this property is "experimentally verified in a broad range of natural scenarios and it is so robust that it is generally assumed as reference to calibrate polarimetric SARs (van Zyl, 1990)".

In the following, we therefore assume I_{VV} and I_{VH} to be uncorrelated.

This way, the distribution of the geometric mean of the two Gamma distributed variables I_{VV} and I_{VH} can be expressed using equation 7.6 (see 7.2.2). However, this expression relies on Meijer functions and can hardly be used in practice for example to derive a likelihood. In the following, we will look for a simpler expression for the distribution of $y_{VV} = \log(I_{VV})$ and $y_{VH} = \log(I_{VH})$ by using a Gaussian approximation. Then, we will compare the distribution derived from these assumptions with numerical simulations order to determine if the assumption results in a large discrepancy between our approximate distribution and the actual distribution derived from Goodman's speckle model.

A.2.1 Simple expression for $y_{VVVH} = \log(I_{VVVH})$

As presented in 2.1.3.1, y_{VV} follows Fisher-Tippett distribution:

$$\mathrm{p}(y_{VV}|x_{VV}) = \frac{L^L}{\Gamma(L)} e^{L(y_{VV}-x_{VV})} \exp(-Le^{y_{VV}-x_{VV}}), \tag{A.3}$$

with $x_{VV} = \log(R_{VV})$ and similarly for y_{VH} with $x_{VH} = \log(R_{VH})$.

These distributions can be approximated with a Gaussian distribution if the value of L is high enough (see 2.1.3.1):

$$\mathrm{p}(y|\mu_{VV}) \simeq \frac{1}{\sigma_{VV}\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{y_{VV}-\mu_{VV}}{\sigma_{VV}}\right)^2}, \tag{A.4}$$

where $\sigma_{VV} = \sqrt{\psi'(L)}$, and $\mu_{VV} = x_{VV} - \log(L) + \psi(L)$ to match the standard deviation and the expected value of the previous distribution.

As y_{VV} and y_{VH} are uncorrelated, the arithmetic mean of their Gaussian approximation follows a Gaussian distribution of mean μ_{VVVH}:

$$\begin{aligned}\mu_{VVVH} =& \frac{\mu_{VV} + \mu_{VH}}{2} \\ =& \frac{x_{VV} + x_{VH}}{2} - \log(L) + \psi(L),\end{aligned} \tag{A.5}$$

and of variance σ^2_{VVVH}:

$$\begin{aligned}\sigma^2_{VVVH} =& \left(\frac{\sigma_{VV} + \sigma_{VH}}{2}\right)^2 \\ =& \frac{\psi'(L)}{2}\end{aligned} \tag{A.6}$$

If we assume that this arithmetic mean y_{VVVH}, which approximately follows a Gaussian distribution, can instead be modeled with a Fisher-Tippett distribution of parameters x_{VVVH} and L_{VVVH}, we have the following equations:

$$\begin{cases}\mu_{VVVH} = x_{VVVH} - \log(L_{VVVH}) + \psi(L_{VVVH}) = \frac{x_{VV}+x_{VH}}{2} - \log(L) + \psi(L) \\ \sigma^2_{VVVH} = \psi'(L_{VVVH}) = \frac{\psi'(L)}{2}\end{cases} \tag{A.7}$$

$\psi'(L_{VVVH}) = \frac{\psi'(L)}{2}$ leads to $L_{VVVH} = \psi'^{-1}(\frac{\psi(L)}{2})$, with ψ'^{-1} being the inverse of ψ' function. This yields the expression of x_{VVVH}:

$$\begin{cases}L_{VVVH} = \psi'^{-1}(\frac{\psi(L)}{2}) \\ x_{VVVH} = \frac{x_{VV}+x_{VH}}{2} + \log(\frac{L_{VVVH}}{L}) + \psi(L) - \psi(L_{VVVH})\end{cases} \tag{A.8}$$

We assumed that y_{VVVH} follows a Fisher-Tippett distribution and its exponential, I_{VVVH} follows a Gamma distribution:

$$\mathrm{p}(I_{VVVH}|R_{VVVH}) = \frac{L_{VVVH}^{L_{VVVH}} I_{VVVH}^{L_{VVVH}-1}}{\Gamma(L_{VVVH}) R_{VVVH}^{L_{VVVH}}} \exp\left(-L_{VVVH}\frac{I_{VVVH}}{R_{VVVH}}\right) \tag{A.9}$$

A.2.2 Comparison with numerical simulations

In this section, we compare the distribution for I_{VVVH} we obtained above using several approximation with the distribution of the geometric mean of simulated I_{VV} and I_{VH} variables. We compare the theoretical and simulated distributions (**Figure** A.2) as well as the values for L_{VVVH} and R_{VVVH} (table A.1). This comparison is done for $L = 1, L = 4$, and $L = 8$ using $R_{VV} = R_{VH} = 1$. These simulation are done with independent I_{VV} and I_{VH} but do not make any Gaussian approximation on the log-transformed intensities.

Value of L	Theoretical L_{VVVH}	Measured L_{VVVH}	Theoretical R_{VVVH}	Measured R_{VVVH}
1	1.650	1.609	0.783	0.784
4	7.53	7.54	0.940	0.939
8	15.5	15.5	0.969	0.969

Table A.1 – Comparison of theoretical and measured L_{VVVH} and R_{VVVH} for various values of L

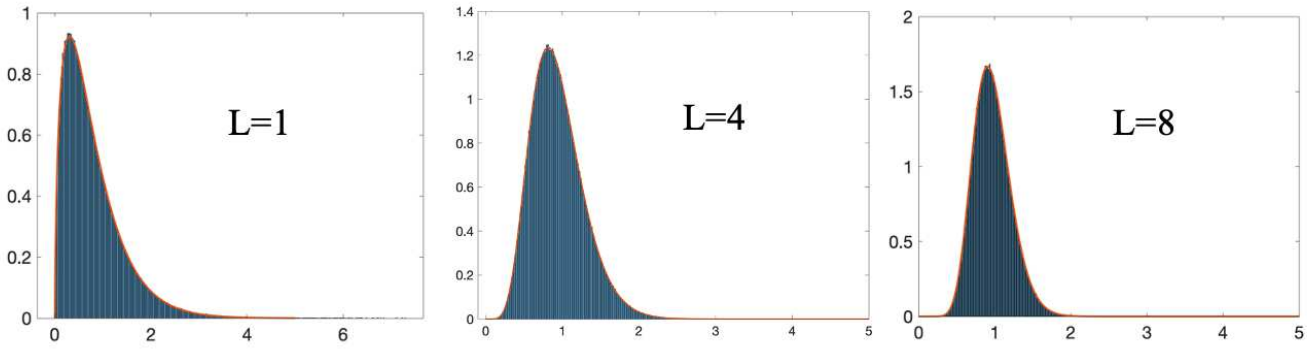


Figure A.2 – Comparison of the histograms of the simulated $\sqrt{I_{VV}I_{VH}}$ combination (blue) and proposed theoretical distributions (orange) for $L = 1, L = 4$ and $L = 8$.

These experiment show a good fit between the simulations and the proposed approximate distribution for I_{VVVH}.

A.3 Comparison on Sentinel-1 images

We complemented the previous experiments on simulated data with experiments on actual Sentinel-1 images. Figure A.3 shows a good fit between the histogram of the pixel-wise geometric mean $\sqrt{I_{VV}I_{VH}}$ on a homogeneous region of Lake Der and the corresponding theoretical distribution. We did also confirm that both the measured L_{VVVH} and R_{VVVH} fall within 1% of the value that is expected with the proposed formula given L, R_{VH}, and R_{VV}.

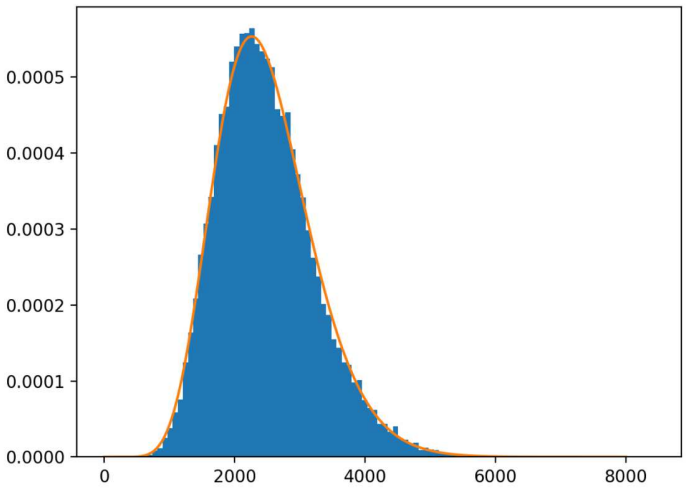


Figure A.3 – Histogram of the pixel-wise geometric mean $\sqrt{I_{VV}I_{VH}}$ on a homogeneous region of a Sentinel-1 GRD image of Lake Der (blue). Corresponding theoretical distribution (orange).

A.4 Conclusion

The geometric mean combination of the co- and cross-polarized channels $\sqrt{I_{VV} \cdot I_{VH}}$ was first proposed by (Nunziata et al., 2016). It enhances the contrast between water and land, which makes it promising for water detection. We propose to model it with a Gamma distribution (hence a Fisher-Tippett distribution for its logarithm) using our proposed formulas for L_{VVVH} and x_{VVVH}:

$$\begin{cases} L_{VVVH} = \psi'^{-1}(\frac{\psi(L)}{2}) \\ x_{VVVH} = \frac{x_{VV}+x_{VH}}{2} + \log(\frac{L_{VVVH}}{L}) + \psi(L) - \psi(L_{VVVH}) \end{cases} \tag{A.10}$$

For Sentinel-1 GRD images (L=4.4), this boils down to:

$$\begin{cases} L_{VVVH} = 8.33 \\ x_{VVVH} = \frac{x_{VV}+x_{VH}}{2} - 0.0567 \end{cases} \tag{A.11}$$

Our experiments on both simulations and actual Sentinel-1 images show that the proposed distribution very well fits the actual geometric mean $\sqrt{I_{VV}I_{VH}}$.

Appendix B

Comparison of the different graph cut models we propose

In this appendix, we compare the Graph-Cut based segmentation methods that are presented through the previous chapters and some reference Graph-Cut based segmentation models:

1. The segmentation step of the SWOT baseline water detection method, proposed in (Lobry, 2017).
2. The segmentation step of the narrow river guided detection method we propose in chapter 5.
3. The basic MRF method used in section 6.1.4 as a comparison.
4. The segmentation step of the original GrabCut method (Rother et al., 2004).
5. The segmentation step of our proposed 2D GrabCut method (section 6.1.3).
6. The segmentation step of our proposed 2D+T GrabCut method (section 6.3).
7. Our combined SAR-Optical water detection method (section 6.4).

Methods 1, 2 and 3 are presented in table B.1 and methods 4, 5, 6 and 7 are presented in table B.2.

Table B.1 – Comparison of methods 1, 2, and 3

	1 Segmentation step of the SWOT Baseline water detection method (Lobry, 2017)	2 CRF method for river segmentation around its centerline (section 5.2.4)	3 Basic MRF used as a comparison against our GrabCut approaches (section 6.1.4)
Input data	SWOT coherent power amplitude image	SAR intensity image	SAR log-intensity image
Data term	Derived from the neg-log-likelihood for a Rayleigh-Nakagami distribution	For water: derived from neg-log-likelihood for a Gamma distribution For land: Spatially uniform value for the data term.	Quadratic term based on a Gaussian assumption for the log-transformed intensities.
Water class prior	Estimated water reflectivity value in every pixel in parameter the estimation step	Estimated water reflectivity from the centerline	True water reflectivity estimated using the ground truth
Land class prior	Estimated land reflectivity value in every pixel in parameter the estimation step	-	True land reflectivity estimated using the ground truth
Hard constraints	-	The centerline pixels have to be classified as water	Pixels outside of the polygon have to be classified as land
Regularization term	Independant of the data (MRF)	Asymmetric CRF term, depends on magnitude, direction and orientation of the local gradient	Independent of the data (MRF)
Gradient used for the regularization	-	ROEWA gradient	-
Flux term	No	Yes	No

Table B.2 – Comparison of methods 4, 5, 6, and 7

	4 Segmentation step of the original GrabCut method (Rother et al., 2004)	5 The segmentation step of our proposed 2D GrabCut method (section 6.1.3).	6 The segmentation step of our proposed 2D+T GrabCut method (section 6.3)	7 Our combined SAR-Optical water detection method (section 6.4).
Input data	2D RGB natural image	2D SAR intensity image	2D+T stack of SAR log-intensity images	2D+T stack of SAR log-intensity and optical NDWI images
Data term	Linear combination of the neg-log-likelihoods for each of the components of a three-dimension Gaussian mixture model (GMM)	Linear combination of the neg-log-likelihoods for each of the components of a Fisher-Tippett mixture model (FTMM)		FTMM-derived term for SAR images. Quadratic term for Optical NDWI images
Water and land class prior	One estimated GMM for each	One estimated FTMM for each		One FTMM for SAR images for each. One NDWI water mean for optical images for each.
Hard constraints	Pixels outside of the initial polygon have to be classified as background	Pixels outside of the initial polygon have to be classified as land		-
Spatial regularization term	Symmetric CRF term, depends on magnitude and orientation of the local spatial gradient, not on its direction			Same for SAR images. No regularization for optical images
Gradient used for the spatial regularization	Finite differences	ROEWA gradient		ROEWA gradient for the SAR images
Temporal regularization term	-	-	Symmetric CRF term	MRF regularization
Flux term	No	Yes	Yes	Yes (SAR images only)

Appendix C

Notations used in the document

The following list condenses most of the notations used in this document.

In this document, bold variables corresponds to array or vectors, such as the intensity image $\boldsymbol{I}$, while scalars variables are written in roman typeface: the intensity of a particular pixel can be written I.

A pixel in a 2D image can be designated either by its coordinates (i, j) or, more compactly, by its index k.

Notation	Meaning	Type or Dimension
$\mathbf{1}$	A vector of ones	$(2N+1)(2n+1)$
$\widehat{\boldsymbol{d}_v}$	Variable used in the ADMM iterations	$W \times H$
f	Function $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$	Function
i	Index of a pixel along the first dimension	Scalar
j	Index of a pixel along the second dimension	Scalar
k	Designation of a pixel	Scalar
$\boldsymbol{\ell}$	Label field	$W \times H$
$n_{\mathcal{C}_W}$	Number of sub-classes for water	Scalar
$n_{\mathcal{C}_L}$	Number of sub-classes for land	Scalar
$\widehat{\boldsymbol{p}}_{k,\theta}$	Estimator for the reflectivity profile at pixel k with orientation θ.	Vector
$\hat{r}_k$	Biased estimator of the log of the uniform reflectivity (under $\mathcal{H}_1$) for the patch centered in k	Vector
$\widehat{\boldsymbol{r}}_{k,\hat{\theta}}$	Estimator of the log-reflectivities of the patch considering an orientation $\hat{\theta}$	Vector
s_c	cross-section of a corner reflector	Scalar
t	Discrete coordinate of a pixel along the temporal dimension	Scalar
u	Continuous coordinate in the image	Scalar
$\boldsymbol{w}_{\text{asym}}$	Asymmetric weighting term	Scalar
$\boldsymbol{w}_{\text{sym}}$	Symmetric weighting term	Scalar

$\boldsymbol{x}$	$\log(\boldsymbol{R})$	Scalar
$\boldsymbol{x_G}$	Log of the noise-free geometric mean	$W \times H$
$\widehat{\boldsymbol{x_G}}$	Log of the denoised geometric mean	$W \times H$
$x_{\mathcal{K}}$	Reflectivity parameter for the class $\mathcal{K}$	Scalar
$\boldsymbol{y}$	Log of the intensity	$W \times H$
$\boldsymbol{y}_v$	Variational variable	$W \times H$
$\widehat{\boldsymbol{z}}$	Variational split variable	$W \times H$
$\boldsymbol{y}_k$	Logarithm of $\boldsymbol{I}_k$	$W \times H$
A	Amplitude	Scalar
B	Radar baseline	Scalar
$B_G(L,T)$	Bias of the geometric mean	Scalar
$\boldsymbol{C}$	Cost array	$W \times H$
$\boldsymbol{C}_L$	Boolean raster for the presence of the centerline	$W \times H$
$\mathcal{C}$	Class (GrabCut)	Object
$\mathcal{C}_W$	Water class (GrabCut)	Object
$\mathcal{C}_L$	Land class (GrabCut)	Object
$\mathcal{K}$	Subclass (GrabCut)	Object
$\boldsymbol{D}$	Linear structure detector response	$W \times H$
D_{max}	Global maximum of $\boldsymbol{D}$	Scalar
$D_1(.,.)$	First derivative of the log-likelihood with respect to y	Function
$D_2(.,.)$	Second derivative of the log-likelihood with respect to y	Function
E_0	Reconstruction error under hypothesis H_0	Scalar
E_1	Reconstruction error under hypothesis H_1	Scalar
E_{Global}	Global energy	Scalar
H	Image size along second dimension	Scalar
$\mathcal{H}_0$	Null hypothesis	
$\mathcal{H}_1$	Alternative hypothesis	
$\boldsymbol{I}$	Intensity in one pixel	Scalar
$\boldsymbol{I}$	Image intensity	$W \times H$
$\boldsymbol{I}_k$	Vector formed by the intensity of all pixels in the patch $\boldsymbol{I}_{\square}(k)$	$W \times H$
$\boldsymbol{I}_{\mathrm{OR}}$	Time series combining SAR and optical images	$W \times H \times T$
$\boldsymbol{I}_{VV}$	Image intensity in VV polarization	$W \times H$
$\boldsymbol{I}_{VH}$	Image intensity in VH polarization	$W \times H$
$\boldsymbol{I}_{VVVH}$	Pixelwise geometric mean of $\boldsymbol{I}_{VV}$ and $\boldsymbol{I}_{VH}$	$W \times H$
$\boldsymbol{I}$	Image intensity	$W \times H$
I_0	First-kind modified Bessel function with order zero	Function
$\boldsymbol{I}_{\square}(k)$	Patch centered in the pixel k	$(2N+1)\times(2n+1)$

K	Multiplying factor for a temporal reflectivity	Scalar
K_C	High constant energy for U_{data}^C term.	Scalar
L	Equivalent number of looks	Scalar
$\mathscr{L}$	Neg-log-likelihood	Function
$\mathscr{L}_0$	Neg-log-likelihood for the land class	Function
$\mathscr{L}_1$	Neg-log-likelihood for the water class	Function
$\boldsymbol{\mathcal{L}}$	Likelihood	Function
$\mathrm{LoG}(\boldsymbol{I}, \sigma_L)$	Laplacian of Gaussian applied to the image $\boldsymbol{I}$	$W \times H$
$\boldsymbol{M}_\theta$	Matrix for the linear profile $\longrightarrow$ patch transform	Matrix
$\boldsymbol{M}_\theta^{\mathrm{pinv}}$	Moore-Penrose pseudoinverse of $\boldsymbol{M}_\theta$	Matrix
N	Interger. The patch size is $(2N+1) \times (2N+1)$	Scalar
N_{pow}	Parameter for the computation of the cost array	Scalar
$\boldsymbol{P}$	Reflectivity profile	Vector
$\boldsymbol{R}$	Reflectivity	xx
$\boldsymbol{R}_H k$	Reflectivity of an homogeneous patch	xx
$\widehat{\boldsymbol{R}}$	Estimator of the reflectivity	xx
R_1	Homogenous water reflectivity	Scalar
$\boldsymbol{I_G}$	Estimator for the geometric mean of the temporal reflectivities	$W \times H$
$\widetilde{\boldsymbol{I}}_{\boldsymbol{G}}$	Debiased estimator for the geometric mean of the temporal reflectivities	$W \times H$
S	Surface area in a continuous image	
$\boldsymbol{S}$	Vector giving the sensor for each image of the stack	T
S_{min}	Minimum scale for linear structures detection	Scalar
S_{max}	Minimum scale for linear structures detection	Scalar
T	Total number of images in the time series	Scalar
U_{data}^I	Data energy	Scalar
U_{data}^C	Centerline energy	Scalar
U_{flux}^C	Flux energy	Scalar
U_{reg}	Spatial regularization energy	Scalar
U_{TR}	Temporal regularization energy	Scalar
$\boldsymbol{V}_{n_{it}}$	Vector indicating which subclass a pixel belongs to	Vector
W	Image size along first dimension	Scalar
β	Regularization hyper-parameter (Markovian model)	Scalar
β_S	Spatial regularization hyper-parameter (Markovian model)	Scalar
β_T	Temporal regularization hyper-parameter (Markovian model)	Scalar

β_v	Speed of convergence tuning (variational method)	Scalar
γ_I	σ_I/μ_I ratio, called coefficient of variation of I	Scalar
γ_G	σ_G/μ_G ratio: the coefficient of variation of the geometric mean	Scalar
γ_A	σ_A/μ_A ratio: the coefficient of variation of the arithmetic mean	Scalar
η	Hyper-parameter for the flux term	Scalar
λ	Hyper-parameter for the exponential weighting of the CRF spatial regularization.	Scalar
λ_S	Water surface riddles wavelength.	Scalar
λ_R	Radar electromagnetic pulse wavelength	Scalar
λ_T	Hyper-parameter for the exponential weighting of the CRF temporal regularization.	Scalar
μ_I	Mean of the intensity distribution	Scalar
μ_R	Mean geometric parameter for the temporal distribution of the reflectivity	Scalar
$\boldsymbol{\mu}$	Vector containing the reflectivity of each each subclass of the GMM	Vector
$\boldsymbol{\pi}$	Vector containing the proportion of the GMM corresponding to each subclass	Vector
σ_{GA}	StD for the Gaussian approximation	Scalar
σ_A	StD of the temporal arithmetic mean	Scalar
σ_G	StD of the temporal geometric mean	Scalar
σ_I	StD of the intensity distribution	Scalar
σ_L	kernel size of the Laplacian of Gaussian filter	Scalar
σ_R	Sigma geometric parameter for the log-normal temporal distribution of the reflectivity	Scalar
σ_0	Radar backscattering coefficients	Scalar
$\widehat{\theta}$	Estimated orientation of the line	Scalar
θ_i	Sensor incidence angle	Scalar
$\psi(.)$	Digamma function	Function
$\psi'(.)$	Trigamma function	Function
$\Gamma(.)$	Gamma function	Function
$\Phi(.)$	Outward flux	Scalar

Appendix D

Publications

International journal

- Nicolas Gasnier, Loïc Denis, Florence Tupin
  SAR time-series denoising based on temporal geometric means
  IEEE Geoscience and Remote Sensing Letters (GRSL)
  Published 28 January 2021
  https://doi.org/10.1109/LGRS.2021.3051936
  Code available: https://gitlab.telecom-paris.fr/ring/geometric_mean_denoising

- Nicolas Gasnier, Loïc Denis, Roger Fjørtoft, Frédéric Liege, Florence Tupin
  Narrow River Extraction from SAR Images Using Exogenous Information
  IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS)
  Published 25 May 2021
  https://doi.org/10.1109/JSTARS.2021.3083413
  Code available: https://gitlab.telecom-paris.fr/ring/guided-river-detection

International conference with proceeding

- Nicolas Gasnier, Loïc Denis, Florence Tupin
  Generalized Likelihood Ratio Tests for Linear Structure Detection in SAR Images
  13th European Conference on Synthetic Aperture Radar 2021 (EUSAR)
  29 March 2021-4 April 2021 (Online)
  https://ieeexplore.ieee.org/document/9472518
  Code available: https://gitlab.telecom-paris.fr/ring/glrt_based_lines_detector

- Nicolas Gasnier ,Emanuele Dalsasso, Loïc Denis, Florence Tupin
  Despeckling Sentinel-1 GRD Images by Deep-learning and Application to Narrow River Segmentation
  IEEE's International Geoscience and Remote Sensing Symposium 2021 (IGARSS 2021)
  11-16 July 2021 (Online) https://doi.org/10.1109/IGARSS47720.2021.9554350

- Béatrice Pinel-Puysségur, Luca Maggiolo, Michel Roux, Nicolas Gasnier, David Solarna, Gabriele Moser, Sebastiano Bruno Serpico, Florence Tupin
  Experimental Comparison of registration methods for Multisensor SAR-optical Data
  IEEE's International Geoscience and Remote Sensing Symposium 2021 (IGARSS 2021)
  11-16 July 2021 (Online)
  https://doi.org/10.1109/IGARSS47720.2021.9553640

Other conferences

- Nicolas Gasnier, Roger Fjørtoft, Loïc Denis, Florence Tupin, Lucie Labat-Allée, Frédéric Liège.
  Robust Approaches for Water Detection in SWOT HR Interferometric SAR images
  HYDROSPACE-GEOGloWS, ESA-ESRIN 7-11 juin 2021, (Online)

Appendix E

Résumé en français

E.1 Contexte

Les eaux des lacs et des rivières représentent respectivement 0,006 % et 0,0002 % de la quantité totale d'eau sur Terre et n'occupent que 3,7 % et 0,6 % des surfaces terrestres non recouvertes de glace (Verpoorter et al., 2014; Allen and Pavelsky, 2018). Pourtant, elles jouent un rôle clé pour certains des principaux enjeux auxquels l'humanité est confrontée. L'eau est une ressource essentielle pour l'agriculture, l'industrie, ou même la vie quotidienne. Sa demande augmente plus rapidement que la population mondiale et sa pénurie représente une menace pour la santé et la sécurité alimentaire de plus de la moitié de la population mondiale (FAO, 2020, 2020; Mekonnen and Hoekstra, 2016). Les eaux de surface représentent également une menace directe, une proportion croissante de la population mondiale étant exposée au risque d'inondations (Tellman et al., 2021), ainsi qu'une menace indirecte, car les lacs et les barrages sont des facteurs majeurs de transmission du paludisme (Kibret et al., 2021).

De plus, dans un contexte de réchauffement climatique causé par les émissions de gaz à effet de serre, des changements majeurs dans le cycle de l'eau sont attendus (IPCC, 2021). Pour cette raison, nos connaissances des systèmes hydriques doivent être mises à jour en permanence.

Les lacs et les rivières jouent également un rôle longtemps sous-estimé dans les flux de ces gaz, en particulier de dioxyde de carbone, de méthane et de dioxyde d'azote une évaluation globale précise de leurs surfaces est nécessaire pour une bonne évaluation de ces émissions.

Pour relever ces défis, une surveillance et une gestion efficaces des ressources en eau douce sont nécessaires, mais seuls quelques pays développés y sont partiellement parvenus (UN-Water, 2021) à travers de très nombreux et coûteux relevés sur le terrain.

Dans ce contexte, il est crucial d'améliorer la collecte de données hydrologiques et la télédétection spatiale est essentielle, car elle permet l'acquisition de données à l'échelle mondiale. Les données spatiales ont été utilisées pour des applications en hydrologie dès qu'elles ont été disponibles. Par exemple, les images optiques et les images SAR (Synthetic Aperture Radar) ont permis de cartographier les surfaces en eau à l'échelle globale, notamment avec Landsat depuis 1972 et ERS depuis 1991, ou plus récemment avec les constellations radar Sentinel-1 et optique Sentinel-2. Les instruments SAR et optiques sont et resteront des sources de données essentielles pour l'hydrologie, mais il leur manque les informations de hauteur d'eau qui sont indispensables pour évaluer le débit des rivières et l'évolution des réserves d'eau contenues dans les lacs. Cette

information peut actuellement être mesurée depuis un satellite avec des altimètres nadir tels que ceux de la série Poséidon (satellites TOPEX puis Jason) ou le SRAL (Sar Radar ALtimeter) de Sentinel-3, qui fournissent l'information sur la hauteur d'eau à une échelle spatiale grossière (résolution spatiale kilométrique) et ont révolutionné l'océanographie. Cependant, leur résolution spatiale limite leurs applications hydrologiques aux très grands lacs et fleuves.

La mission SWOT (Surface Water and Ocean Topography) vise à dépasser cette limite avec un altimètre de fauchée qui mesurera l'élévation de l'eau sur une grille spatiale en deux dimensions à haute résolution. L'instrument fonctionnera en effectuant des calculs interférométriques sur une paire d'images SAR. Ces images seront acquises simultanément avec un angle d'observation proche du nadir. SWOT fournira également des données pour des applications océanographiques qui ne seront pas mentionnées dans cette thèse.

La mission SWOT est une collaboration entre le Centre National d'Études Spatiales (CNES) français et le Jet Propulsion Laboratory (JPL) de la NASA, avec des contributions de l'Agence Spatiale Canadienne (ASC) et de l'Agence Spatiale du Royaume-Uni (UKSA). Les tâches sont réparties entre ces agences spatiales, avec le soutien de partenaires universitaires et industriels locaux. Le CNES est responsable des algorithmes de détection de l'eau, qui constituent une étape clé dans le traitement des données à haut débit (HR) de SWOT. La méthode opérationelle pour la détection de l'eau a été développée au cours des travaux de doctorat de Sylvain Lobry à Télécom Paris en collaboration avec le CNES et est en cours de calibration et de test sur des images issues de simulations avec le soutien de CS Group France.

La mission SWOT repose sur une technologie de capteur qui n'avait jamais été embarquée sur un satellite. Dans ce contexte, il existe encore des incertitudes sur ses performances, notamment en termes de rapport signal/bruit (SNR), avec des conséquences sur la capacité à détecter les surfaces d'eau et à mesurer leur élévation. L'objectif de ce travail est de fournir des méthodes alternatives de détection de l'eau qui soient plus robustes pour être capables de détecter les surfaces d'eau même dans les situations où la méthode opérationnelle échoue. Cette approche fait partie d'une démarche d'atténuation des risques pour la mission SWOT, avec un potentiel d'application à d'autres capteurs SAR également. Pour cela, nous nous sommes intéressés à trois stratégies pour de rendre la détection plus robuste :

- Utilisation de données externes aux images pour guider la détection
- Approches multi-temporelles et multi-capteurs
- Utilisation d'une étape préalable de débruitage

Au-delà de la mission SWOT, nos travaux sur la détection de l'eau dans les images SAR pourront être utilisés pour d'autres capteurs SAR tels que Sentinel-1, qui resteront utiles pour l'hydrologie en complément des données SWOT : SWOT ne rendra pas les autres capteurs obsolètes pour les applications en hydrologie. Au contraire, leurs données pourront être combinées, par exemple dans le cadre de centres de données et de services tels que le futur centre de données HYSOPE II du CNES. Notre travail peut également être utile dans le contexte de la mission d'altimétrie interférométrique WiSA (actuellement à l'état de concept) qui pourrait remplacer SWOT après sa fin de vie, avec un capteur moins complexe et potentiellement moins performant.

Les méthodes que nous proposons sont conçues pour être exécutées sur des images à une seule polarisation, comme les images de *coherent power* de SWOT. Pour traiter des images Sentinel-1 en double polarisation, elles pourront être appliquées soit sur un seul canal, soit sur la moyenne géométrique des canaux, calculée pixel par pixel, comme on le présente en annexe A.

E.2 Contributions

Les principales contributions de ces travaux sont liées à trois applications :

1. Une méthode de détection robuste des rivières fines guidée par une base de données exogène des lignes centrales des rivières.
2. Des approches guidées pour la détection de lacs sur des images SAR en utilisant des informations exogènes, des données multitemporelles, ou même en combinant des informations provenant d'images optiques et SAR.
3. Dans la perspective d'utiliser une étape de débruitage préalable à la détection de l'eau, nous avons travaillé sur l'utilisation et le débruitage de la moyenne géométrique temporelle d'une série d'images SAR.

E.2.1 Détection de rivières fines guidée par des données exogènes

Contribution (1): Détection de rivières fines guidée des données exogènes

Cette contribution est présentée dans notre article (Gasnier et al., 2021b) et dans le chapitre 5.

Compte tenu de la disponibilité d'une base de données mondiale des rivières, nous avons proposé un cadre pour la détection guidée de rivières fines qui est robuste à la fois à un faible rapport signal sur bruit dans les images et à des erreurs dans l'information *a priori* issue de cette base de données.
Cette méthode repose sur trois étapes (cf. Figure E.1) :

1. Un nouveau détecteur de structures linéiques (contribution 1-A)
2. Une étape de repositionnement de la ligne centrale de la rivière basée sur des informations exogènes et sur la réponse du détecteur de structures linéiques, en utilisant un algorithme qui calcule le chemin de moindre coût
3. Une segmentation autour de la ligne centrale qui utilise un nouveau modèle de champ aléatoire conditionnel (CRF) (contribution 1-B)

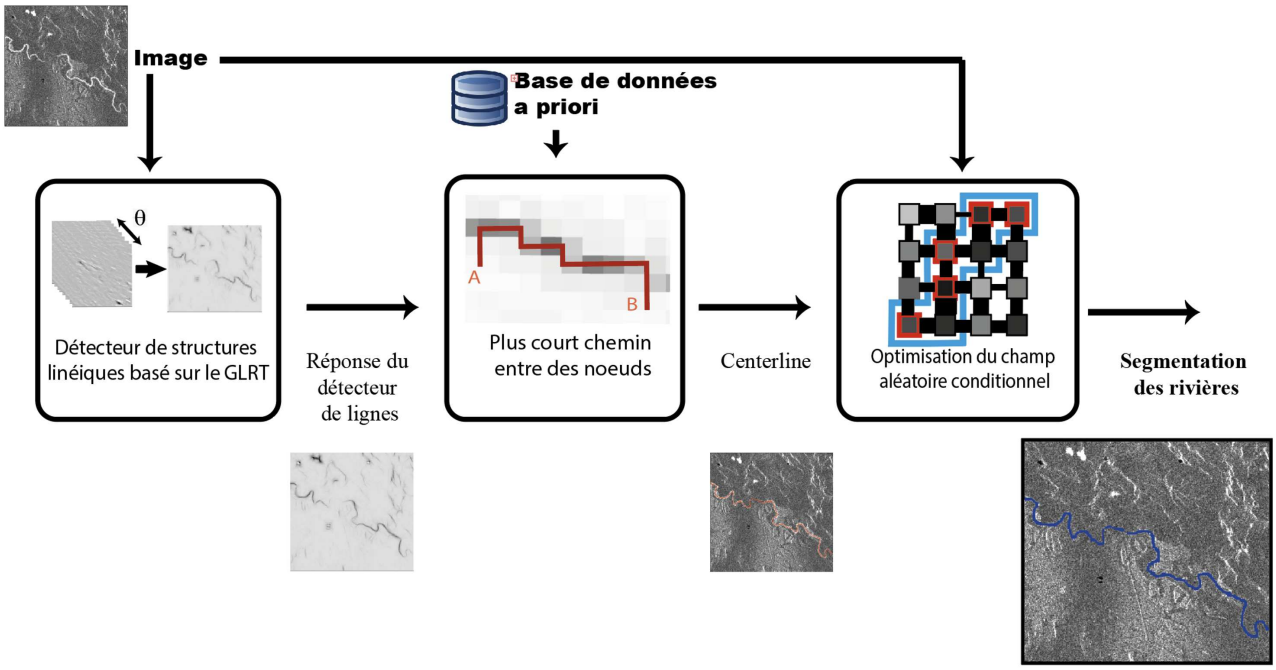


Figure E.1 – Principe général de la méthode de détection des rivières fines que nous proposons

Contribution (1-A): Détection de structures linéiques

Cette contribution est détaillée dans nos articles (Gasnier et al., 2021c; Gasnier et al., 2021b) et dans la section 5.2.2.

Nous proposons un nouveau détecteur de structures linéiques pour les images SAR, basé sur le rapport de vraisemblance généralisé (GLR). Il compare la vraisemblance d'un patch en considérant les réflectivités estimées sous deux hypothèses :

- $\mathcal{H}_0$: il n'y a pas de structure linéique
- $\mathcal{H}_1$: il existe une structure linéique

L'application pratique de cette méthode pour les images SAR est rendue possible par l'approche optimisée que nous présentons. Nous comparons cette méthode au détecteur de structures linéiques de (Tupin et al., 1998) et montrons qu'elle donne lieu à moins de fausses détections.

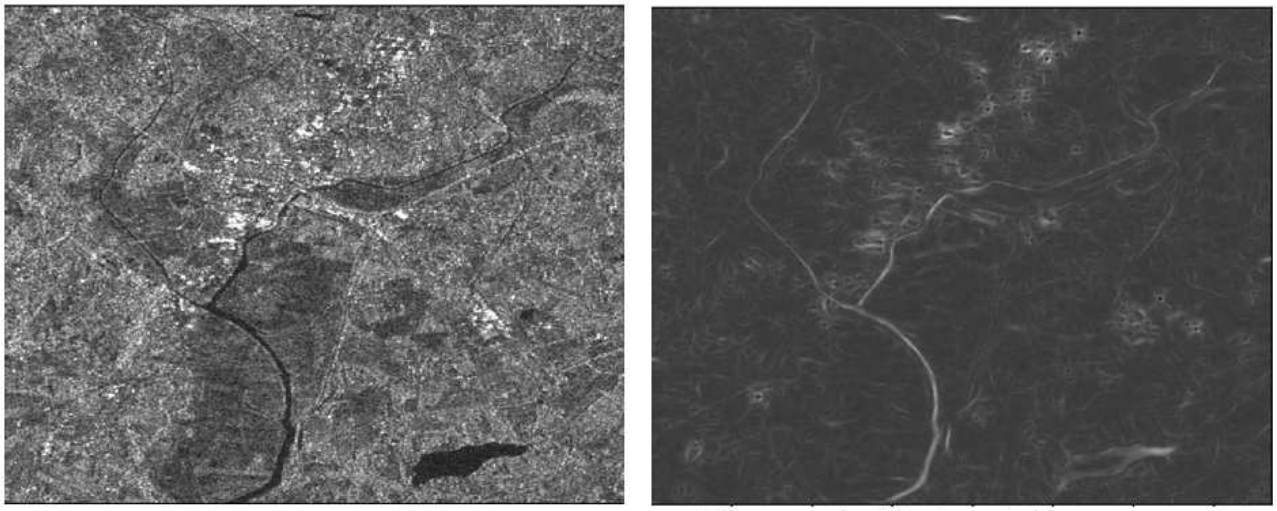


Figure E.2 – Exemple de résultat obtenu avec le détecteur de structures linéiques proposé

Contribution (1-B): Modèle de segmentation CRF pour la détection de rivières fines autour d'une ligne centrale

Cette contribution est détaillée dans notre article (Gasnier et al., 2021b) et dans la section 5.2.4.

Pour la dernière étape de la méthode de détection de rivières fines que nous proposons, nous devons détecter la surface de l'eau autour de la ligne centrale reconstruite à l'étape précédente, sans connaissance préalable des réflectivités de la terre et de l'eau. Pour cela, nous avons proposé un nouveau modèle CRF (conditional random field). Ce nouveau modèle CRF combine quatre termes :

- Un terme d'attache aux données qui est différent pour la classe d'eau et pour la classe de terre. Pour l'eau, il dérive d'un modèle statistique qui considère une réflectivité de l'eau estimée à partir des pixels de la ligne centrale. Pour la terre, comme nous ne disposons d'aucune information sur les réflectivités sous-jacentes, nous utilisons un *a priori* non informatif qui se traduit par un terme de données spatialement uniforme dont la valeur a été choisie de manière à ne pas introduire de biais en faveur de l'une des classes.
- Un terme de ligne centrale qui empêche la ligne centrale d'être classée comme terre.
- Un terme de régularisation asymétrique basé sur un modèle CRF, qui prend en compte le fait que l'eau est plus claire que la terre (pour SWOT) ou plus sombre que la terre (pour les systèmes SAR conventionnels tels que Sentinel-1). Ainsi, il tient compte du signe du gradient et pas uniquement de son module.
- Un terme de flux du gradient, qui compense les conséquences de la régularisation dans certaines situations où c'est nécessaire.

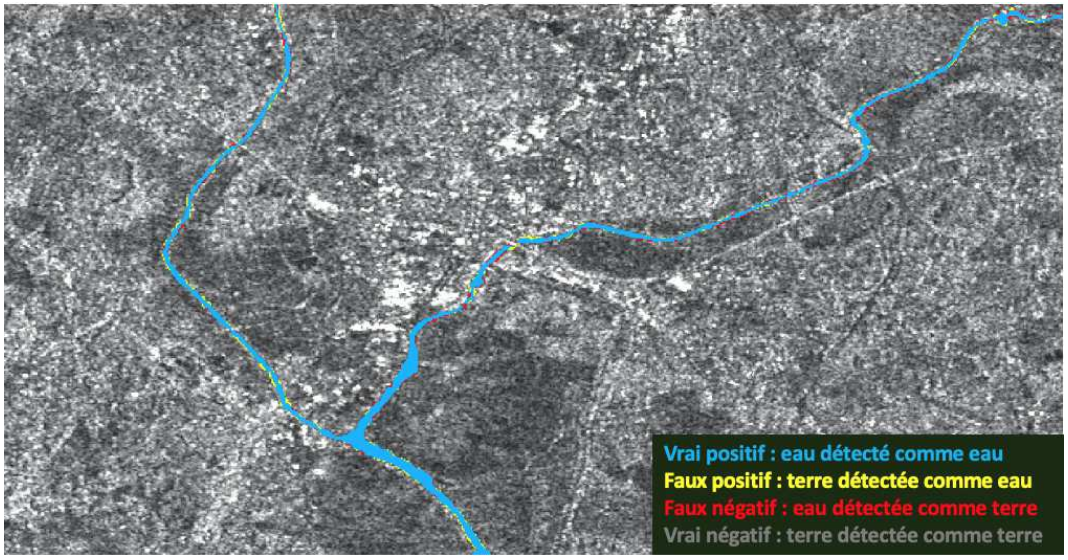


Figure E.3 – Résultat de segmentation en utilisant la méthode proposée sur une image Sentinel-1 GRD (Redon, France). Quatre points *a priori* ont été utilisés : un à chaque extrémité des rivières et un à leur intersection.

E.2.2 Approche de segmentation des lacs dérivée de GrabCuts

Pour améliorer la détection des lacs, en particulier ceux de petite superficie ou de forme irrégulière, nous avons proposé d'utiliser des informations de type *a priori* sous la forme d'un polygone de délimitation grossier pour chaque lac et de combiner des données multitemporelles et multi-capteurs. Pour utiliser un polygone de délimitation grossier en entrée, une approche dérivée de la méthode GrabCut (Rother et al., 2004) est tout à fait adaptée.

Pour cette tâche, nous avons proposé trois méthodes :

1. Une adaptation de la méthode GrabCut pour la détection de l'eau dans les images SAR à date unique. Cette méthode n'utilise pas de connaissances préalables sur les réflectivités de l'eau et de la terre, mais prend en entrée un polygone de délimitation grossier.

2. Une extension multitemporelle de la méthode précédente. Un exemple de résultat est présenté Figure E.4.

3. Une méthode multitemporelle et multi-capteurs qui traite une série temporelle combinée d'images SAR et optiques. Contrairement aux deux méthodes précédentes, elle ne prend pas en entrée un polygone de délimitation, mais utilise des modèles statistiques antérieurs pour les classes d'eau et de terre dans les images SAR et optiques.

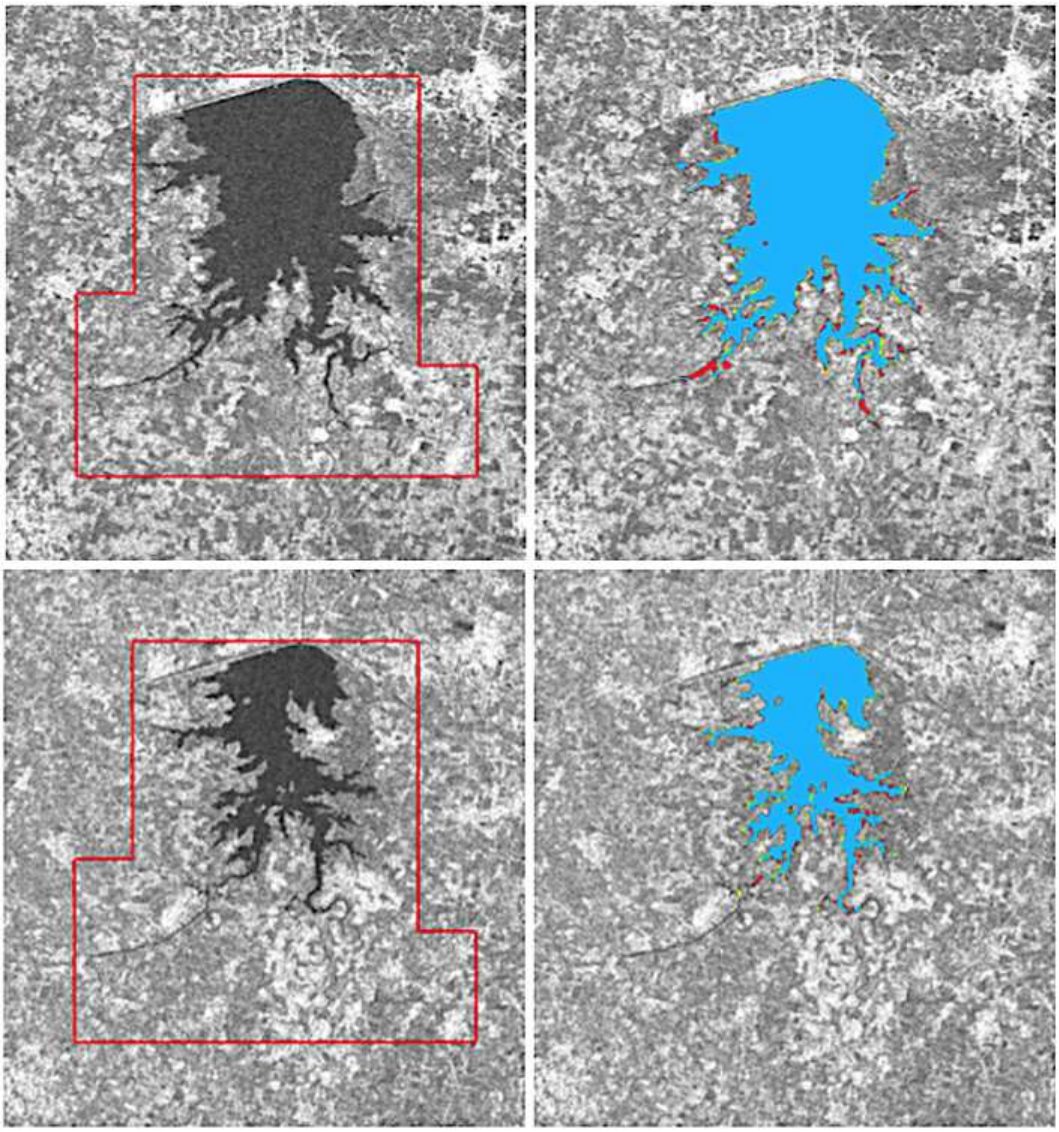


Figure E.4 – Pour deux dates différentes de la série d'images Sajnam (haut et bas), masque a priori (à gauche) et résultat de détection d'eau (à droite). Les pixels bien classés comme eau apparaissent en bleu, les faux positifs et faux négatifs respectivement en jaune et en rouge.

Contribution (2): Approches multitemporelles et multi-capteurs guidées pour la détection des lacs dans les images SAR

Cette contribution est présentée dans le chapitre 6.

Nous avons adapté la méthode GrabCut, proposée à l'origine pour la détection de n'importe quel objet dans des images RVB naturelles, à la détection d'eau dans des images SAR. Pour ce faire, nous avons adapté les modèles de mélange utilisés aux statistiques des images SAR et ajouté un terme de flux au modèle de segmentation afin de favoriser la détection de structures plus sombres (pour Sentinel-1 ou d'autres capteurs SAR conventionnels) ou plus claires (pour SWOT) que leur arrière-plan.

Nous avons ensuite adapté ce modèle aux séries temporelles SAR et ajouté un terme de régularisation temporelle qui améliore la localisation des contours du lac pour une date donnée tout en préservant les évolutions temporelles de la surface du lac.

Enfin, nous avons proposé une approche de segmentation non guidée pour des séries temporelles combinées d'images SAR et optiques. Les distributions statistiques pour la terre et l'eau qui sont utilisées comme entrée peuvent être déterminées pour les images SAR en utilisant la méthode précédente, et pour les images optiques, par une approche de clustering externe telle que (Cordeiro et al., 2021). Cette approche non guidée pourrait également être intégrée dans une méthode GrabCut multitemporelle combinée SAR et optique.

E.2.3 Débruiter la moyenne géométrique temporelle pour faciliter la détection des surfaces d'eau

Une étape préalable de débruitage peut faciliter la détection des surfaces d'eau. Par exemple, nous avons présenté à la conférence IGARSS 2021 (Gasnier et al., 2021a) des résultats préliminaires d'amélioration de la détection de rivières fines en utilisant une étape préalable de débruitage des images SAR. Cette étape de débruitage pourrait bénéficier de l'information temporelle contenue dans la série temporelle, et les propriétés de la moyenne géométrique en font un bon candidat pour ça, par exemple en utilisant une méthode de débruitage par ratio (Zhao et al., 2019).

Contribution (3): Propriétés statistiques et débruitage de la moyenne géométrique temporelle d'une série temporelle images SAR

Cette contribution est détaillée dans notre article (Gasnier et al., 2021d) et dans la section 7

.

Nous avons étudié les propriétés statistiques de la moyenne géométrique temporelle et l'avons comparée à la moyenne arithmétique. Nous avons montré que la moyenne géométrique est plus intéressante dans les situations avec des valeurs très claires transitoires ou avec une réflectivité fluctuant dans le temps. Comme il n'existe pas d'expression analytique pour la distribution de cette moyenne avec des images SAR, nous avons proposé une approche numérique pour son estimation et l'avons utilisée dans un cadre variationnel pour débruiter l'image de la moyenne géométrique (Figure E.5). Nous avons utilisé cette moyenne géométrique temporelle débruitée dans plusieurs applications telles que la détection des changements ou le débruitage par ratio et avons montré qu'elle améliore le résultat du débruitage dans certaines situations par rapport à la moyenne arithmétique (Figure E.6).

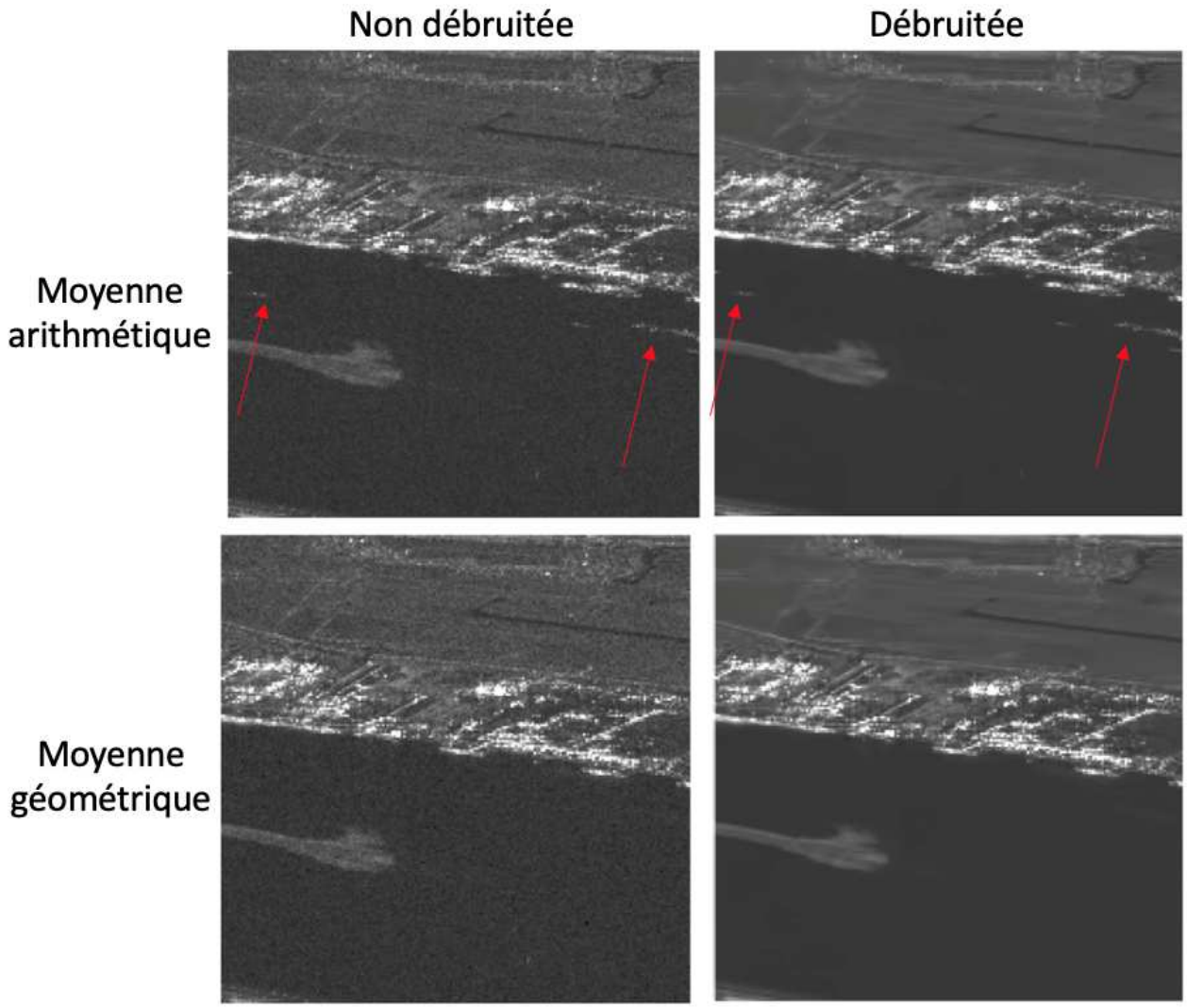


Figure E.5 – Pour une série temporelle d'images Sentinel-1 SLC (port de Saint-Nazaire), moyenne temporelle arithmétique (en haut) et géométrique (en bas) non débruitées (à gauche) et débruitées (à droite). Les flèches en rouge mettent en évidence des structures claires correspondant à des bateaux qui ne sont présents que sur quelques images de la série.

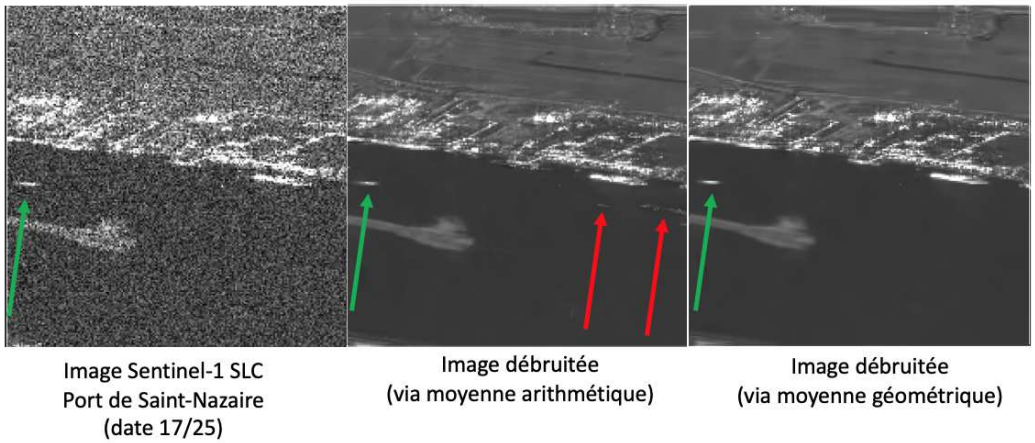


Figure E.6 – Résultats obtenus par débruitage par ratio de la 17ème image de la série (gauche) en utilisant comme super-image la moyenne arithmétique (centre) ou la moyenne géométrique (droite). On remarque que le bateau qui est présent sur l'image (flèche verte) est bien présent sur les images débruitées par les deux méthodes. Par contre, avec l'utilisation du débruitage par ratio avec la moyenne arithmétique comme super-image, des structures correspondant à des bateaux absents de l'image à cette date apparaissent.

E.2.4 Synthèse

La répartition de nos contributions entre les trois stratégies mentionnées plus haut (guidage par des données exogènes, combinaison multi-temporelles et multi-capteurs, et débruitage préalable) est décrite par la Figure E.7.

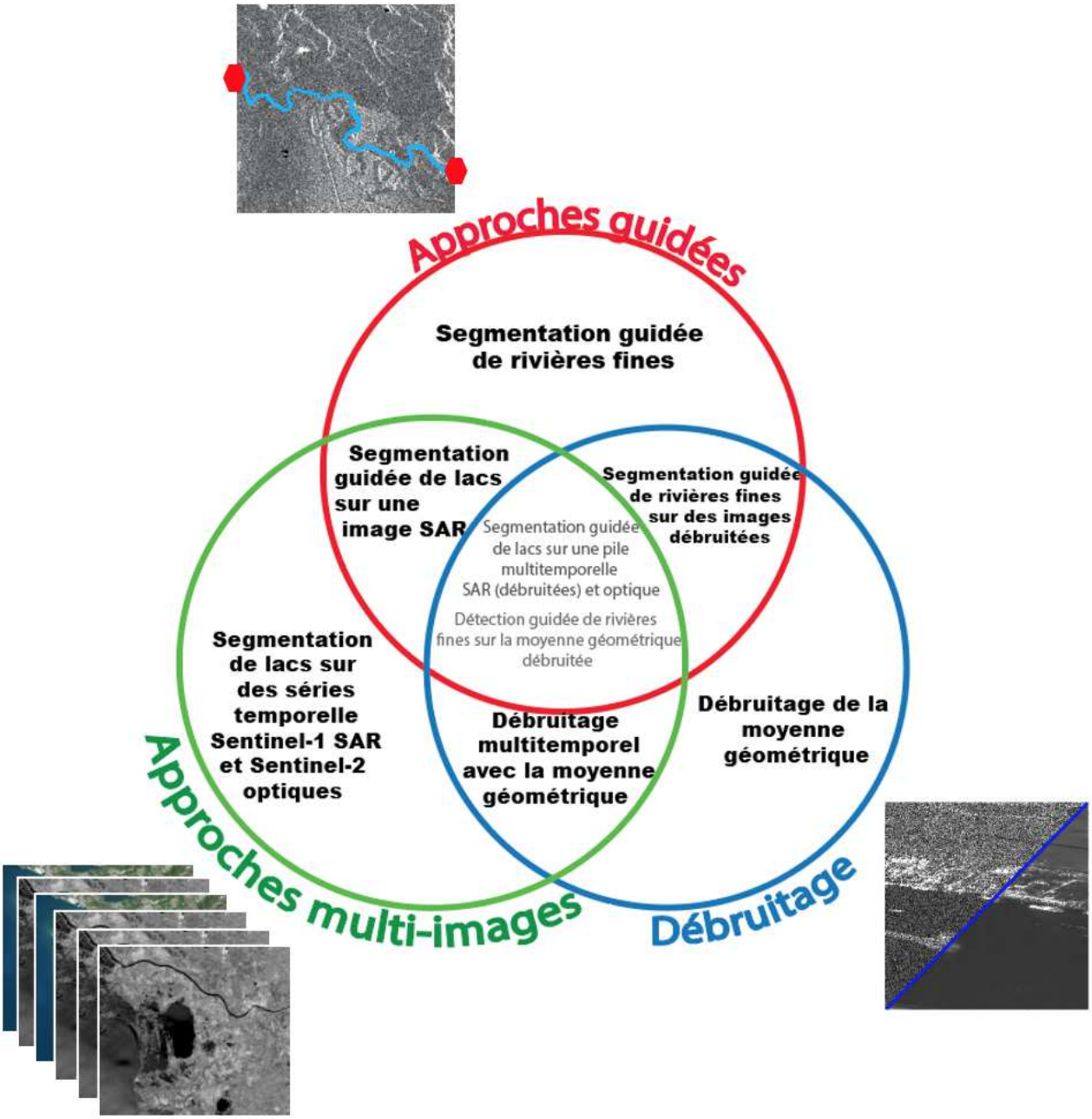


Figure E.7 – Répartition des approches proposées entre les trois stratégies mentionnées plus haut. La détection guidée des rivières fines avec des images débruitées de leur contenu n'est pas détaillée dans ce manuscrit, mais a été présentée à la conférence IGARSS 2021. Les approches correspondant à des perspectives non encore publiées apparaissent en gris.

Bibliography

Allen, G. H. and Pavelsky, T. M. (2018). Global extent of rivers and streams. *Science*, 361(6402):585–588. 13, 49, 179

Almhana, J., Liu, Z., Choulakian, V., and McGorman, R. (2006). A recursive algorithm for gamma mixture models. In *2006 IEEE International Conference on Communications*, volume 1, pages 197–202. 111

Altenau, E. H., Pavelsky, T. M., Durand, M. T., Yang, X., Frasson, R. P. d. M., and Bendezu, L. (2021). The surface water and ocean topography (SWOT) mission river database (SWORD): A global river network for satellite data products. *Water Resources Research*, 57(7):e2021WR030054. 49, 50

Altenau, E. H., Pavelsky, T. M., Moller, D., Lion, C., Pitcher, L. H., Allen, G. H., Bates, P. D., Calmant, S., Durand, M., and Smith, L. C. (2017). AirSWOT measurements of river water surface elevation and slope: Tanana River, AK. *Geophysical Research Letters*, 44(1):181–189. 35

Amhaz, R., Chambon, S., Idier, J., and Baltazart, V. (2016). Automatic crack detection on two-dimensional pavement images: An algorithm based on minimal path selection. *IEEE Transactions on Intelligent Transportation Systems*, 17(10):2718–2729. 66

Argenti, F., Lapini, A., Bianchi, T., and Alparone, L. (2013). A tutorial on speckle reduction in synthetic aperture radar images. *IEEE Geoscience and Remote Sensing Magazine*, 1(3):6–35. 42

Atto, A. M., Trouvé, E., Nicolas, J., and Lê, T. T. (2016). Wavelet operators and multiplicative observation models—application to SAR image time-series analysis. *IEEE TGRS*, 54(11):6606–6624. 145

Aubert, G. and Aujol, J.-F. (2008). A variational approach to removing multiplicative noise. *SIAM Journal of Applied Mathematics*, 68:925–946. 57, 58

Beaulieu, J. J., Tank, J. L., Hamilton, S. K., Wollheim, W. M., Hall, R. O., Mulholland, P. J., Peterson, B. J., Ashkenas, L. R., Cooper, L. W., Dahm, C. N., Dodds, W. K., Grimm, N. B., Johnson, S. L., McDowell, W. H., Poole, G. C., Valett, H. M., Arango, C. P., Bernot, M. J., Burgin, A. J., Crenshaw, C. L., Helton, A. M., Johnson, L. T., O'Brien, J. M., Potter, J. D., Sheibley, R. W., Sobota, D. J., and Thomas, S. M. (2011). Nitrous oxide emission from denitrification in stream and river networks. *Proceedings of the National Academy of Sciences*, 108(1):214–219. 13

Bernauer, T. and Böhmelt, T. (2020). International conflict and cooperation over freshwater resources. *Nature Sustainability*, 3(5):350–356. 13

Bioucas-Dias, J. M. and Figueiredo, M. A. T. (2010). Multiplicative noise removal using variable splitting and constrained optimization. *IEEE Transactions on Image Processing*, 19(7):1720–1730. 58, 59

Bousbih, S., Zribi, M., Lili-Chabaane, Z., Baghdadi, N., El-Hajj, M., Gao, Q., and Mougenot, B. (2017). Potential of Sentinel-1 Radar Data for the Assessment of Soil and Cereal Cover Parameters. In *Sensors*. 148

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3:1–122. 58

Boykov, Y. and Jolly, M.-P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 1, pages 105–112 vol.1. 56, 87

Boykov, Y. and Kolmogorov, V. (2004). An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137. 53, 54, 91, 95, 107, 117, 118

Boykov, Y., Veksler, O., and Zabih, R. (1998). Markov random fields with efficient approximations. In *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231)*, pages 648–655. 53

Brekke, C. and Solberg, A. H. (2005). Oil spill detection by satellite remote sensing. *Remote Sensing of Environment*, 95(1):1–13. 28

Brisco, B., Ahern, F., Murnaghan, K., White, L., Canisus, F., and Lancaster, P. (2017). Seasonal change in wetland coherence as an aid to wetland monitoring. *Remote Sensing*, 9(2). 127

Buzzard, G. T., Chan, S. H., Sreehari, S., and Bouman, C. A. (2018). Plug-and-play unplugged: Optimization-free reconstruction using consensus equilibrium. *SIAM Journal on Imaging Sciences*, 11(3):2001–2020. 59

Canisius, F., Brisco, B., Murnaghan, K., Van Der Kooij, M., and Keizer, E. (2019). SAR backscatter and InSAR coherence for monitoring wetland extent, flood pulse and vegetation: A study of the amazon lowland. *Remote Sensing*, 11(6). 127

Cao, F., Tupin, F., Nicolas, J., Fjørtoft, R., and Pourthié, N. (2011). Extraction of water surfaces in simulated Ka-band SAR images of KaRIn on SWOT. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 3562–3565. 43, 66

Carayon, G., Steunou, N., Courriere, J.-L., and Thibaut, P. (2002). Poseidon 2 radar altimeter design and in flight preliminary performances. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 774–776 vol.2. 33

Carroll, M., Townshend, J., DiMiceli, C., Noojipady, P., and Sohlberg, R. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2(4):291–308. 48

Cazals, C., Rapinel, S., Frison, P.-L., Bonis, A., Mercier, G., Mallet, C., Corgne, S., and Rudant, J.-P. (2016). Mapping and characterization of hydrological dynamics in a coastal marsh using high temporal resolution Sentinel-1A images. *Remote Sensing*, 8(7):570. 42

Chambolle, A. and Lions, P.-L. (1997). Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188. 57

Chan, S. H., Wang, X., and Elgendy, O. A. (2017). Plug-and-Play ADMM for Image Restoration: Fixed-Point Convergence and Applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98. 58, 152

Chan, T. and Vese, L. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277. 43

Chanussot, J., Issa, I., and Lambert, P. (1998). Adaptive directional order filters and mathematical morphology for road network extraction on SAR images. In *9th European Signal Processing Conference (EUSIPCO 1998)*, pages 1–4. 66

Chanussot, J., Mauris, G., and Lambert, P. (1999). Fuzzy fusion techniques for linear features detection in multitemporal SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1292–1305. 68

Chaturvedi, S., Banerjee, S., and Lele, S. (2019). An assessment of oil spill detection using Sentinel-1 SAR-C images. *Journal of Ocean Engineering and Science*, 5. 29

Christiano, P., Kelner, J. A., Madry, A., Spielman, D. A., and Teng, S.-H. (2011). Electrical flows, laplacian systems, and faster approximation of maximum flow in undirected graphs. New York, NY, USA. Association for Computing Machinery. 54, 115

Cordeiro, M., Martinez, J., and Peña-Luque, S. (2021). Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. *Remote Sensing of Environment*, 253. 17, 48, 136, 187

Coulthard, T. and Van De Wiel, M. (2012). Modelling river history and evolution. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1966):2123–2142. 49, 64

Cumming, I. and Wong, F. (2005). *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Number vol. 1 in Artech House remote sensing library. Artech House. 21

Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K. (2007). Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095. 42, 143, 152

Dabov, K., Foi, A., Katkovnik, V., and Egiazarian, K. (2007). Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095. 58

D'Addabbo, A., Refice, A., Pasquariello, G., Lovergine, F. P., Capolongo, D., and Manfreda, S. (2016). A Bayesian Network for Flood Detection Combining SAR Imagery and Ancillary Data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6):3612–3625. 43

Dalsasso, E., Denis, L., and Tupin, F. (2021). SAR2SAR: A semi-supervised despeckling algorithm for SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:4321–4329. 42, 158, 160

Dantzig, G. B. and Fulkerson, D. R. (1955). On the max flow min cut theorem of networks. 54

Delagnes, P. and Barba, D. (1995). A markov random field for rectilinear structure extraction in pavement distress image analysis. In *Proceedings., International Conference on Image Processing*, volume 1, pages 446–449 vol.1. 66

Deledalle, C., Denis, L., Tabti, S., and Tupin, F. (2017a). MuLoG, or how to apply gaussian denoisers to multi-channel SAR speckle reduction? *IEEE Transactions on Image Processing*, 26(9):4389–4403. 42, 58, 59, 142, 143

Deledalle, C., Denis, L., Tabti, S., and Tupin, F. (2017b). MuLoG, or How to Apply Gaussian Denoisers to Multi-Channel SAR Speckle Reduction? *IEEE Transactions on Image Processing*. 152

Deledalle, C.-A., Denis, L., Poggi, G., Tupin, F., and Verdoliva, L. (2014). Exploiting patch similarity for SAR image processing: The nonlocal paradigm. *IEEE Signal Processing Magazine*, 31(4):69–78. 42

Delignon, Y., Garello, R., and Hillion, A. (1992). Parametrisation of sea state from SAR images. In *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 29–32 vol.3. 28

Dellinger, F., Delon, J., Gousseau, Y., Michel, J., and Tupin, F. (2015). SAR-SIFT: A SIFT-Like Algorithm for SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1):453–466. 88, 114

Delong, A. and Boykov, Y. (2008). A scalable graph-cut algorithm for n-d grids. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. 54

Dempster, A. P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics*, 38(2):325 – 339. 127

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22. 107

Deng, H. and Clausi, D. (2005). Unsupervised segmentation of synthetic aperture radar sea ice imagery using a novel Markov random field model. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):528–538. 42

Deschênes, F. and Ziou, D. (2000). Detection of line junctions and line terminations using curvilinear features. *Pattern Recognition Letters*, 21(6):637 – 649. 68

Desroches, D., Fjørtoft, R., Massonnet, D., Duro, J., Gaudin, J., and Pourthié, N. (2016). Inland water height estimation without ground control points for near-nadir InSAR data. *IEEE Geoscience and Remote Sensing Letters*, 13:1354–1358. 24, 44

Dillabaugh, C. R., Niemann, K. O., and Richardson, D. E. (2002). Semi-automated extraction of rivers from digital imagery. *GeoInformatica*, 6(3):263–284. 63, 82

Downing, J., Cole, J., Duarte, C., Middelburg, J., Melack, J., Prairie, Y., Kortelainen, P., Striegl, R., McDowell, W., and Tranvik, L. (2012). Global abundance and size distribution of streams and rivers. *Inland Waters*, 2(4):229–236. 49

Eckstein, J. and Bertsekas, D. P. (1992). On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55:293–318. 58

Edmonds, J. and Karp, R. (2003). Theoretical improvement in algorithmic efficiency for network flow problems. *Journal of the ACM*, 19:248–264. 54

FAO, 2020 (2020). The State of Food and Agriculture 2020. Overcoming water challenges in agriculture. Rome. 13, 179

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45(2). 23

Feng, M., Sexton, J. O., Channan, S., and Townshend, J. R. (2016). A global, high-resolution (30-m) inland water body dataset for 2000: first results of a topographic–spectral classification algorithm. *International Journal of Digital Earth*, 9(2):113–133. 48

Ferrentino, E., Nunziata, F., Buono, A., Urciuoli, A., and Migliaccio, M. (2020). Multipolarization time series of sentinel-1 SAR imagery to analyze variations of reservoirs' water body. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:840–846. 163

Ferrentino, E., Nunziata, F., and Migliaccio, M. (2017). Full-polarimetric SAR measurements for coastline extraction and coastal area classification. *International Journal of Remote Sensing*, 38(23):7405–7421. 163

Fischler, M. A., Tenenbaum, J. M., and Wolf, H. C. (1987). Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. In *Readings in Computer Vision*, pages 741–752. Elsevier. 68

Fjortoft, R., Delignon, Y., Pieczynski, W., Sigelle, M., and Tupin, F. (2003). Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 41(3):675–686. 42

Fjortoft, R., Lopes, A., Marthon, P., and Cubero-Castan, E. (1998). An optimal multiedge detector for SAR image segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3):793–802. 43, 88, 114

Fjørtoft, R., Gaudin, J.-M., Pourthie, N., Lion, C., Mallet, A., Souyris, J.-C., Ruiz, C., Koudogbo, F., Duro, J., Ordoqui, P., and Arnaud, A. (2010). KaRIn - the Ka-band radar interferometer on SWOT: Measurement principle, processing and data specificities. In *2010 IEEE International Geoscience and Remote Sensing Symposium*, pages 4823–4826. 34

Fjørtoft, R., Gaudin, J.-M., Pourthié, N., Lalaurie, J.-C., Mallet, A., Nouvel, J.-F., Martinot-Lagarde, J., Oriot, H., Borderies, P., Ruiz, C., and Daniel, S. (2014). KaRIn on SWOT: Characteristics of Near-Nadir Ka-Band Interferometric SAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(4):2172–2185. 24, 35, 37

Ford, L. R. and Fulkerson, D. R. (1956). Maximal flow through a network. *Canadian Journal of Mathematics*, 8:399–404. 54

Gabay, D. and Mercier, B. (1976). A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics With Applications*, 2:17–40. 58

Garello, R., Proust, S., and Chapron, B. (1993). 2D ocean surface SAR images simulation: a statistical approach. In *Proceedings of OCEANS '93*, pages III7–II12 vol.3. 28

Gasnier, N., Dalsasso, E., Denis, L., and Tupin, F. (2021a). Despeckling Sentinel-1 GRD images by deep-learning and application to narrow river segmentation. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 2995–2998. 17, 158, 160, 161, 188

Gasnier, N., Denis, L., Fjørtoft, R., Liège, F., and Tupin, F. (2021b). Narrow river extraction from SAR images using exogenous information. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:5720–5734. 15, 16, 61, 99, 181, 182, 184

Gasnier, N., Denis, L., and Tupin, F. (2021c). Generalized likelihood ratio tests for linear structure detection in SAR images. In *EUSAR 2021; 13th European Conference on Synthetic Aperture Radar*, pages 1–6. 16, 61, 72, 182

Gasnier, N., Denis, L., and Tupin, F. (2021d). On the use and denoising of the temporal geometric mean for SAR time series. *IEEE Geoscience and Remote Sensing Letters*, pages 1–5. 18, 142, 188

Geman, D. and Jedynak, B. (1996). An active testing model for tracking roads in satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):1–14. 68

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741. 42, 52, 53

Glowinski, R. and Marroco, A. (1975). Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 9(R2):41–76. 58

Goldberg, A. V. (2008). The partial augment–relabel algorithm for the maximum flow problem. In Halperin, D. and Mehlhorn, K., editors, *Algorithms - ESA 2008*, pages 466–477, Berlin, Heidelberg. Springer Berlin Heidelberg. 54

Goldberg, A. V. and Tarjan, R. E. (1988). A new approach to the maximum flow problem. *JOURNAL OF THE ACM*, 35(4):921–940. 54

Goldstein, R. M., Zebker, H. A., and Werner, C. L. (1988). Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio Science*, 23(4):713–720. 23

Goodman, J. W. (1976). Some fundamental properties of speckle∗. *J. Opt. Soc. Am.*, 66(11):1145–1150. 24, 25

Grady, L. (2006). Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783. 87

GRDC (2020). GRDC (2020): Major River Basins of the World / Global Runoff Data Centre, GRDC. 2nd, rev. ext. ed. Koblenz, Germany: Federal Institute of Hydrology (BfG). 49

Greig, D., Porteous, B., and Seheult, A. H. (1989). Exact maximum a posteriori estimation for binary images. *Journal of the royal statistical society series b-methodological*, 51:271–279. 42, 53

Gruen, A. and Li, H. (1994). Semiautomatic road extraction by dynamic programming. In Ebner, H., Heipke, C., and Eder, K., editors, *ISPRS Commission III Symposium: Spatial Information from Digital Photogrammetry and Computer Vision*, volume 2357, pages 324 – 332. International Society for Optics and Photonics, SPIE. 63

Hammersley, J. M. and Clifford, P. (1971). Markov fields on finite graphs and lattices. 53

Han, B. and Wu, Y. (2018). River extraction of SAR images via active contours driven by adaptive global fitting energies. *International Journal of Remote Sensing*, 39(20):6995–7013. 43, 66

He, X., Zemel, R., and Carreira-Perpinan, M. (2004). Multiscale conditional random fields for image labeling. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, volume 2, pages II–II. 56

Hellwich, O., Laptev, I., and Mayer, H. (2002). Extraction of linear objects from interferometric SAR data. *International Journal of Remote Sensing*, 23(3):461–475. 68

Hong, S., Jang, H., Kim, N., and Sohn, H.-G. (2015). Water area extraction using RADARSAT SAR imagery combined with Landsat imagery and terrain information. *Sensors (Basel)*, 15(3):6652–6667. 43

Hooke, J. (1984). Changes in river meanders: a review of techniques and results of analyses. *Progress in Physical Geography: Earth and Environment*, 8(4):473–508. 49, 64

Hua Xie, Pierce, L. E., and Ulaby, F. T. (2002). Statistical properties of logarithmically transformed speckle. *IEEE TGRS*. 26, 145

Huang, W., DeVries, B., Huang, C., Lang, M. W., Jones, J. W., Creed, I. F., and Carroll, M. L. (2018). Automated extraction of surface water extent from Sentinel-1 Data. *Remote Sensing*, 10(5). 42

IEA (2021). Hydropower special market report, International Energy Agency, Paris. https://www.iea.org/reports/hydropower-special-market-report. 13

IEEE Std 521-2002 (2020). IEEE Standard Letter Designations for Radar-Frequency Bands. *IEEE Std 521-2019 (Revision of IEEE Std 521-2002)*, pages 1–15. 22

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]*. Cambridge University Press. In Press. 13, 179

Irwin, K., Beaulne, D., Braun, A., and Fotopoulos, G. (2017). Fusion of SAR, optical imagery and airborne LiDAR for surface water detection. *Remote Sensing*, 9(9). 127

Isikdogan, F., Bovik, A. C., and Passalacqua, P. (2017). Surface water mapping by deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(11):4909–4918. 43

JPL-D-61923-rev-B (2018). SWOT mission science requirements document. JPL D-61923 rev B. 45, 98

JPL D-79123 (2014). E. Peral, E. Rodríguez, D. Moeller, M. McAdams, M. Johnson, K. Andreadis, D. Arumugam, B. Williams, SWOT Simulator Quick User Guide, JPL D-79123. 33, 96

Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *INTERNATIONAL JOURNAL OF COMPUTER VISION*, 1(4):321–331. 43, 82

Khalil, M. A. K. and Rasmussen, R. A. (1983). Sources, sinks, and seasonal cycles of atmospheric methane. *Journal of Geophysical Research: Oceans*, 88(C9):5131–5144. 13

Kibret, S., McCartney, M., Lautze, J., Nhamo, L., and Yan, G. (2021). The impact of large and small dams on malaria transmission in four basins in africa. *Scientific Reports*, 11(1):13355. 13, 179

Kittler, J. and Illingworth, J. (1986). Minimum error thresholding. *Pattern Recognition*, 19(1):41–47. 42

Klemenjak, S., Waske, B., Valero, S., and Chanussot, J. (2012). Automatic detection of rivers in high-resolution SAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(5):1364–1372. 43, 66

Kohli, P., Ladicky, L., and Torr, P. H. S. (2008). Robust higher order potentials for enforcing label consistency. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. 56

Kohli, P. and Torr, P. H. S. (2005). Efficiently solving dynamic Markov random fields using graph cuts. *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, 2:922–929 Vol. 2. 125

Kolmogorov, V. and Boykov, Y. (2005). What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 1, pages 564–571 Vol. 1. 116

Kolmogorov, V. and Zabih, R. (2004). What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159. 53, 116

Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M., and Zink, M. (2007). TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11):3317–3341. 23

Lafferty, J. D., McCallum, A., and Pereira, F. C. N. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, page 282–289, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc. 42, 56, 87

Landuyt, L., Van Wesemael, A., Schumann, G. J.-P., Hostache, R., Verhoest, N. E. C., and Van Coillie, F. M. B. (2019). Flood mapping based on synthetic aperture radar: An assessment of established approaches. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2):722–739. 127

Le Hegarat-Mascle, S., Bloch, I., and Vidal-Madjar, D. (1997). Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):1018–1031. 127

Lee, J.-S. (1981). Speckle analysis and smoothing of synthetic aperture radar images. *Computer Graphics and Image Processing*, 17(1):24–32. 42

Lee, J.-S. (1983). Digital image smoothing and the sigma filter. *Computer vision, graphics, and image processing*, 24(2):255–269. 42

Lee, J.-S., Grunes, M., Pottier, E., and Ferro-Famil, L. (2004). Unsupervised terrain classification preserving polarimetric scattering characteristics. *IEEE Transactions on Geoscience and Remote Sensing*, 42(4):722–731. 23

Lee, J.-S. and Pottier, E. (2009). *Polarimetric Radar Imaging: From Basics to Applications*. 23

Lee, Y. T., Rao, S., and Srivastava, N. (2013). A new approach to computing maximum flows using electrical flows. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '13, page 755–764, New York, NY, USA. Association for Computing Machinery. 54, 115

Lehner, B. and Grill, G. (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, 27(15):2171–2186. 49

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wisser, D. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, 9(9):494–502. 49

Lehner, B., Verdin, K., and Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10):93–94. 49

Lermé, N., Le Hégarat-Mascle, S., Malgouyres, F., and Lachaize, M. (2020). Multilayer joint segmentation using MRF and graph cuts. *Journal of Mathematical Imaging and Vision*, 62(6):961–981. 129

Lermé, N., Malgouyres, F., and Létocart, L. (2010). Reducing graphs in graph cut segmentation. In *2010 IEEE International Conference on Image Processing*, pages 3045–3048. 54

Li, H.-C., Krylov, V. A., Fan, P.-Z., Zerubia, J., and Emery, W. J. (2016). Unsupervised learning of generalized gamma mixture model with application in statistical modeling of high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(4):2153–2170. 112

Li, J., Kuang, G., and He, Z. (2011). Mapping water bodies from radarsat imagery. *Geomatica*, 65. 43

Li, Z., Bao, Z., and Suo, Z. (2007). A joint image coregistration, phase noise suppression, and phase unwrapping method based on subspace projection for multibaseline insar systems. *Geoscience and Remote Sensing, IEEE Transactions on*, 45:584 – 591. 23

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., and Thieme, M. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Scientific Data*, 6(1):283. 49, 161

Liu, C., Abergel, R., Gousseau, Y., and Tupin, F. (2020a). LSDSAR, a Markovian a contrario framework for line segment detection in SAR images. *Pattern Recognition*, 98:107034. 43

Liu, C., Tupin, F., and Gousseau, Y. (2020b). Training CNNs on speckled optical dataset for edge detection in SAR images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 170:88–102. 43

Liu, H. and Jezek, K. C. (2004). Automated extraction of coastline from satellite imagery by integrating Canny edge detection and locally adaptive thresholding methods. *International Journal of Remote Sensing*, 25(5):937–958. 42

Liu, J. and Sun, J. (2010). Parallel graph-cuts by adaptive bottom-up merging. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2181–2188. 54

Lobry, S. (2017). *Markovian models for SAR images: Application to water detection in SWOT satellite images and multi-temporal analysis of urban areas.* PhD thesis, Télécom ParisTech. 14, 46, 170, 171

Lobry, S., Denis, L., Williams, B., Fjørtoft, R., and Tupin, F. (2019). Water Detection in SWOT HR Images Based on Multiple Markov Random Fields. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(11):4315–4326. 36, 42, 46, 47, 97, 101

Lobry, S., Tupin, F., and Fjørtoft, R. (2017). Unsupervised detection of thin water surfaces in SWOT images based on segment detection and connection. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3720–3723. 66

Lombardo, P. and Oliver, C. J. (2001). Maximum likelihood approach to the detection of changes between multitemporal SAR images. *IEE Proceedings - Radar, Sonar and Navigation.* 145

Lopes, A., Nezry, E., Touzi, R., and Laur, H. (1990). Maximum a posteriori speckle filtering and first order texture models in SAR images. In *10th Annual International Symposium on Geoscience and Remote Sensing*, pages 2409–2412. 42

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. 111

Martinis, S., Twele, A., and Voigt, S. (2009). Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. *Natural Hazards and Earth System Sciences*, 9(2):303–314. 42

Martinis, S., Twele, A., and Voigt, S. (2011). Unsupervised extraction of flood-induced backscatter changes in SAR data using Markov image modeling on irregular graphs. *IEEE Transactions on Geoscience and Remote Sensing*, 49(1):251–263. 127

Massonnet, D. and Feigl, K. (1998). Radar interferometry and its application to changes in the earth's surface. *Reviews of Geophysics*, 36. 24

Matthews, B. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 405(2):442 – 451. 98

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7):1425–1432. 48, 136, 139

Mekonnen, M. M. and Hoekstra, A. Y. (2016). Four billion people facing severe water scarcity. *Science Advances*, 2(2):e1500323. 13, 179

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7(1):13603. 50

Migliaccio, M., Ferrara, G., Gambardella, A., Nunziata, F., and Sorrentino, A. (2007). A physically consistent speckle model for marine SLC SAR images. *Oceanic Engineering, IEEE Journal of*, 32:839 – 847. 28

Molini, A. B., Valsesia, D., Fracastoro, G., and Magli, E. (2021). Speckle2Void: Deep self-supervised SAR despeckling with blind-spot convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, pages 1–17. 42, 158

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., and Papathanassiou, K. P. (2013). A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, 1(1):6–43. 23, 25

Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. In Watson, G. A., editor, *Numerical Analysis*, pages 105–116, Berlin, Heidelberg. Springer Berlin Heidelberg. 42

Movaghati, S., Moghaddamjoo, A., and Tavakoli, A. (2010). Road extraction from satellite images using particle filtering and extended Kalman filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 48(7):2807–2817. 68

Negri, M., Gamba, P., Lisini, G., and Tupin, F. (2006). Junction-aware extraction and regularization of urban road networks in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2962–2971. 66

Nemni, E., Bullock, J., Belabbes, S., and Bromley, L. (2020). Fully convolutional neural network for rapid flood segmentation in synthetic aperture radar imagery. *Remote Sensing*, 12(16). 43

Nghiem, S., Yueh, S., Kwok, R., and Li, F. (1992). Symmetry properties in polarimetric remote sensing. *Radio Science*, 27. 166

Nicolas, J. and Tupin, F. (2016). Statistical models for SAR amplitude data: A unified vision through Mellin transform and Meijer functions. In *2016 24th European Signal Processing Conference (EUSIPCO)*. 146

Nicolas, J.-M. and Tupin, F. (2020). A New Parameterization for the Rician Distribution. *IEEE Geoscience and Remote Sensing Letters*, 17(11):2011–2015. 31

Nieuwenhuis, G. and Schotten, C. (1992). Land cover monitoring with multi-temporal ERS-1 SAR observations in the Netherlands. In *First ERS-1 Symposium, Cannes, France*. 145

Nunziata, F., Buono, A., Migliaccio, M., and Benassai, G. (2016). Dual-polarimetric C- and X-band SAR data for coastline extraction. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(11):4921–4928. 163, 166, 169

Nunziata, F., Migliaccio, M., and Brown, C. E. (2012). Reflection symmetry for polarimetric observation of man-made metallic targets at sea. *IEEE Journal of Oceanic Engineering*, 37(3):384–394. 166

Oliveira, H. and Correia, P. L. (2013). Automatic road crack detection and characterization. *IEEE Transactions on Intelligent Transportation Systems*, 14(1):155–168. 66

Oliver, C. (1993). Optimum texture estimators for SAR clutter. *Journal of Physics D: Applied Physics*, 26(11):1824. 148

Olver, F. W., Lozier, D. W., Boisvert, R. F., and Clark, C. W. (2010). *NIST Handbook of Mathematical Functions*. Cambridge University Press, USA, 1st edition. 27, 31, 145, 146

Orchard, M. T. and Bouman, C. A. (1991). Color quantization of images. *IEEE Trans. Signal Process.*, 39:2677–2690. 107

Ouled Sghaier, M., Hammami, I., Foucher, S., and Lepage, R. (2018). Flood extent mapping from time-series SAR images based on texture analysis and data fusion. *Remote Sensing*, 10(2). 127

Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633):418–422. 48

Pelizzari, S. and Bioucas-Dias, J. (2007). Oil spill segmentation of SAR images via graph cuts. In *2007 IEEE International Geoscience and Remote Sensing Symposium*, pages 1318–1321. 42

Perciano, T., Tupin, F., Jr., R. H., and Jr., R. M. C. (2016). A two-level Markov random field for road network extraction and its application with optical, SAR, and multitemporal data. *International Journal of Remote Sensing*, 37(16):3584–3610. 66

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639. 42

Peña-Luque, S., Ferrant, S., Cordeiro, M. C. R., Ledauphin, T., Maxant, J., and Martinez, J.-M. (2021). Sentinel-1&2 multitemporal water surface detection accuracies, evaluated at regional and reservoirs level. *Remote Sensing*, 13(16). 127

Quesney, A., Le Hégarat-Mascle, S., Taconet, O., Vidal-Madjar, D., Wigneron, J., Loumagne, C., and Normand, M. (2000). Estimation of watershed soil moisture index from ERS/SAR data. *Remote Sensing of Environment*, 72(3):290–303. 148

Quin, G. (2014). *Etude des séries temporelles en imagerie satellitaire SAR pour la détection automatique de changements*. PhD thesis, ENST, Paris. 147

Quin, G., Pinel-Puysségur, B., Nicolas, J., and Loreaux, P. (2014). MIMOSA: An Automatic Change Detection Method for SAR Time Series. *IEEE TGRS*, 52(9):5349–5363. 145, 148, 156

Rambour, C., Audebert, N., Koeniguer, E., Le Saux, B., Crucianu, M., and Datcu, M. (2020). Flood detection in time series of optical and SAR images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2020:1343–1346. 127

Rignot, E. and Chellappa, R. (1992). Segmentation of polarimetric synthetic aperture radar data. *IEEE Transactions on Image Processing*, 1(3):281–300. 42

Rogers, A. S. and Kearney, M. S. (2004). Reducing signature variability in unmixing coastal marsh thematic mapper scenes using spectral indices. *International Journal of Remote Sensing*, 25(12):2317–2335. 136

Rossant, F., Badellino, M., Chavillon, A., Bloch, I., and Pâques, M. (2011). A morphological approach for vessel segmentation in eye fundus images, with quantitative evaluation. *Journal of Medical Imaging and Health Informatics*, 1:42–49. 66

Rother, C., Kolmogorov, V., and Blake, A. (2004). "GrabCut": Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.*, 23(3):309–314. 17, 106, 107, 170, 172, 185

Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268. 57

Ruzon, M. and Tomasi, C. (2000). Alpha estimation in natural images. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, volume 1, pages 18–25 vol.1. 107

S1-RS-MDA-52-7440 (2016). Sentinel-1 product definition, S1-RS-MDA-52-7440, MacDonald, Dettwiler and Associates Ltd. Bourbigot, Matthieu and Johnsen, Harald and Piantanida, Riccardo and Hajduch, Guillaume. 38

Sghaier, M. O., Foucher, S., and Lepage, R. (2017). River extraction from high-resolution SAR images combining a structural feature set and mathematical morphology. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(3):1025–1038. 43, 66

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press. 127

Sheng, Y., Song, C., Wang, J., Lyons, E. A., Knox, B. R., Cox, J. S., and Gao, F. (2016). Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. *Remote Sensing of Environment*, 185:129–141. Landsat 8 Science Results. 50

Silveira, M. and Heleno, S. (2009). Separation between water and land in SAR images using region-based level sets. *IEEE Geoscience and Remote Sensing Letters*, 6(3):471–475. 42

Skriver, H., Mattia, F., Satalino, G., Balenzano, A., Pauwels, V. R. N., Verhoest, N. E. C., and Davidson, M. (2011). Crop classification using short-revisit multitemporal sar data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(2):423–431. 127

Sohn, H.-G. and Jezek, K. C. (1999). Mapping ice sheet margins from ERS-1 SAR and SPOT imagery. *International Journal of Remote Sensing*, 20(15-16):3201–3216. 42

Sun, J., Wang, X., Yuan, X., Zhang, Q., Guan, C., and Babanin, A. V. (2018). The dependence of sea SAR image distribution parameters on surface wave characteristics. *Remote Sensing*, 10(11). 30

Sun, Z., Geng, H., Lu, Z., Scherer, R., and Woźniak, M. (2021). Review of road segmentation for SAR images. *Remote Sensing*, 13(5). 66

Talbot, J., Young, B., and Xu, X. (2004). Implementing GrabCut. 107

Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., Erickson, T. A., and Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. *Nature*, 596(7870):80–86. 13, 179

Tison, C. et al. (2004). A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(10):2046–2057. 31

Tupin, F., Houshmand, B., and Datcu, M. (2002). Road detection in dense urban areas using SAR imagery and the usefulness of multiple views. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11):2405–2414. 66

Tupin, F., Maitre, H., Mangin, J.-F., Nicolas, J.-M., and Pechersky, E. (1998). Detection of linear features in SAR images: application to road network extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 36(2):434–453. 16, 66, 68, 74, 81, 182

Uddin, K., Matin, M. A., and Meyer, F. J. (2019). Operational flood mapping using multi-temporal Sentinel-1 SAR images: A case study from bangladesh. *Remote Sensing*, 11(13). 127

UN-Water (2021). Progress on integrated water resources management. Technical report, United Nations Water. 13, 179

Valenzuela, G. R. (1978). Theories for the interaction of electromagnetic and oceanic waves —a review. *Boundary-Layer Meteorology*, 13(1):61–85. 28

Valero, S., Chanussot, J., Benediktsson, J., Talbot, H., and Waske, B. (2010). Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images. *Pattern Recognition Letters*, 31(10):1120 – 1127. Pattern Recognition in Remote Sensing. 43, 66

Van Trees, H. L. (2004). *Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory*. John Wiley & Sons. 69

Van Tricht, K., Gobin, A., Gilliams, S., and Piccard, I. (2018). Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: A case study for belgium. *Remote Sensing*, 10(10). 127

van Zyl, J. (1990). Calibration of polarimetric radar images using only image parameters and trihedral corner reflector responses. *IEEE Transactions on Geoscience and Remote Sensing*, 28(3):337–348. 166

Vanderbrug, G. J. (1976). Line detection in satellite imagery. *IEEE Transactions on Geoscience Electronics*, 14(1):37–44. 68

Venkatakrishnan, S. V., Bouman, C. A., and Wohlberg, B. (2013). Plug-and-play priors for model based reconstruction. In *2013 IEEE Global Conference on Signal and Information Processing*, pages 945–948. 59

Verpoorter, C., Kutser, T., Seekell, D. A., and Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, 41(18):6396–6402. 13, 179

Wang, X., Xie, S., Zhang, X., Chen, C., Guo, H., Du, J., and Duan, Z. (2018). A robust multi-band water index (MBWI) for automated extraction of surface water from Landsat 8 OLI imagery. *International Journal of Applied Earth Observation and Geoinformation*, 68:73–91. 136

Waske, B. and Braun, M. (2009). Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(5):450–457. Theme Issue: Mapping with SAR: Techniques and Applications. 127

Weinberg, G. and Tran, C. (2018). Burr distribution for X-band maritime surveillance radar clutter. *Progress In Electromagnetics Research B*, 81:183–201. 28, 30

Whittemore, A., Ross, M. R. V., Dolan, W., Langhorst, T., Yang, X., Pawar, S., Jorissen, M., Lawton, E., Januchowski-Hartley, S., and Pavelsky, T. (2020). A participatory science approach to expanding instream infrastructure inventories. *Earth's Future*, 8(11):e2020EF001558. e2020EF001558 2020EF001558. 49

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., Vogelmann, J., Wulder, M. A., and Wynne, R. (2008). Free access to Landsat imagery. *Science*, 320(5879):1011–1011. 47

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., and Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122:2–10. Landsat Legacy Special Issue. 47

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14):3025–3033. 48, 136

Xu, L. and Jordan, M. I. (1996). On Convergence Properties of the EM Algorithm for Gaussian Mixtures. *Neural Computation*, 8(1):129–151. 107

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M. (2019). Merit hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resources Research*, 55(6):5053–5073. 49

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., and Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 44(11):5844–5853. 49

Yamazaki, D., O'Loughlin, F., Trigg, M. A., Miller, Z. F., Pavelsky, T. M., and Bates, P. D. (2014). Development of the global width database for large rivers. *Water Resources Research*, 50(4):3467–3480. 49

Yamazaki, D., Trigg, M. A., and Ikeshima, D. (2015). Development of a global 90m water body map using multi-temporal Landsat images. *Remote Sensing of Environment*, 171:337–351. 48

Yang, X., Qin, Q., Yésou, H., Ledauphin, T., Koehl, M., Grussenmeyer, P., and Zhu, Z. (2020). Monthly estimation of the surface water extent in france at a 10-m resolution using sentinel-2 data. *Remote Sensing of Environment*, 244:111803. 136

Yim, P. (2015). Method and system for image segmentation. US Patent. US9.214,029B2. 54, 115

Yoruk, U., Hargreaves, B. A., and Vasanawala, S. S. (2018). Automatic renal segmentation for MR urography using 3D-GrabCut and random forests. *Magnetic Resonance in Medicine*, 79(3):1696–1707. 129

Zhao, W., Deledalle, C., Denis, L., Maître, H., Nicolas, J., and Tupin, F. (2018). RABASAR: A Fast Ratio Based Multi-Temporal SAR Despeckling. In *IGARSS 2018*, pages 4197–4200. 142, 143, 145

Zhao, W., Deledalle, C., Denis, L., Maître, H., Nicolas, J., and Tupin, F. (2019). Ratio-Based Multitemporal SAR Images Denoising: RABASAR. *IEEE TGRS*, 57(6):3552–3565. 17, 42, 143, 156, 188

Zieger, A., Hancock, D., Hayne, G., and Purdy, C. (1991). NASA radar altimeter for the TOPEX/POSEIDON Project. *Proceedings of the IEEE*, 79(6):810–826. 33

ECOLE DOCTORALE

Titre : Exploitation de données multi-temporelles et multi-capteurs pour l'extraction de surfaces d'eau continentales dans le contexte de la mission SWOT

Mots clés : SWOT, SAR, Télédétection, Extraction de surfaces d'eau, MRF

Résumé : La télédétection spatiale fournit aux hydrologues et aux décideurs des données indispensables à la compréhension du cycle de l'eau et à la gestion des ressources et risques associés. Le satellite SWOT, qui est une collaboration entre les agences spatiales françaises (CNES) et américaine (NASA, JPL), et dont le lancement est prévu en 2022 vise à mesurer la hauteur des lacs, rivières et océans avec une grande résolution spatiale. Il complétera ainsi les capteurs existants, comme les constellations SAR et optique Sentinel-1 et 2 et les relevés in situ. SWOT représente une rupture technologique car il est le premier satellite qui embarque un altimètre de fauchée quasi-nadir. Le calcul des hauteurs d'eau est fait par interférométrie sur les images SAR acquises par SWOT. La détection d'eau dans ces images est donc une étape essentielle du traitement des données SWOT, mais qui peut être difficile, en particulier avec un faible rapport signal sur bruit ou en présence de radiométries inhabituelles. Dans cette thèse, nous cherchons à développer de nouvelles méthodes pour rendre la détection d'eau plus robustes. Pour cela, nous nous intéressons à l'utilisation de données exogènes pour guider la détection, à la combinaison de données multi-temporelles et multi-capteurs et à des approches de débruitage. La première méthode proposée exploite les informations de la base de donnée des rivières utilisée par SWOT pour détecter les rivières fines dans l'image de façon robuste à la fois aux bruit dans l'image, aux erreurs éventuelles de la base de données et aux changements survenus. Cette méthode s'appuie sur un nouveau détecteur de structures linéiques, un algorithme de chemin de moindre coût et une nouvelle méthode de segmentation par CRF qui combine des termes d'attache aux données et de régularisation adaptés au problème. Nous avons également proposé une méthode dérivée des GrabCut qui utilise un polygone a priori contenant un lac pour le détecter sur une image SAR ou une série temporelle. Dans ce cadre, nous avons également étudié le recours à une combinaison multi-temporelle et multi-capteurs (optique et SAR). Enfin, dans le cadre d'une étude préliminaire sur les méthodes de débruitage pour la détection d'eau nous avons étudié les propriétés statistiques de la moyenne géométrique temporelle et proposé une adaptation de la méthode variationelle MuLoG pour la débruiter.

Title : Use of multi-temporal and multi-sensor data for continental water body extraction in the context of the SWOT mission

Keywords : SWOT, SAR, Remote Sensing, Water body extraction, MRF

Abstract : Spaceborne remote sensing provides hydrologists and decision-makers with data that are essential for understanding the water cycle and managing the associated resources and risks. The SWOT satellite, which is a collaboration between the French (CNES) and American (NASA, JPL) space agencies, is scheduled for launch in 2022 and will measure the height of lakes, rivers, and oceans with high spatial resolution. It will complement existing sensors, such as the SAR and optical constellations Sentinel-1 and 2, and in situ measurements. SWOT represents a technological breakthrough as it is the first satellite to carry a near-nadir swath altimeter. The estimation of water levels is done by interferometry on the SAR images acquired by SWOT. Detecting water in these images is therefore an essential step in processing SWOT data, but it can be very difficult, especially with low signal-to-noise ratios, or in the presence of unusual radiometries. In this thesis, we seek to develop new methods to make water detection more robust. To this end, we focus on the use of exogenous data to guide detection, the combination of multi-temporal and multi-sensor data and denoising approaches. The first proposed method exploits information from the river database used by SWOT to detect narrow rivers in the image in a way that is robust to both noise in the image, potential errors in the database, and temporal changes. This method relies on a new linear structure detector, a least-cost path algorithm, and a new CRF segmentation method that combines data attachment and regularization terms adapted to the problem. We also proposed a method derived from GrabCut that uses an a priori polygon containing a lake to detect it on a SAR image or a time series of SAR images. Within this framework, we also studied the use of a multi-temporal and multi-sensor combination (optical and SAR). Finally, as part of a preliminary study on denoising methods applied to water detection, we studied the statistical properties of the geometric temporal mean and proposed an adaptation of the variational method MuLoG to denoise it.

Institut Polytechnique de Paris
91120 Palaiseau, France