



EDITE - ED 130

Doctorat ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Signal et images »

présentée et soutenue publiquement par

Sylvain Lobry

soutenance prévue le 16 novembre 2017

Modèles Markoviens pour les images SAR : Application à la détection de l'eau dans les images satellitaires SWOT et analyse multi-temporelle de zones urbaines

Directeurs de thèse : **Florence Tupin et Roger Fjørtoft**

Co-encadrement de la thèse : **Loïc Denis**

Jury

M. Ronan Fablet, Professeur, LabSTICC, IMT Atlantique

M. Jean-François Giovannelli, Professeur, IMS Lab, Université de Bordeaux

M. Eric Pottier, Professeur, IETR, Université de Rennes 1

Mme Marie Chabert, Professeur, INP - ENSEEIHT Toulouse

Mme Gilda Schirinzi, Professeur, Dipartimento di Ingegneria, Université de Naples Parthénope

M. Loïc Denis, Maître de conférence, Laboratoire Hubert Curien, Université de Saint-Etienne

M. Brent Williams, Ingénieur de recherche, NASA's Jet Propulsion Laboratory

M. Roger Fjørtoft, Ingénieur de recherche, CNES

Mme Florence Tupin, Professeur, LTCI, Télécom ParisTech

Rapporteur

Rapporteur

Président

Examinateuse

Examinateuse

Examinateur

Invité

Directeur de thèse

Directrice de thèse

T
H
È
S
E

TELECOM ParisTech

école de l'Institut Mines-Télécom - membre de ParisTech

Remerciements

Je souhaite tout d'abord remercier le Centre National d'Études Spatiales (CNES) ainsi que le programme Futur & Ruptures de l'institut Mines-Télécom (IMT) de m'avoir donné l'opportunité d'effectuer cette thèse. Les programmes d'accompagnements (notamment les journées Jeunes Chercheurs pour le CNES et les journées de restitution des travaux côté Futur & Ruptures) m'ont permis de découvrir de nombreux travaux, de présenter les miens et surtout ont été l'occasion de belles rencontres.

Je tiens à remercier M. Eric Pottier de m'avoir fait l'honneur de présider mon Jury et M. Ronan Fablet ainsi que M. Jean-François Giovannelli d'avoir rapporté (et donc d'avoir contribué à ma présentation de soutenance et à ce document). I also want to thank Mme Marie Chabert and Mme Gilda Schirinzi who accepted to examine my PhD and M. Brent Williams who accepted our invitation. Thank to all of you, the defense that I was dreading turned out to be really interesting and an unforgettable moment.

Bien sûr, je souhaite tout particulièrement remercier mes encadrants : Florence Tupin, qui en plus de l'encadrement quotidien et des conseils scientifiques, m'a donné l'envie de poursuivre dans la recherche académique ; Roger Fjørtoft qui m'a investi sur un projet de grande envergure tout en me laissant une liberté dans les axes de recherche explorés, m'a accueilli toujours chaleureusement au CNES et avec qui c'est un réel plaisir de travailler ; enfin Loïc Denis pour ses nombreuses idées, les expériences dans un Saint-Étienne/Paris à 7h du matin et les répétitions dans un hall d'hôtel du Texas à 23h. Sur des points différents, vous êtes tous trois une source d'inspiration. Je n'aurais pu souhaiter meilleur encadrement, un grand merci à vous.

Durant ma thèse, j'ai effectué un monitorat à l'UPMC. J'ai eu la chance de rencontrer de nombreux enseignant-chercheurs, ATER et moniteurs dans des UE bien organisées ; merci à eux.

Merci aux membres du LRDE de m'avoir accueilli dans leur équipe et m'avoir initié à la recherche. En particulier, je souhaite remercier Théo qui m'a en plus suivi pendant la thèse et dont les encouragements sont toujours précieux. Christelle et Vincent, pour m'avoir fait confiance et pour tout ce que vous m'avez appris.

Lors de mes séjours au CNES, j'ai trouvé une super ambiance. La journée, j'ai beaucoup appris, autant d'un point de vue scientifique que technique. Les discussions politiques autour d'un café, théorie des jeux et sondages lorsque c'est la fin des haricots, un blind-test contrepèteries et super-héros ou plus classiquement les bières du dispensary resteront de très bon souvenirs. J'espère recroiser votre route bientôt. Les réunions de l'ADT m'ont permis de rencontrer de nombreux scientifiques impliqués sur la mission ; je souhaite les remercier pour les discussions qui ont permis d'enrichir ce travail.

La majeur partie de ma thèse se déroulait dans l'équipe IMAGES du département IDS dans le laboratoire LTCI à Télécom ParisTech. Je tiens à remercier tous les permanents pour leur disponibilité et les échanges toujours fructueux. En particulier Jean-Marie pour les nombreuses discussions ainsi que pour l'aide aussi bien théorique que technique que tu m'as apportée; les radaristes te doivent beaucoup! Henri et Michel, merci pour les répétitions, les questions en SAR meeting, votre aide m'a été précieuse. Au niveau du département, merci à Salimatou, Marie-Laure et Laurence pour leur aide importante. J'ai eu le plaisir d'assister au conseil de département pendant deux années, aussi j'aimerais remercier ses membres, en particulier Gaël. J'ai aussi une pensée pour Caroline. Aussi, je souhaite remercier Florence et Mariana pour leur disponibilité.

J'ai eu la chance de croiser de nombreux collègues doctorants largement responsables de la bonne humeur qui règne dans cette équipe. Je me souviens de celles qui m'ont accueilli dans leur bureau; bureau où j'ai eu la chance de trouver des collègues aux compétences variées, avec de nombreuses discussions qui m'ont permis d'avancer très vite. Il y avait notre maman à tous, toujours prête à nous soutenir et à nous aider. J'espère avoir une bonne évaluation dans ton fichier. Il y a celui que j'ai vu (presque) au quotidien pendant toute ma thèse, avec qui j'ai apprécié discuter. J'attends toujours que tu m'apprennes à jouer à LoL cependant. Celle avec qui je peux discuter de tout et souvent en hauteur (dans un château, en haut d'un mur ou essoufflé sur une colline). Tu m'as beaucoup appris et fait évoluer (en bien j'espère). Celui qui est le boss du "crew". J'espère te retrouver pour un nouveau petit-déjeuner vegan, bio, sans gluten dans un quartier hipster de Londres. Celle dont j'aurais souhaité qu'elle soutienne sa thèse un peu plus tard afin d'avoir une autre collègue parisienne qui a des goûts cools.

Plus tard, j'ai rencontré celui avec qui j'ai toujours d'intéressantes discussions, que ce soit sur le boulot ou autre et que ce soit au boulot ou dans de belles soirées. Tu m'as aidé à passer une bonne fin de thèse même si tu as bien trop élagué ma playlist. Celle qui, d'abord vue à EUSAR, a égayé un bureau de 3^e année. Celui qui m'invite dans des salons où on se sent pas très bien à 14h. Celui qui, jamais avare de questions, a dynamisé le groupe. Celui qui, bien que discret, est quelqu'un de talentueux avec un humour que je ne comprends pas (donc sûrement drôle).

Plus loin, Celui qui donne TPLM, qui quand il rentre sur le dancefloor TLMSMAL et qui aime bien les frites parce que c'est très très bon; enfin bref TMTC. Plus sérieusement celui qui gère au boulot et qui en plus a un humour de type drôle. Celle que j'ai vu (cette fois vraiment) au quotidien pendant toute ma thèse et qui même sans mojito ni parapluie, est toujours dynamique, prête à aider et de bonne humeur. Celui avec qui on a couru un marathon (et 2-3 autres courses au passage), le roi des gâteaux et du karaoké et surtout un mec (trop) gentil. Celui qui n'a pas été très efficace pour m'apprendre l'espagnol et qui se met à parler français seulement à la fin de sa thèse. Du coup tu comprendras : bonne chance pour ta soutenance! Celui que je vois plus en Chine ou aux USA qu'à Paris, avec qui je garde de très bons souvenirs de ces voyages. Celui qui

a dû supporter 2-3 fois une mauvaise acceptation de ma part d'un certain événement de 2006, et qui malgré cela ne me déteste pas (j'espère). Celui dont la connaissance en optimisation, la capacité à apprendre et la gentillesse m'impressionne. Merci pour les nombreuses fois où tu m'as aidé. Celui dont la bonne humeur a accompagné une conférence à Juan-Les-Pins. Celui qui aime faire des pauses où il inverse les lettres des claviers et d'un enthousiasme rare. Celui qui dit venir des ténèbres, mais qui est bien trop sympa pour ça. Celui qui m'a offert un médiator au hasard d'un concert. Celle qui pose avec une bouteille de rouge. Celui qui, discret, aime toujours rire. J'espère que tu te plais en Suisse. Celle qui a partagé avec moi les joies du rendu d'un manuscrit de thèse et de la préparation d'une soutenance. Celui, dont j'ai encore l'accent du sud-ouest dans les oreilles. Celui qui a une manière... Anglaise de débattre. Ceux impliqués dans un conseil des doctorants dynamique. Sorti des murs du labo, ceux rencontrés en conférence. Notamment celle dont le soutien en fin de thèse a été important et que j'espère recroiser, dans un bus à Toulouse ou à Melbourne.

Tous les autres aussi.

Bref, j'ai rencontré ceux qui font que je garderai un très bon souvenir de cette période de ma vie.

Enfin, je tiens à remercier ma famille qui m'a toujours encouragé et supporté dans mes études. J'ai de la chance de vous avoir. Je souhaite aussi remercier ma compagne qui me supporte au quotidien (et ce n'est pas facile!) et sa famille avec qui j'aime beaucoup passer du temps. Enfin, je dédie ce travail à ma grand-mère, Simone, à qui je dois beaucoup.

Contents

Synthèse des travaux	VII
1 Introduction	1
1.1 Context	1
1.2 Contributions	2
1.3 Organization of this manuscript	5
I Image models	7
2 Physics and statistics of SAR images	9
2.1 Introduction	9
2.2 Acquisition	10
2.3 Statistics of SAR data	12
2.4 Particularities of SWOT	17
3 Image models	21
3.1 Bayesian modeling	21
3.2 Markov Random Fields	22
3.2.1 Notation and definitions	22
3.2.2 Optimization	24
II Water/land classification	31
4 Water/land classification	35
4.1 Introduction	35
4.2 Binary classification	35
4.3 Markovian approaches	42
4.4 Estimation and spatial regularization of class parameters	45
4.4.1 Constant parameters	47
4.4.2 Non-constant parameters with no re-estimation	48
4.4.3 Region-based parameters estimation	52
4.4.4 MRF estimation of the parameter images	56
4.5 Results	59

4.5.1	Results on Garonne	62
4.5.2	Results on Po	63
4.5.3	Results on Camargue	64
4.5.4	Results on Kaw (SETHI)	67
4.6	Conclusion	68
5	Detection of narrow rivers	71
5.1	Introduction	71
5.2	Two-step approach	72
5.2.1	Low-level step	72
5.2.2	Connection and regularization	75
5.3	Results	81
5.4	Conclusion	84
III	Processing of multi-temporal series of SAR images	87
6	Decomposition models for multi-temporal series of SAR images	89
6.1	Introduction	89
6.2	Decomposition models	93
6.2.1	A general framework	93
6.2.2	Interpretation of the decomposition as a detection problem	95
6.2.3	Reformulation as an estimation problem	99
6.3	Optimization	101
6.3.1	Exact optimization using graph-cuts	101
6.3.2	Efficient optimization	105
6.4	Parameters computation	107
6.5	Conclusion	110
7	Applications	113
7.1	Strong scatterers detection	113
7.2	Image regularization	118
7.3	Change detection	126
7.4	Conclusion	136
Conclusion and perspectives		137
Acronyms		143

Synthèse des travaux

Introduction

Contexte L'eau est omniprésente sur terre. D'après (Gleick, 1993), 71% de la surface terrestre est couverte d'eau. Seulement, 97,5% de celle-ci est salée. La plupart des réserves d'eau douce sont contenues dans des glaciers ou sont souterraines, les lacs ne comptant que pour 0,26%. Alors que cette ressource n'est présente qu'en quantité limitée, la demande d'eau a crû deux fois plus rapidement que la population sur le siècle dernier d'après l'OCDE¹. Les conséquences sont dramatiques: d'après l'OMS, 3,4 millions de personnes meurent chaque année de maladies liées à l'eau² et d'ici 2025, l'ONU estime que deux tiers de la population mondiale pourrait avoir des difficultés pour accéder à l'eau³. Il apparaît alors primordial de surveiller ces ressources. Pour cela, les hydrologues utilisent principalement des données acquises sur place, dont la couverture spatiale est limitée par un coût élevé.

Afin d'obtenir des mesures globales et régulières, les scientifiques ont commencé à utiliser des données satellitaires. On peut par exemple citer les missions Jason qui utilisent l'altimétrie radar afin de mesurer la topographie des surfaces océaniques. En hydrologie, la mission GRACE a été utilisée pour mesurer les changements dans les réserves d'eau (Ramillien *et al.*, 2008).

La future mission SWOT (Surface Water and Ocean Topography) a pour objectif de faire la première mesure globale des hauteurs d'eau de surface. Dans la lignée de la collaboration entre le CNES et le laboratoire JPL de la NASA, cette mission est aussi menée en collaboration avec l'agence spatiale Canadienne (CSA) et celle du Royaume-Uni (UKSA). Le lancement du satellite est prévu pour Avril 2021. Afin de mener à bien l'objectif de mesurer les variations de hauteur d'eau, le radar à synthèse d'ouverture (SAR) est particulièrement adapté: en plus de sa capacité à acquérir des données quelles que soient les conditions climatiques, sa stabilité radiométrique permet d'avoir des mesures comparables dans le temps et les capacités interférométriques permettent l'obtention d'une mesure de hauteur. Afin d'obtenir ces mesures, une étape cruciale est la détection de l'eau. Dans le cadre de la collaboration, une distribution précise des tâches a été effectuée, et le CNES est en charge de cette partie. Le développement d'algorithmes pour la détection d'eau sur les images continentales de

¹voir <http://www.oecd.org/publications/factbook/34416097.pdf>

²see http://www.who.int/water_sanitation_health/takingcharge.html

³see <http://www.un.org/waterforlifedecade/scarcity.shtml>

SWOT est l'objectif principal du travail proposé. Certaines des méthodes présentées dans ce document ont été implémentées dans la chaîne de traitement du CNES pendant le déroulement de cette thèse.

Nous avons vu que l'un des objectifs de SWOT est de détecter les changements, ce qui implique l'utilisation de données acquises à différentes dates. De telles données n'étant pas disponibles, nous avons utilisé des données SAR issues d'autres capteurs avec le même type de méthodes mais pour d'autres applications. Nous montrons dans la suite que certaines parties de ce travail sont néanmoins également utiles dans le cadre de SWOT.

Contributions Les contributions de ce manuscrit peuvent être divisées en 3 catégories:

- des méthodes dédiées pour la détection de l'eau dans les images SWOT;
- des méthodes de traitement de séries temporelles d'images SAR urbaines;
- des méthodes génériques pour le traitement de grandes quantités de données et/ou de séries temporelles.

Alors que la première et la deuxième catégories visent des applications spécifiques, la dernière comprend les techniques développées pour le traitement de séries temporelles qui peuvent être utilisées dans le cadre de SWOT.

Les objets d'intérêt en hydrologie peuvent être classifiés en se basant sur leur forme générale: soit cet objet est de taille importante et compact (lac), soit il est d'aspect linéaire (rivière). Ainsi, deux méthodes ont été développées pour détecter ces deux types d'objets:

Contribution (1): Méthodes pour la détection de grands plan d'eau

À cause du diagramme d'antenne, des variations dans la rugosité et du faible rapport signal sur bruit (SNR) de la terre dans les images SWOT, les paramètres de la classe eau et ceux de la classe terre ne peuvent être considérés constants à travers l'image. Pour prendre en compte ces variations, trois méthodes d'estimation des paramètres sont proposées:

- une estimation se basant sur une connaissance a priori du X-factor (obtenu à partir d'un modèle numérique de terrain, la forme attendue du diagramme d'antenne et le SNR);
- une estimation itérative basée sur un partitionnement de l'image;
- une estimation dense itérative utilisant un champ de Markov (MRF) sur les cartes de paramètres.

La classification est obtenue conjointement avec l'estimation des paramètres en utilisant un modèle d'Ising optimisé exactement.

Contribution (2): Détection des rivières fines

La méthode proposée pour la détection de rivières fines dans les images SWOT est une approche en deux temps:

1. une détection au niveau du pixel de segments de rivière;
2. la connexion de ces segments et leur sélection d'après des propriétés globales de réseaux de rivières.

La détection s'inspire des travaux de ([Tupin et al., 1998](#); [Cao et al., 2011](#)) et les connexions se font en utilisant l'algorithme de Dijkstra. Enfin, la sélection des connexions utilise un MRF afin d'introduire des a priori de forme sur l'ensemble du réseau.

Une des principales difficultés dans l'interprétation d'images SAR est la présence de speckle pouvant être interprété comme un bruit multiplicatif. Bien que de nombreuses techniques de filtrage de ce speckle aient été développées, la plupart n'arrivent pas à prendre en compte la présence de rétro-diffuseurs forts, typiquement présents en zone urbaine. Dans cette optique, ([Denis et al., 2010](#)) présente une technique de décomposition qui modélise l'image comme une somme de deux composantes: un fond (spatialement régulier) et les rétro-diffuseurs forts. Dans ce manuscrit on propose une extension de ce modèle aux séries temporelles avec des applications pour la détection de cibles, de changement et la régularisation d'images.

Contribution (3): Modèles de décomposition pour les séries temporelles d'images SAR

On présente un cadre général pour décomposer une série temporelle en deux composantes: un ou plusieurs fonds et une composante de cibles fortes pour chaque image dans la série. Cette formulation permet de mettre des a priori différents sur chaque composante, donnant trois modèles:

- TVR: autant d'images de fond et de cibles que d'images en entrée. Modèle adapté pour la régularisation et la détection de cibles fortes;
- TV1BG: un fond pour la totalité de la série, et autant de composantes de cibles que d'images en entrée. Ce modèle est adapté pour la détection de cibles;
- TV1C: un fond pour la totalité de la série, et autant de composantes de cibles que d'images en entrée, avec une contrainte supplémentaire sur les cibles afin de détecter les changements.

Un des avantages de la formulation proposée est que ces modèles peuvent être optimisés exactement. Une étude présentée dans cette thèse montre aussi que relâcher la pseudo-norme L0 vers L1 dégrade les résultats dans ce contexte.

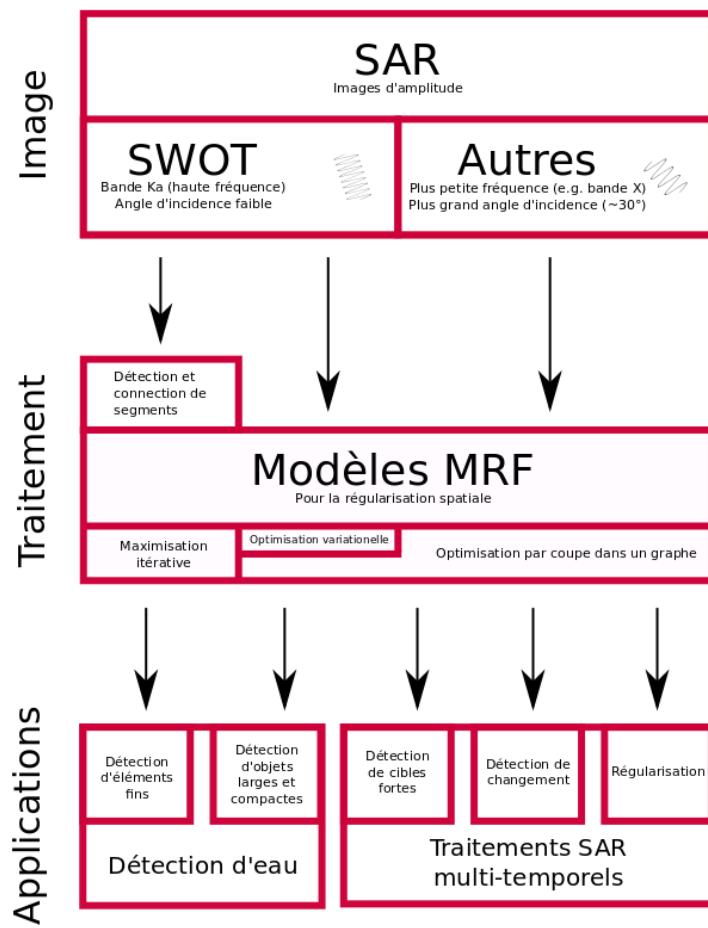


Figure 1 – Vue schématique des contributions de cette thèse.

Certaines méthodes développées pour le traitement des séries temporelles peuvent être utilisées dans le cadre de SWOT lorsque des données temporelles seront disponibles. Par ailleurs, la contribution suivante peut être directement utilisée:

Contribution (4): Optimisation efficace des MRF

Les techniques d'optimisation par coupe dans un graphe peuvent être utilisées pour obtenir le minimum global de certaines énergies mais demandent généralement une grande quantité de mémoire, ce qui peut être un obstacle à leur utilisation sur des applications réelles. Ainsi, une méthode simple d'optimisation par blocs permettant un compromis entre qualité du résultat et quantité de mémoire est présentée. Dans le cas de SWOT, cette technique pourrait être utilisée pour la détection d'eau dans des grandes images ou des séries temporelles.

Enfin, on montre dans la Figure 1 un diagramme contextualisant les différentes contributions présentées dans ce manuscrit.

Classification eau/terre

En ce qui concerne les applications en hydrologie, un des objectifs principaux de SWOT est d'estimer les hauteurs d'eau. Pour ce faire, le satellite utilise l'interférométrie SAR. Cependant, pour appliquer les traitements interférométriques il est nécessaire de savoir où se trouve l'eau. Ainsi, l'objectif principal de cette section est de détecter l'eau à partir d'une image d'amplitude acquise par SWOT et notée \mathbf{v} . En d'autres termes, on cherche une carte \mathbf{u} telle que:

$$u_i = \begin{cases} 1 & \text{si } i \text{ est un pixel d'eau,} \\ 0 & \text{si } i \text{ est un pixel de terre.} \end{cases}$$

La classification binaire est un problème classique (dont on présente un état-de l'art dans la [section 4.2](#)), mais dans le cas de SWOT deux spécificités doivent être prises en compte:

- Les paramètres de classes ne sont pas constants à travers l'image. Avoir des paramètres constants est souvent une hypothèse des algorithmes de classification. Dans ce document est présenté un algorithme pour conjointement estimer les paramètres de classes et réaliser la classification.
- La plupart des méthodes de classification intègrent un a priori de régularité spatiale. Cependant, dans le cas de l'hydrologie, une rivière ne répond généralement pas au critère de compacité spatiale qui convient pour un lac. Ainsi, nous proposons un algorithme dédié à la détection de rivières fines.

Deux approches distinctes ont donc été étudiées.

Classification binaire avec a priori de régularité spatiale pour des classes dont les paramètres ne sont pas constants

Aperçu de l'état de l'art Dans un cadre Bayesien, la solution la plus simple pour trouver une classification \mathbf{u} à partir d'une observation \mathbf{v} est d'utiliser le maximum de vraisemblance:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} p(\mathbf{v}|\mathbf{u}). \quad (1)$$

La vraisemblance, dans le cas d'une image d'amplitude SAR, suit une loi de Rayleigh-Nakagami de paramètres μ et L :

$$p(v|\mu) = \frac{2\sqrt{L}}{\Gamma(L)\mu} \left(\frac{v\sqrt{L}}{\mu} \right)^{2L-1} e^{-\left(\frac{v\sqrt{L}}{\mu} \right)^2}. \quad (2)$$

Toujours dans un cadre Bayesien, il est possible d'intégrer des a priori en utilisant

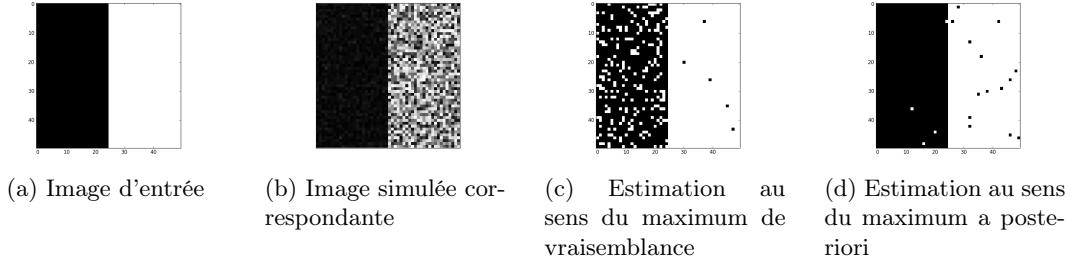


Figure 2 – Exemple sur une image SAR simple simulée. Les deux classes suivent une loi de Rayleigh-Nakagami de paramètres $\mu_0 = 10$ et $\mu_1 = 20$, $L = 1$.

le théorème de Bayes:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} \frac{p(\mathbf{v}|\mathbf{u})p(\mathbf{u})}{p(\mathbf{v})} \quad (3)$$

$$= \arg \max_{\mathbf{u}} p(\mathbf{v}|\mathbf{u})p(\mathbf{u}) \quad (4)$$

On parle alors de la solution au sens du maximum a posteriori (MAP). Cet a priori peut-être statistique (proportion relative des classes) ou spatial par exemple. Lorsque l'on souhaite intégrer des a priori spatiaux, une approche classique est d'utiliser un champ de Markov, ce qui sera la solution proposée dans cette partie. En effet, que l'on utilise l'estimateur au sens du maximum de vraisemblance ou du maximum a posteriori avec des a priori définis au niveau du pixel, on peut aisément montrer que cela revient à appliquer un seuil. Or, nous avons vu que les images SAR sont sujettes au speckle, qui introduit de fortes variations. Une approche purement pixellique donnera donc un résultat bruité comme on le montre dans la Figure 2.

D'autres solutions existent pour intégrer des a priori spatiaux: on peut par exemple citer des approches débruitant l'image dans un premier temps (apportant donc une régularité spatiale) puis faisant la classification dans un second ([Cazals et al., 2016](#); [Cao et al., 2011](#)). Une approche suivant une idée similaire consiste à faire une segmentation (qui intégrera des critères de régularité) puis à classifier les régions segmentées ([Touzi et al., 1988](#); [Fjørtoft et al., 1998](#)).

Les champs de Markov, depuis les travaux de ([Geman and Geman, 1984](#)) sont fréquemment utilisés pour introduire des relations spatiales dans un modèle. Cette formulation permet de prendre en compte des relations spatiales à l'échelle de l'image en ne les modélisant qu'à l'échelle du voisinage d'un pixel. Dans un tel cadre la solution optimale est donnée par l'équation suivante:

$$\hat{\mathbf{u}} = \mathcal{E}(\mathbf{u}) = \text{DT}(\mathbf{v}, \mathbf{u}) + \mathcal{P}(\mathbf{u}), \quad (5)$$

où $\text{DT}(\mathbf{v}, \mathbf{u})$ est le terme d'attache aux données (le plus souvent lié à la physique d'acquisition), et $\mathcal{P}(\mathbf{u})$ est l'a priori sur l'image. Ainsi, modéliser un problème avec un champ de Markov revient à définir ces deux termes.

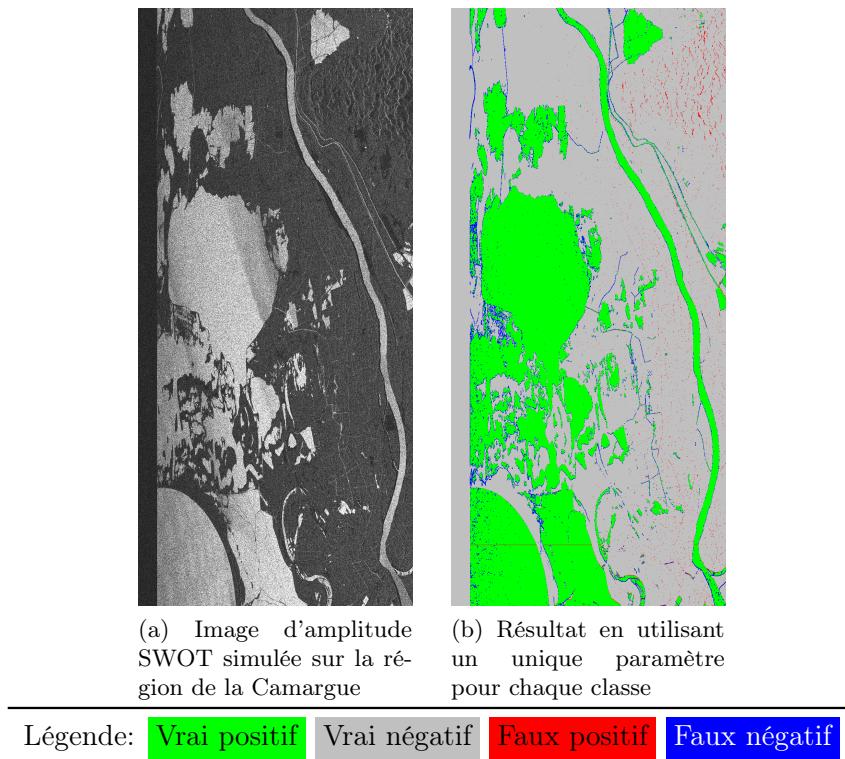


Figure 3 – Extrait d'un résultat obtenu en utilisant un champ de Markov avec des paramètres constants sur une image SWOT simulée. On note une plus grande présence de faux négatifs en début de fauchée qu'en milieu.

Dans un cas binaire, et lorsque l'on souhaite avoir une solution compacte (d'un point de vue spatial), il est courant d'utiliser comme a priori le modèle d'Ising qui pénalise des pixels voisins n'appartenant pas à la même classe. Lorsqu'on applique un tel champ de Markov (en prenant la log-vraisemblance négative de Rayleigh-Nakagami comme attache aux données) sur une image d'amplitude SWOT, on obtient le résultat présenté dans la [Figure 3](#). Comme on peut le constater, le nombre de faux positifs n'est pas le même en début de fauchée qu'en milieu. Ce phénomène s'explique en partie en regardant la [Figure 4](#). On voit que les paramètres de classes ne peuvent être efficacement considérés constants le long de la fauchée à cause du diagramme d'antenne. A cela s'ajoutent les variations locales dues aux variations dans la rugosité des surfaces d'eau.

Prise en compte de paramètres non constants Une façon d'améliorer les résultats de classification est alors de considérer des paramètres non constants dans l'image. Afin d'estimer ces paramètres, quatre méthodes sont proposées dans ce document:

1. estimation basée sur le Xfactor (sous-produit de l'acquisition qui lie le signal reçu à la réflectivité des classes). Cette méthode est une simple inversion de formule mais fait d'hypothèses sur la réflectivité des classes et la topographie (estimée depuis un modèle numérique de terrain potentiellement imprécis et daté).
2. Estimation par fenêtre glissante: en considérant que la plupart des variations

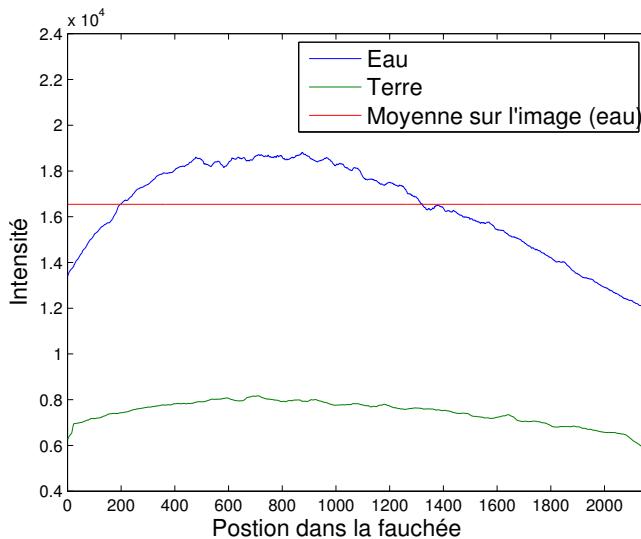


Figure 4 – Evolution moyenne des paramètres le long de la fauchée. Les courbes bleue et verte montrent les paramètres de l'eau et de la terre respectivement, et la courbe rouge montre l'estimation au sens du maximum de vraisemblance du paramètre de l'eau lorsque l'on fait l'hypothèse que celui-ci est constant dans l'image. La courbe de l'eau montre l'effet du diagramme d'antenne alors que les paramètres de la terre sont plus constants car le signal est dominé par le bruit thermique.

de paramètres proviennent du diagramme d'antenne (et ne se produisent donc que le long de la fauchée), il semble naturel, à partir d'une première classification, d'utiliser une fenêtre glissante de la même hauteur que celle de l'image afin d'obtenir les paramètres. Cette méthode a l'inconvénient de ne prendre qu'un type de variation en compte (tout du moins explicitement) et de dépendre d'une classification.

3. Estimations itératives: les deux prochaines méthodes s'appuient sur un schéma itératif tel que présenté dans la [Figure 5](#):

- approche basée régions: l'idée de cette méthode est de partitionner l'image en régions sur lesquelles sont calculés les paramètres. Ainsi, pour chaque région, on obtient un paramètre pour chacune des classes. Pour obtenir une estimation variable des paramètres de bonne qualité, on souhaite respecter deux critères contradictoires: (i) un critère de localité, afin de prendre en compte les variations; (ii) un critère imposant des régions suffisamment grandes, afin que l'estimation soit peu bruitée. Ainsi, on propose une approche utilisant un partitionnement en arbre quaternaire à chaque itération. Le partitionnement est poursuivi jusqu'à ce que le critère (ii) ne soit plus respecté pour toutes les régions (état de convergence).
- approche champs de Markov: pour cette méthode, on modélise les cartes de paramètres par des champs de Markov imposant une régularité spatiale. Le terme d'attache aux données pour la carte des paramètres de la classe eau (resp. terre) n'est défini que pour les pixels marqués comme appartenant à la

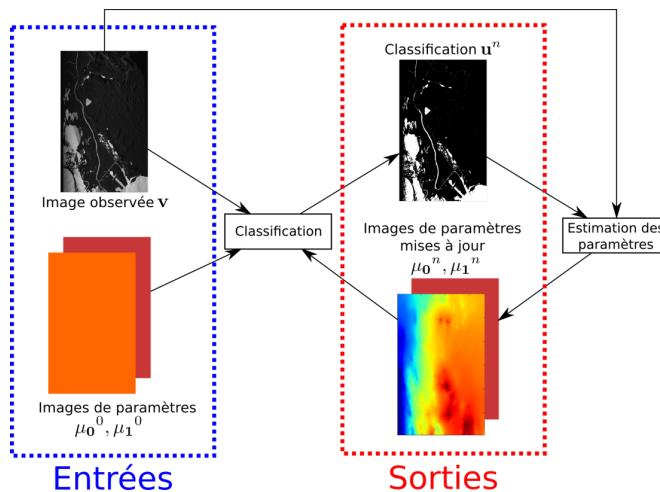


Figure 5 – Estimation itérative des paramètres: à partir de paramètres constants, une première classification est obtenue. De nouveaux paramètres, variables cette fois-ci, sont alors estimés et une nouvelle classification est obtenue. Ce processus est répété jusqu'à convergence ou pour un nombre fixé d'itérations.

Méthode	TPR	FPR	Taux d'erreur (détection d'eau)	MCC
MLE	83.26%	7.54%	54.85%	0.7023
MAP	39.94%	0.32%	61.69%	0.5817
MRF	91.27%	2.11%	19.41%	0.8847
MRF (approche région)	93.16%	2.95%	21.74%	0.8748
MRF (approche Markov)	91.80%	2.18%	19.22%	0.8863

Table 1 – Performances sur l'image de Camargue

classe eau (resp. terre) dans la classification courante. Afin d'avoir une énergie convexe, le modèle est défini sur le logarithme des cartes de paramètres. Dans ce cas-là, les données suivent une distribution de Fisher-Tipett qui peut être approximée par une distribution Normale. Ainsi, le terme d'attache aux données est une simple distance L2. Cette même distance est utilisée pour les termes d'a priori spatiaux et de rappel aux paramètres théoriques définis par le Xfactor.

Résultats Des résultats sont montrés sur l'image de Camargue, à la fois visuellement (Figure 6) et quantitativement (Table 1). Pour ce qui est des performances chiffrées, le Taux de Vrais Positifs (TPR) et le Taux de Faux Négatifs (FPR) montrent que selon le niveau de faux positifs acceptables, différentes méthodes peuvent être utilisées. Lorsque l'on regarde les performances globales (MCC, entre -1 (mauvaise classification partout) et 1 (classification parfaite) et le taux d'erreur (0% pour une classification parfaite)), on constate que l'approche utilisant un champ de Markov pour l'estimation des paramètres donne les meilleurs résultats. Dans la section 4.5, des résultats sur 3 autres images sont montrés. Globalement, on constate que l'estimation des paramètres, qui représente

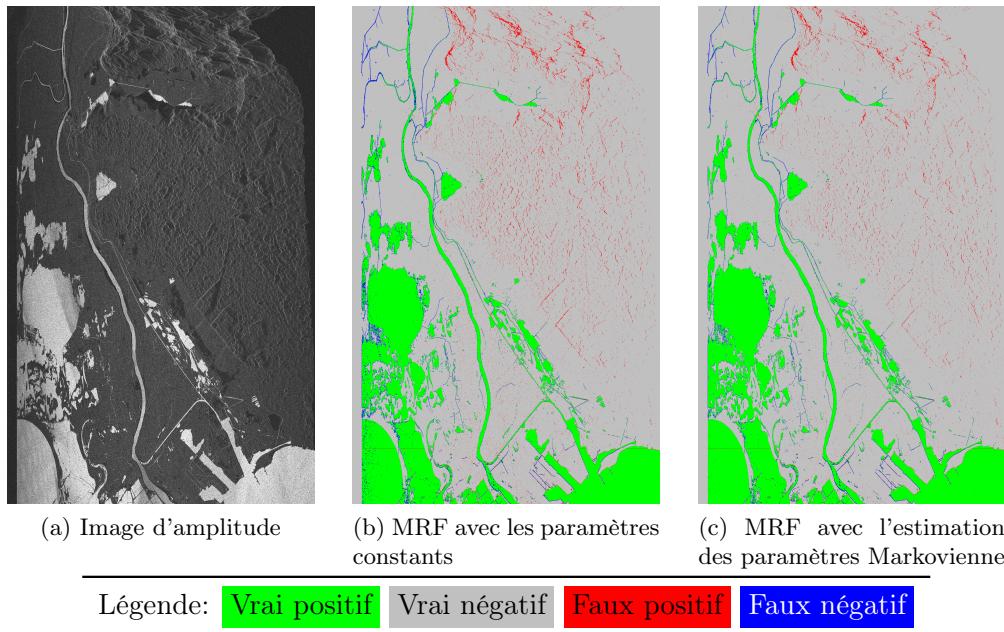


Figure 6 – Résultats sur l'image de Camargue

l'originalité des méthodes proposées, améliore les résultats. Cette amélioration peut-être faible (comme sur le cas de la Camargue) ou beaucoup plus spectaculaire (voir le cas de Kaw dans le Chapitre 4 par exemple).

Détection des rivières fines

De par l'objectif (détection des grands plans d'eau), les méthodes vues à la sous-section précédente ne sont pas capables de détecter les rivières fines. Pourtant, les réseaux de rivières présentent un intérêt pour les utilisateurs finaux de SWOT. Ainsi, un travail dédié à la détection de ces éléments a été effectué. Alors que peu de travaux sur ce sujet en particulier sont présents dans la littérature, des problèmes similaires ont été traités: détection de fissures dans les routes ([Amhaz et al., 2016](#)), la détection de veines dans une image de fond d'œil ([Rossant and others, 2011](#)) ou encore la détection de routes dans les images SAR ([Tupin et al., 1998](#)). Cette dernière méthode utilise une approche en deux temps: d'abord une détection de segments au niveau du pixel, puis un travail sur ces segments dans une étape au niveau des objets. À cela, ([Cao et al., 2011](#)) a ajouté un traitement multi-échelle pour être capable de détecter aussi des plans d'eau dans les images SWOT. Ici encore, on propose une approche basée sur deux étapes, inspirées de ([Tupin et al., 1998](#)). On montre une vue d'ensemble de la méthode proposée dans la Figure 7.

Détection de segments L'objectif de cette étape, résumée dans la [Figure 8](#), est de détecter des segments susceptibles d'appartenir à une rivière. Ces segments pourront alors être utilisés comme entrées dans la partie haut niveau de la chaîne de traitement. Ce détecteur combine deux tests afin de déterminer si le rectangle noté 1 dans la [Figure 8](#)

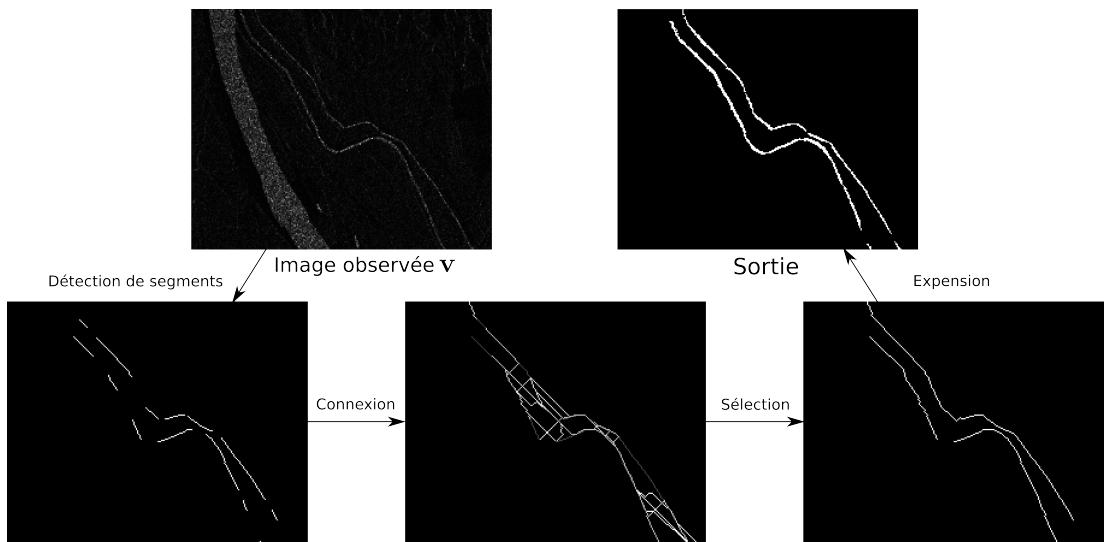


Figure 7 – Vue d’ensemble de la méthode proposée pour la détection des rivières

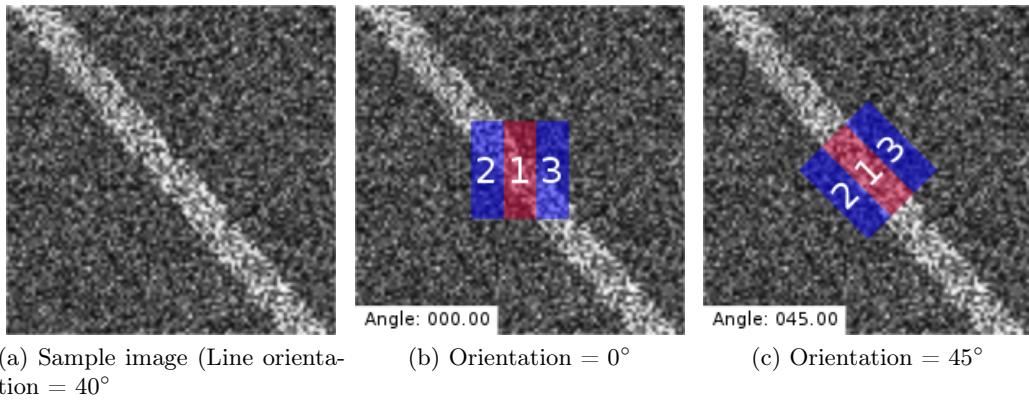


Figure 8 – Illustration du détecteur de segments. Une fenêtre rectangulaire est appliquée à tous les pixels de l’image avec à chaque fois 16 directions et 5 largeurs testées. Ici, lors du test des différentes orientations, on obtiendra un score plus élevé pour la direction 45° que pour 0° . Une animation est disponible sur www.sylvainlobry.com/phd.

suit une distribution différente des rectangles 2 et 3.

Le premier de ces tests, D1 est un détecteur de segment basé sur le détecteur de côtés proposé par (Touzi *et al.*, 1988):

$$re(r_1, r_2) = 1 - \min \left(\frac{\mu_{r_1}}{\mu_{r_2}}, \frac{\mu_{r_2}}{\mu_{r_1}} \right), \quad (6)$$

où μ_{r_x} est le paramètre de distribution dans le rectangle r_x . A partir de ce test, on obtient D1:

$$D1(r_1) = \min (re(r_1, r_2), re(r_1, r_3)). \quad (7)$$

Ainsi, D1 contraint d’avoir une réponse forte du détecteur de côtés entre les trois rectangles.

D2 est un détecteur basé sur la corrélation croisée normalisée entre r_1 et r_2, r_3 . Dans

(Tupin *et al.*, 1998), il est défini ainsi:

$$D2(r_1) = \min(cc(r_1, r_2), cc(r_1, r_3)) , \quad (8)$$

avec cc la corrélation croisée normalisée donnée par:

$$cc(r_x, r_y) = \sqrt{\frac{n_x n_y (\mu_x - \mu_y)^2}{1 + (n_x + n_y)(n_x \sigma_x^2 + n_y \sigma_y^2)}}, \quad (9)$$

avec n_x le nombre de pixels dans r_x et σ_x^2 la variance de r_x .

Enfin, les deux détecteurs sont fusionnés avec une somme associative et symétrique (définie par (Bloch, 1996)). Cela permet d'avoir un score, lié à la probabilité d'appartenir à une rivière, pour chaque pixel. Ce score est ensuite seuillé, puis les détections isolées sont enlevées et enfin les pixels détectés sont regroupés en segments.

Connexion et régularisation À partir des segments détectés, les traitements suivants sont effectués:

1. Une étape de connexion: les segments suffisamment proches (en termes de distance et d'orientation) sont connectés à l'aide de l'algorithme de Dijkstra (Dijkstra, 1959) défini sur une distance inversement proportionnelle aux scores définis dans l'étape précédente. Ainsi, les connexions passent par des points susceptibles d'appartenir à des rivières.
2. À l'étape précédente, de nombreuses connexions sont effectuées. L'objectif de l'étape de sélection est de ne garder que celles qui ont une faible distance (soit une forte probabilité d'appartenir à des rivières) et qui répondent à des a priori que l'on peut placer sur la forme globale d'un réseau de rivières. Ces a priori sont les suivants:
 - un réseau de rivières contient peu de nœuds terminaux;
 - peu d'intersections;
 - peu de bifurcations;
 - un réseau est principalement rectiligne;
 - les connexions ne sont censées que remplir des trous, elles doivent donc être courtes;
 - les segments longs (détectés lors de la première étape) appartiennent très probablement au réseau.

Ces différentes propriétés sont modélisées par un champ de Markov dont l'objectif est de donner une étiquette à chaque connexion indiquant si celle-ci fait partie du réseau. Celui-ci est optimisé avec un algorithme simple : ICM (Modes Conditionnels Itérés). Cela nous permet de trouver un minimum local rapidement. L'initialisation est alors un aspect important : elle est obtenue avec un seuil sur les distances des connexions.

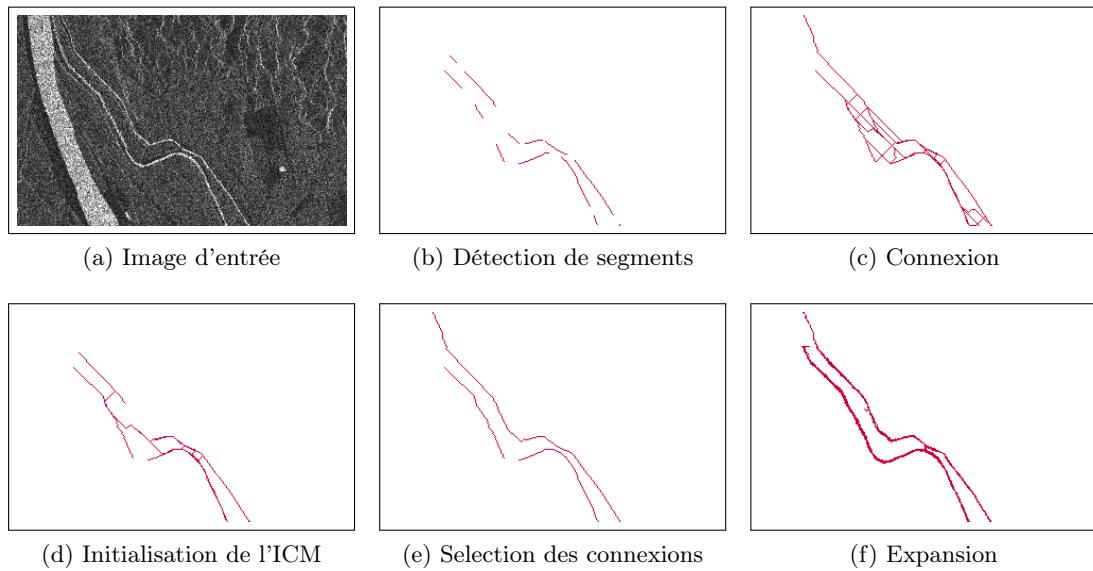


Figure 9 – Illustration des différentes étapes sur un extrait de l'image de la Camargue.

Méthode	TPR	FPR	Taux d'erreur (détection d'eau)	MCC
MRF (Markovian)	92.78%	1.64%	15.52%	0.9074
MRF + Rivières	93.08%	1.69%	15.46%	0.9080

Table 2 – Résultats pour l'image de la Camargue

3. À l'issue de l'étape précédente, le réseau est représenté par des chaînes de pixels de largeur 1. Cependant, nous sommes d'abord intéressés par une classification pixellique. Afin de l'obtenir, mais aussi d'améliorer le positionnement des rivières détectées, nous proposons une approche simple:

- (a) le réseau est étiqueté en composantes connexes;
- (b) la boîte englobante de chaque composante connexe est débruitée à l'aide de l'algorithme NL-SAR ([Deledalle et al., 2015](#));
- (c) l'image débruitée est seuillée par rapport à la radiométrie théorique de l'eau, puis un étiquetage en composantes connexes est effectué;
- (d) on garde les composantes connexes ayant une intersection avec le réseau initial.

Résultats On montre dans la [Figure 9](#) un extrait du résultat obtenu aux différentes étapes de la méthode proposée et des résultats chiffrés dans la [Table 2](#) (notons que les résultats de MRF (approche Markov) seuls ne sont pas les mêmes que dans la [Table 1](#) car on utilise une initialisation présentée dans la suite du document pour les paramètres). On peut noter que l'amélioration des performances est faible, mais cela est attendu lorsque l'on considère la faible taille relative des rivières par rapport aux lacs.

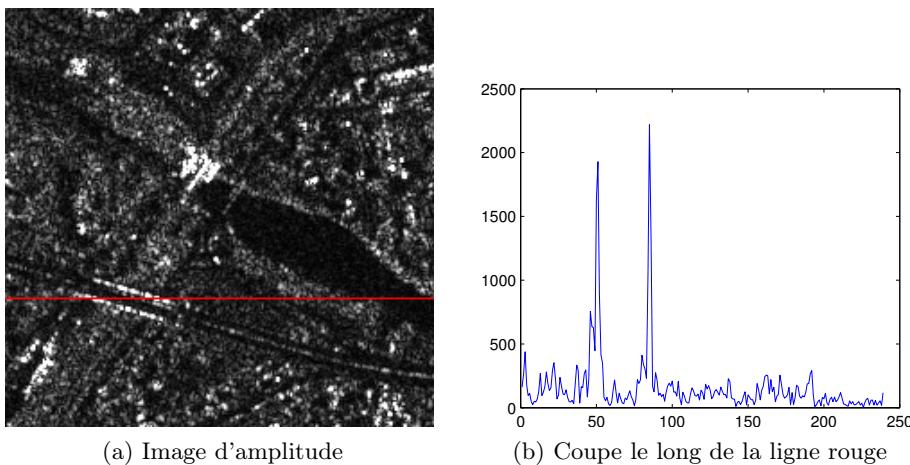


Figure 10 – Illustration du phénomène de speckle et des points brillants sur une image acquise par TerraSAR-X (mode Stripmap) à Saint-Gervais, France (GPS: 45.910877, 6.705635). Les points correspondants au pont ont une radiométrie environ 10 fois supérieure à celle des autres points.

Traitements SAR multi-temporels

Alors que le travail présenté dans la section précédente est dédié à la détection de l'eau dans les images SWOT mono-dates, cette partie traite des piles temporelles d'images urbaines.

Les images SAR sont corrompues par le phénomène de speckle qui peut rendre leur interprétation complexe. Qui plus est, les images urbaines présentent souvent des points brillants, caractéristiques des constructions humaines. Ces deux phénomènes sont illustrés dans la Figure 10.

Il est possible de réduire l'influence du speckle en produisant une image multi-vue (soit en moyennant sur une fenêtre, on parle alors de multi-vue spatial, soit en temps), au prix d'une perte de résolution spatiale ou temporelle. Aussi, cette approche ne donne pas de résultats satisfaisants dans les régions non-homogènes, fréquentes à cause des points brillants notamment. Des approches utilisant des fenêtres adaptatives sont alors souvent adoptées (Lee, 1981; Vasile *et al.*, 2006). D'autres approches, non-locales, comparent des patchs afin de moyenner des pixels similaires. L'idée originale de (Buades *et al.*, 2005) a été étendue par (Deledalle *et al.*, 2011) au SAR, notamment pour la prise en compte des statistiques particulières du speckle et des points brillants.

Une façon différente de procéder est de régulariser l'image. On peut par exemple citer la variation totale (TV) qui pénalise les variations entre pixels voisins tout en préservant les contours (Rudin *et al.*, 1992). Les cibles fortes sont alors un problème, comme illustré dans la Figure 11.

Dans ce travail, on s'intéresse au traitement de séries temporelles d'images SAR. Le lancement récent de constellations de satellites permet d'obtenir des images d'un même lieu à des fréquences importantes. Par exemple, en utilisant les satellites Sentinel-1A et Sentinel-1B, il est possible d'obtenir une nouvelle image tous les 6 jours. Il est

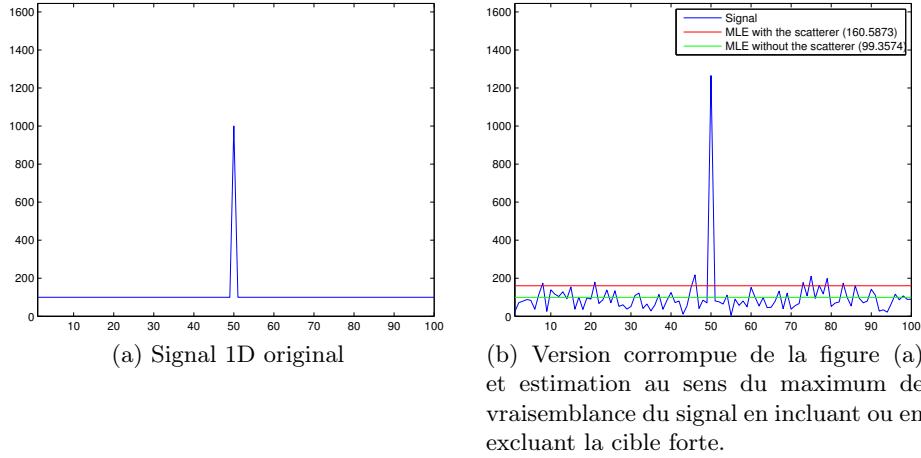


Figure 11 – Illustration de l’effet de la présence d’une cible forte sur la régularisation. Lorsque celle-ci est incluse dans le calcul du maximum de vraisemblance, la vraie radiométrie est sur-estimée. En excluant la cible en question, l’estimation est proche de la vraie radiométrie.

alors possible d’obtenir une meilleure régularisation, ou de faire de la détection de changement entre ces différentes images.

Dans cette section, on propose de résoudre conjointement deux problèmes:

- un problème d’estimation (des radiométries non corrompues);
- un problème de détection (de cibles et éventuellement de changement).

Dans la suite, nous présentons 3 modèles issus de cette idée.

Modèles

L’idée à la base des modèles proposés est de considérer que la pile temporelle observée est la somme de deux composantes. Ainsi, on cherche à expliquer une pile de T images \mathbf{v} où chaque image contient N pixels. On considère alors le modèle suivant:

$$\forall t \in \{1, \dots, T\}, \forall i \in \{1, \dots, N\}, v_{t,i} = u_{t,i} \cdot \xi_{t,i} \quad (10)$$

$$= (b_{t,i} + s_{t,i}) \cdot \xi_{t,i}, \quad (11)$$

où \mathbf{u} est la radiométrie sous-jacente de la scène (soit la version non corrompue de \mathbf{v} que l’on cherche à obtenir), \mathbf{b} est la composante de fond et \mathbf{s} est la composante de cibles fortes. L’idée générale de la méthode proposée est de chercher à obtenir ces deux composantes conjointement au lieu de suivre une approche plus classique, où l’on chercherait à obtenir \mathbf{u} directement. Dans un cadre du Maximum a Posteriori, estimer \mathbf{u} directement serait exprimé comme une énergie à minimiser:

$$\mathcal{E}(\mathbf{u}) = DT(\mathbf{u}, \mathbf{v}) + R(\mathbf{u}), \quad (12)$$

où $\text{DT}(\mathbf{u}, \mathbf{v})$ est un terme pénalisant un modèle \mathbf{u} qui ne représenterait pas fidèlement l'observation \mathbf{v} alors que $R(\mathbf{u})$ encode une connaissance a priori sur les propriétés attendues du modèle. L'approche de décomposition est une réponse au problème que rencontrent les a priori spatiaux classiques lorsque ils sont utilisés en imagerie SAR: en utilisant la décomposition présentée dans [Equation 11](#) et en faisant l'hypothèse que les deux composantes sont indépendantes, on obtient l'énergie suivante à minimiser:

$$\mathcal{E}(\mathbf{u}) = \text{DT}(\mathbf{b} + \mathbf{s}, \mathbf{v}) + \beta_{\text{BG}} R(\mathbf{b}) + \beta_{\text{S}} R(\mathbf{s}), \quad (13)$$

où β_{BG} et β_{S} permettent de pondérer les différents termes. En procédant ainsi, on est capable de mettre des a priori différents sur le fond \mathbf{b} et sur les cibles \mathbf{s} . Dans la suite, on présente trois modèles qui diffèrent par les a priori utilisés. L'a priori sur le fond $R(\mathbf{b})$ est noté $\psi(\mathbf{b})$ et l'a priori sur les cibles $R(\mathbf{s})$ est noté $\lambda(\mathbf{s})$.

Un a priori classiquement utilisé est de dire que l'image devrait avoir une variation totale faible. Alors que cet a priori n'est pas viable lorsque des cibles fortes sont présentes dans l'image, on peut l'appliquer à la composante de fond, où de telles cibles ne sont pas présentes. Aussi, on cherche à travailler sur des piles temporelles. On présente donc une extension simple de la variation totale spatiale au domaine temporel:

$$\text{TV}_{3D}^{\alpha}(\mathbf{b}) = \sum_{t=1}^T \sum_{i \sim j} |b_{t,i} - b_{t,j}| + \alpha \sum_{t=1}^{T-1} \sum_{i=1}^N |b_{t+1,i} - b_{t,i}|, \quad (14)$$

où $i \sim j$ indique que les pixels i et j sont voisins. On fait donc l'hypothèse que les données sont parfaitement calibrées et recalées. En utilisant cette formulation, il est possible de pénaliser les variations spatiales et temporelles de manière indépendante via le paramètre α .

On peut définir une cible forte comme un point ayant une radiométrie largement supérieure à celle des pixels environnants. D'après cette définition, on peut déduire que la composante de cibles fortes devrait être parcimonieuse. La manière directe de modéliser cette hypothèse est d'utiliser la pseudo norme L0:

$$\|\mathbf{s}\|_0 = \sum_{t=1}^T \sum_{i=1}^N \bar{\delta}(s_{t,i}), \quad (15)$$

où $\bar{\delta}(x)$ est égal à 1 si $x \neq 0$. Cependant, lorsque l'on utilise la pseudo-norme L0 comme a priori, l'énergie à minimiser est combinatoire (non-continue et non-convexe). Typiquement, on utilise alors la norme L1. Cependant, on verra dans la suite que l'on est capable, au prix d'une discrétisation du problème et dans le cas où on utilise TV pour le fond, d'optimiser de manière efficace l'énergie définie avec L0.

En utilisant des fonctions différentes pour ψ et λ , trois modèles sont proposés:

- **TVL0 Regularization (TVR):** Dans ce modèle, on utilise $\psi = \text{TV}_{3D}^{\alpha}$ avec α un nombre fini (et dans ce document, $\alpha = 1$) et $\lambda = \|\cdot\|_0$. Ce modèle est adapté à la régularisation et la détection de cibles fortes.

- **TVL0 One Background (TV1BG):** dans ce modèle, on utilise $\psi = \text{TV}_{3D}^\infty$ et $\lambda = \|\cdot\|_0$. En utilisant un poids infini pour la régularisation temporelle, aucune variation dans le temps n'est permise à un pixel donné. Ainsi, ce modèle fait l'hypothèse que les changements ont lieu uniquement dans les cibles fortes. En procédant ainsi, on obtient un modèle avec une occupation mémoire moindre. Ce modèle est particulièrement adapté pour la détection de cibles fortes.
- **TVL0 One Change (TV1C):** Dans ce modèle aussi, on cherche un fond seulement ($\psi = \text{TV}_{3D}^\infty$). Il est différent de TV1BG dans le sens où l'on ne cherche pas seulement une composante de fond et une de cibles, mais aussi une parcimonie dans les changements. Afin que le modèle reste utilisable, on se limite à 4 cas pour les valeurs que peut prendre une cible forte dans le temps:
 - elle peut être toujours égal à 0 (pas de cible forte);
 - elle peut être constante et positive (pas de changement);
 - elle peut être égale à 0 jusqu'à une date t_{app} puis être constante et positive (apparition);
 - elle peut être constante et positive jusqu'à une date t_{disp} puis être égale à 0 (disparition).

On peut ensuite définir une carte de changements $\mathbf{c}(\mathbf{s})$ comme:

$$c_i(\mathbf{s}) = \begin{cases} 0 & \text{si } \exists r > 0, \forall t \in [1, T], s_{t,i} = 0 \text{ ou } s_{t,i} = r; \\ 1 & \text{si } \exists r > 0, \forall t \in [1, t_{\text{app}}[, s_{t,i} = 0 \\ & \text{et } \forall t \in [t_{\text{app}}, T], s_{t,i} = r; \\ 2 & \text{si } \exists r > 0, \forall t \in [1, t_{\text{disp}}[, s_{t,i} = r \\ & \text{et } \forall t \in [t_{\text{disp}}, T], s_{t,i} = 0, \end{cases} \quad (16)$$

avec $r > 0$. Pour garantir la parcimonie des changements, on met un a priori additionnel: $\lambda(\mathbf{s}) = \|\mathbf{s}\|_0 + \frac{\beta_C}{\beta_S} \|\mathbf{c}(\mathbf{s})\|_0$, où β_C est un poids. Par définition, ce modèle est adapté à la détection de changement.

On interprète ensuite la décomposition comme un problème de détection: en un pixel, pour une valeur de fond donnée, la détection de cible peut être vue comme un test de rapport de vraisemblance généralisé (en remplaçant la cible par son estimation au sens du maximum de vraisemblance) et la détection de changement comme un test d'hypothèses hiérarchique, impliquant une succession de tests de rapport de vraisemblance généralisée.

Ce problème de détection est ensuite reformulé comme un problème d'estimation (estimation des valeurs de cibles et du fond). Cette approche permet d'obtenir une énergie à minimiser. D'après l'expression de la décomposition comme un problème de détection, il est possible d'exprimer la valeur optimale des cibles comme une fonction du fond, permettant d'avoir une énergie ne dépendant plus que d'une variable (le fond).

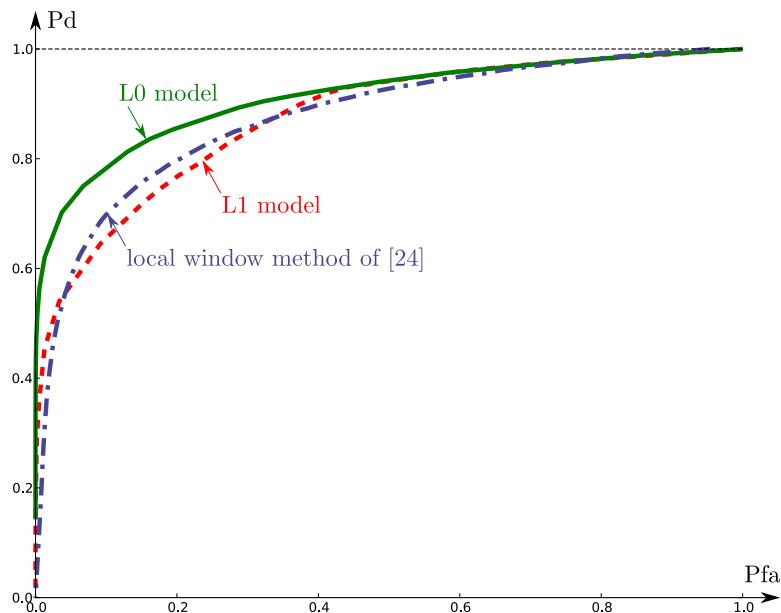


Figure 12 – Fonction d’efficacité du récepteur (ROC, probabilité de détection (P_d) contre Probabilité de fausse alarme (P_{fa})) entre les version L0 et L1 de notre modèle ainsi qu’une méthode basée sur une analyse locale (Lopes *et al.*, 1993).

Comme l’a priori choisi est la variation totale, il est alors possible d’utiliser la méthode d’optimisation proposée par (Ishikawa, 2003) afin d’obtenir le minimum global de notre problème, après quantification des valeurs possibles du fond.

Cette méthode d’optimisation peut être coûteuse en termes de quantité de mémoire utilisée. Afin d’obtenir un compromis entre la qualité du résultat et la mémoire nécessaire, on propose un schéma d’optimisation par blocs avec recouvrement. Cette méthode d’optimisation n’est pas dédiée à notre problème; elle peut être utilisée pour tous les modèles prenant en compte un contexte spatial, comme les méthodes proposées pour la détection de l’eau dans SWOT par exemple.

Applications et résultats

Les modèles proposés sont évalués dans les trois applications visées: détection de cible, régularisation et détection de changement.

Détection de cible Une évaluation simple sur une image synthétique est proposée dans la Figure 12. On note ainsi que pour la détection de cible notre modèle est pertinent. Par ailleurs, cette évaluation nous permet aussi de conclure sur la pertinence de ne pas relâcher la pseudo-norme L0 en norme L1.

Régularisation Des résultats visuels de deux modèles (TVR et TV1BG) sont montrés et comparés à des techniques classiques (variation totale et Multi-vue temporel) dans la Figure 13 sur une pile acquise sur Paris par TerraSAR-X en mode spotlight.

L’image multi-vue présente deux inconvénients majeurs: d’une part, il y a une perte de résolution temporelle (on obtient seulement une image pour la totalité de la série).

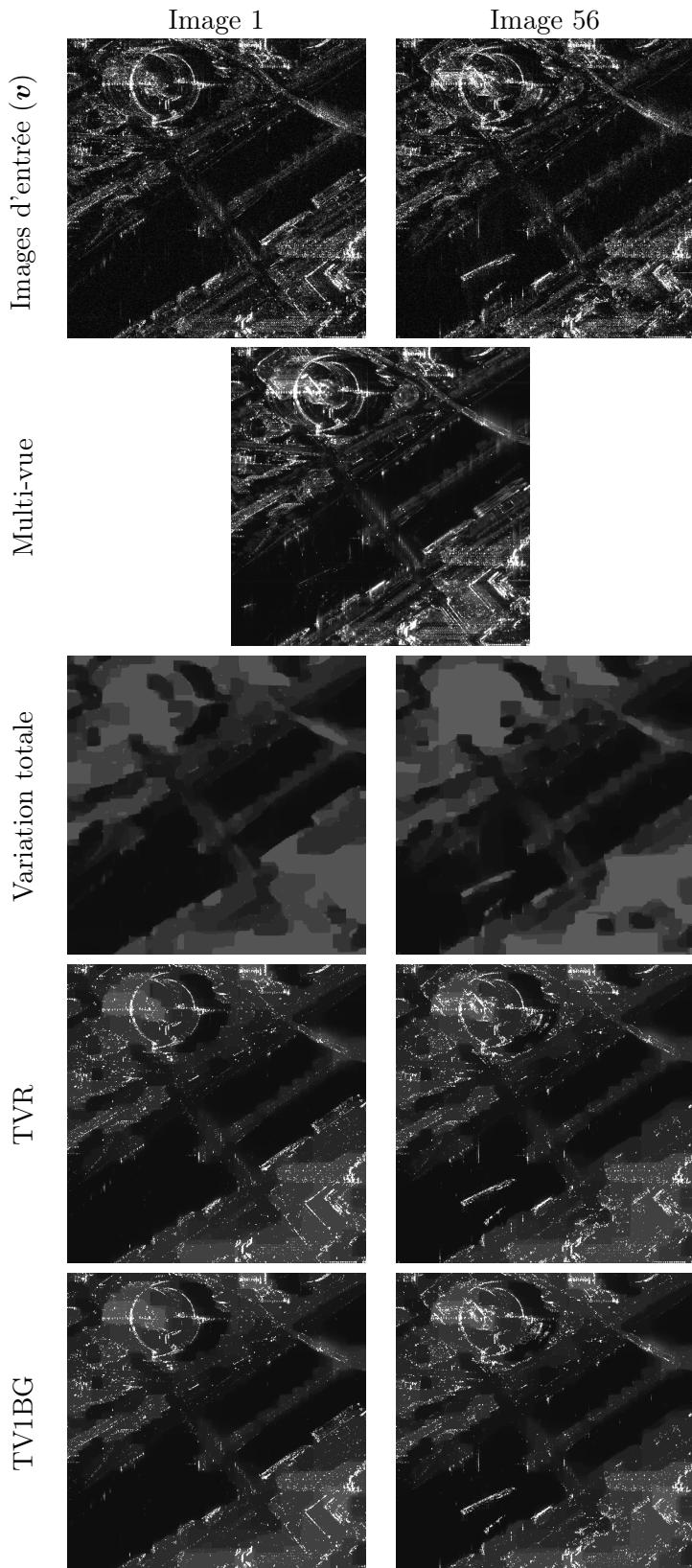


Figure 13 – Comparaison des différents résultats de régularisation.

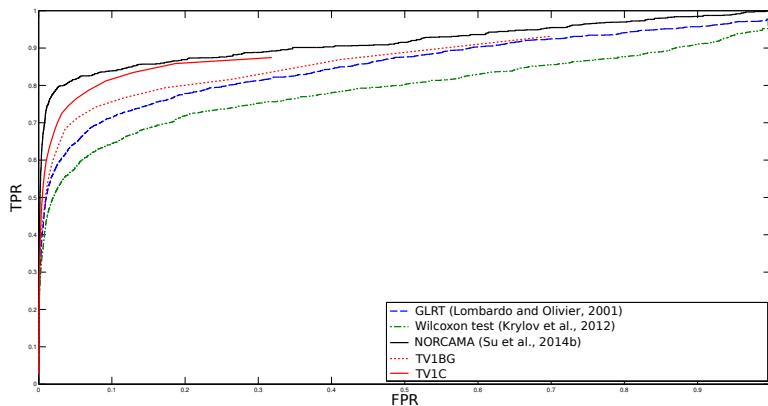


Figure 14 – Courbe ROC des performances de détection de changement sur une série d’images de Saint-Gervais, France, acquise par TerraSAR-X en mode stripmap.

Aussi, on note une persistance des cibles fortes présentes sur seulement une image (visible pour les péniches sur la Seine): leur radiométrie bien plus élevée que les autres pixels fait qu’elles seront toujours visibles après moyennage.

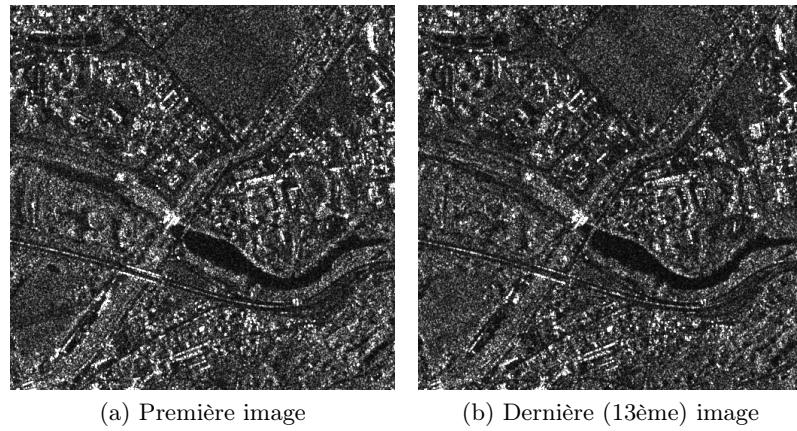
En ce qui concerne la variation totale, la difficulté dans la gestion de ces cibles fortes apparaît clairement. Cela valide la pertinence de traiter l’image comme une somme de composantes.

Détection de changement Les résultats du modèle TV1C pour la détection de changement sont évalués visuellement dans la Figure 15 et comparés à d’autres méthodes dans la Figure 14.

Visuellement, on peut constater que les changements sont regroupés spatialement. Cela n’est pas contraint par le modèle, et le fait de retrouver cette propriété dans les résultats est un bon signe. Enfin, à part la partie supérieure du pont, les changements sont plutôt bien détectés. Lorsque l’on compare nos résultats à l’état de l’art, on peut remarquer que l’on s’approche de la méthode NORCAMA proposée par (*Su et al., 2014a*). Cependant, il faut noter que ce modèle donne dans le même temps une composante de fond et de cibles qui peuvent être utiles pour l’interprétation des images. Qui plus est, les changements détectés ne sont pas les mêmes. Dans la première étape de (*Su et al., 2014a*), un débruitage est effectué, introduisant une régularité spatiale. D’un autre côté seuls les changements dans les cibles fortes (réfléchant une activité humaine) sont détectés par notre modèle. Ainsi, les résultats proposés par ces deux modèles sont assez différents et complémentaires, et selon l’application visée l’un ou l’autre pourra être préférable.

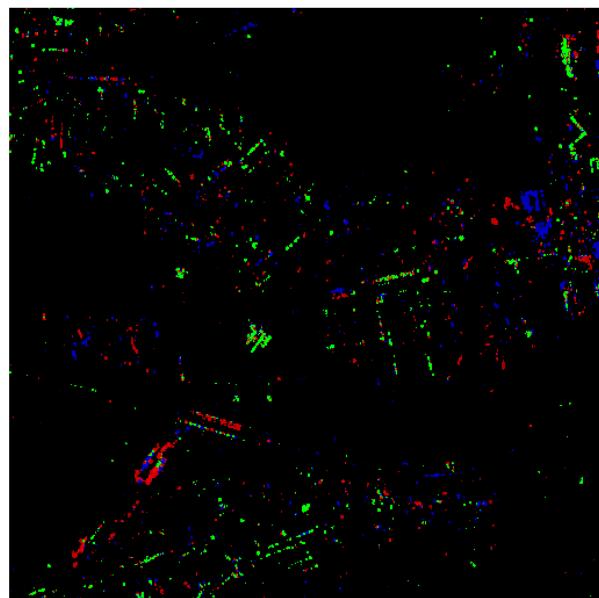
Conclusion

L’objectif principal de cette thèse est de fournir des méthodes fiables pour la classification eau/terre dans les images SWOT. Cette tâche est critique pour l’estimation des hauteurs d’eau. Suite à une division des objets d’intérêt en hydrologie selon leur forme,



(a) Première image

(b) Dernière (13ème) image



(c) carte de changement

légende: Pas de cible forte | Cible forte et constante | Disparition | Apparition

Figure 15 – Résultats sur une série d’images de Saint-Gervais, France, acquise par TerraSAR-X en mode stripmap.

nous proposons deux familles de méthodes:

- les méthodes pour les objets grands et compacts. Les méthodes de classification généralement utilisées ne sont pas adaptées à SWOT car elles font l'hypothèse de paramètres de classes constants. La contribution majeure sur cette partie est un cadre utilisant un modèle bien connu (champs de Markov avec un modèle de Potts comme a priori) utilisant et estimant des paramètres de classe variables.
- une méthode pour l'extraction des rivières fines. Celles-ci ne respectent pas l'hypothèse de régularité spatiale faite par une grande partie des algorithmes de classification. A cet effet, on propose une combinaison de deux étapes: une étape de bas niveau, pour la détection de segments et une étape de haut niveau, pour leur connexion et régularisation.

Une partie de cette thèse a aussi été dédiée à l'implémentation opérationnelle des algorithmes dans la chaîne de traitement du CNES.

Un des intérêts de la mission SWOT sera sa capacité à mettre à jour les données, grâce au temps de revisite du satellite de 21 jours. Il sera alors intéressant d'explorer le traitement multi-temporel des données. Néanmoins, de telles données n'étaient pas disponibles au moment de la thèse. Ainsi, nous travaillons sur des méthodes similaires à celles utilisées dans SWOT dans un contexte différent: les données multi-temporelles en milieu urbain. On présente trois modèles issus d'une même idée de départ (la décomposition de la série d'images) visant trois applications différentes: la régularisation, la détection de cibles et la détection de changement. Ces modèles permettent d'effectuer des traitements couramment demandés sur de grandes séries temporelles.

Perspectives Lorsque des données réalistes seront disponibles, il sera important de vérifier l'adaptation des méthodes multi-temporelles proposées pour la détection de l'eau dans le contexte de SWOT. Il sera aussi important de vérifier le comportement des méthodes lorsque la réflectivité théorique sera différente de celle observée. Ce dernier point devrait surtout impacter les méthodes ne proposant pas de ré-estimation.

Concernant la partie multi-temporelle, le modèle de détection de changement a deux défauts principaux: il n'est pas capable de prendre en compte plus d'un changement par pixel, et la probabilité de détection n'est pas constante avec le temps. Ce dernier point est en cours d'étude. Il serait possible de considérer un nombre fini de changement, mais la complexité algorithmique grandirait alors de manière exponentielle.

Chapter 1

Introduction

1.1 Context

Water is ubiquitous on Earth: according to (Gleick, 1993), 71% of Earth's surface is covered with water. However, 96.5% of the water present on Earth is saline water from the oceans. The fresh water, which is the resource we all need to live only accounts for 2.5% of the total water present on earth (the remaining 1% accounts for other sources of saline water, such as groundwater). Most of it (about 69%) is present in ice caps and glaciers. Fresh groundwater represents 30% of the total fresh water, and lakes only account for 0.26% (or 0.007% of the total water on Earth). While this resource is finite, the growth of demand for water rose by more than the double of the population growth over the last century according to an estimation from OECD¹. This makes access to safe water more difficult, with dramatic effects: according to the WHO, 3.4 million people die annually from water-related diseases². The situation is only getting worse: according to UN, in 2025 two-thirds of the world's population could face problems concerning access to water³. Furthermore, the water cycle implies changes in water locations. It then appears crucial to monitor the changes. Hydrologists mainly use data acquired *in situ*, but these acquisitions are spatially sparse and remain expensive.

In order to obtain global measurements, they have started to use data from earth observation satellites. For instance, we can cite the Jason missions (the latest satellite, Jason-3 was launched on January 2016) which uses radar altimetry to measure the topography of ocean's surfaces (Lambin *et al.*, 2010). In hydrology, the GRACE mission (Tapley and Reigber, 1999) maps the gravity field of the Earth using a microwave ranging system and has been used to monitor changes in the water storage (Ramillien *et al.*, 2008).

The upcoming Surface Water & Ocean Topography (SWOT) mission aims at making the first global survey of the Earth's water. It is a collaboration between the

¹see <http://www.oecd.org/publications/factbook/34416097.pdf>

²see http://www.who.int/water_sanitation_health/takingcharge.html

³see <http://www.un.org/waterforlifedecade/scarcity.shtml>

Centre National D'Études Spatiales (**CNES**) and NASA's Jet Propulsion Laboratory (**JPL**) with contributions from the Canadian space agency (**CSA**) and the United Kingdom space agency (**UKSA**), following the historical collaboration on environmental satellites between the first two agencies. The launch of the satellite is planned in April 2021. To fulfill **SWOT**'s objective of measuring variations in the water elevation, **Synthetic Aperture Radar (SAR)** is particularly adapted: it has the capability to reliably acquire data (regardless of the weather or day/night conditions), a radiometric stability which will provide comparable measurements over time and the interferometry can retrieve the water elevation. Toward the goal of measuring inland water elevations, one of the steps is to detect the water. In this collaboration, a precise distribution of the algorithms development tasks has been proposed, and **CNES** is responsible for the water detection part. Developing algorithms to detect the water in inland **SWOT** images is the main objective of the presented work. Some methods were implemented in the **CNES** processing chain prototype as part of the work of this PhD.

We have seen that one of the objectives of **SWOT** is to monitor changes, which implies processing of data of the same site acquired at different dates. As realistic multi-temporal data are not yet available, we have used **SAR** data coming from other sensors, with the same family of processing methods but for different applications. We studied applications of **SAR** in urban areas, and we will see in the rest of this document that parts of this work could easily be transferred to **SWOT** data.

1.2 Contributions

Based on the context described in section 1.1 the contributions of this manuscript can be divided in three parts:

- dedicated methods for water detection in **SWOT** images;
- methods for the processing of multi-temporal urban **SAR** data;
- generic methods for the processing of multi-temporal data and/or large data.

While the first and second ones target a specific application, the third type of contribution groups the techniques developed for the processing of urban **SAR** data that could be transferred to the processing of **SWOT** data.

Dedicated methods for water detection: Objects of interest in hydrology can generally be classified based on their shape: either it is a large and compact object (lake/reservoir) or it is long and thin (river). Therefore we have developed two methods to account for these two types of objects.

Contribution (1): Method for the detection of large water bodies

Due to the antenna pattern, varying water roughness and low **Signal-to-noise ratio (SNR)** over land in **SWOT** images, the parameters describing the water class and the land class can not be considered constant through the image. To take into account this effect we present three methods for the estimation of the water and land parameters:

- estimation based on the prior knowledge of the so-called Xfactor (using a **Digital Elevation Model (DEM)**, the known shape of the antenna pattern and the **SNR** to retrieve the parameters);
- iterative estimation based on a quad-tree partitioning of the image;
- iterative dense estimation based on a **Markov Random Field (MRF)** model on the parameters maps.

Classification is done jointly with the parameters estimation with an Ising model (enforcing spatial compactness) which is optimized exactly.

Contribution (2): Detection of narrow rivers

We propose a method for the detection of thin water bodies in **SWOT** images based on a two-step approach:

1. pixel-based detection of river segments;
2. connection and selection of the connections based on their contributions to global properties of a river network.

The pixel-based detection is inspired from the work of (*Tupin et al., 1998; Cao et al., 2011*) and the connection step uses Dijkstra's algorithm. Finally the selection of the connections uses a Markovian model to introduce priors on the global shape of the river network.

Methods for the processing of urban SAR data: One of the main challenges for the interpretation of **SAR** images is the presence of speckle which can be modeled as a multiplicative noise. Many speckle filtering techniques have been developed but most of them fail in the presence of strong scatterers (pixels with a radiometry an order of magnitude higher than their surrounding) that are typically present in urban areas. To this effect, (*Denis et al., 2010*) presented a decomposition technique which models the image as a sum of two components: a background (which is spatially regular) and a strong scatterers component, with only a few points. In this work, we propose an extension of this model to multi-temporal series of images and applications for strong

scatterers detection, change detection and regularization.

Contribution (3): A decomposition framework for multi-temporal series of SAR images

We present a general framework for the decomposition of multi-temporal series of SAR images in two components: one or several backgrounds and strong scatterers components for each image of the series. This formulation makes it possible to put different priors on the strong scatterers components, and three models are derived:

- TVR: one background and one strong scatterer components for each input image which can be used for the regularization and strong scatterers detection in time series;
- TV1BG: one background for the whole time series and one strong scatterer image for each input image. This model is adapted for strong scatterers detection;
- TV1C: one background for the whole time series and one strong scatterer image for each input image. An additional constraint is put on the series of strong scatterer images so it can be used for change detection.

We also propose a technique to use results from TVR or TV1BG for change detection. One advantage of the proposed formulation is that it can be optimized exactly once possible values for the background are quantized.

We also show that using a convex relaxation of the L0 pseudo-norm to the L1 norm in this context deteriorates the results.

Generic methods for the processing of multi-temporal and/or large data:

Some of the methods for the processing of urban **SAR** data could be transferred in the case of **SWOT** when multi-temporal data become available. In addition to the simple extension of **Total Variation (TV)** for a multi-temporal stack of images, which could be used in an Ising model for the detection of large and compact objects, we have identified the following contribution:

Contribution (4): Efficient optimization of MRF

Graphcut optimization techniques can be used to obtain the global minimum of some energies (which is the case for contributions 1 and 3) but generally requires a lot of memory which can prevent its application in real cases. To this effect, we developed a simple technique to optimize by blocks. This method provides a trade-off between the needed memory, the computation time and the quality of the results. In the case of **SWOT**, it could be useful for the detection of water in large images or large number of images in a multi-temporal series.

Finally, we present in [Figure 1.1](#) a diagram showing the different contributions presented in this thesis and how they are linked together.

1.3 Organization of this manuscript

This manuscript is divided in three parts. In a first part, we review both the data in [chapter 2](#) and the models in [chapter 3](#). In [chapter 2](#), we explain the acquisition process and the statistics of **SAR** images that are used in this document. We also outline the particular characteristics of **SWOT**. In [chapter 3](#), we review the Bayesian models that are at the core of the developed methods.

The second part is dedicated to water detection in **SWOT** amplitude images. After reviewing methods from the literature that are adapted to the task of binary classification in **SAR** images, we present in [chapter 4](#) methods that take into account the class parameters variations that occur in **SWOT** images. The [chapter 5](#) is dedicated to the detection of narrow curvilinear elements (representing rivers), and a two-step approach is proposed.

A third part is dedicated to the processing of multi-temporal series of **SAR** images. We show that the strong scatterers present in images of urban areas generally prevent the direct application of classical regularization techniques, and we propose in [chapter 6](#) a decomposition model that jointly detects these strong scatterers and estimate the radiometry of the scene. As the proposed model can process multi-temporal series, a change detection method is proposed in this chapter and evaluated in [chapter 7](#), along with the evaluation of the performances in regularization and strong scatterers detection.

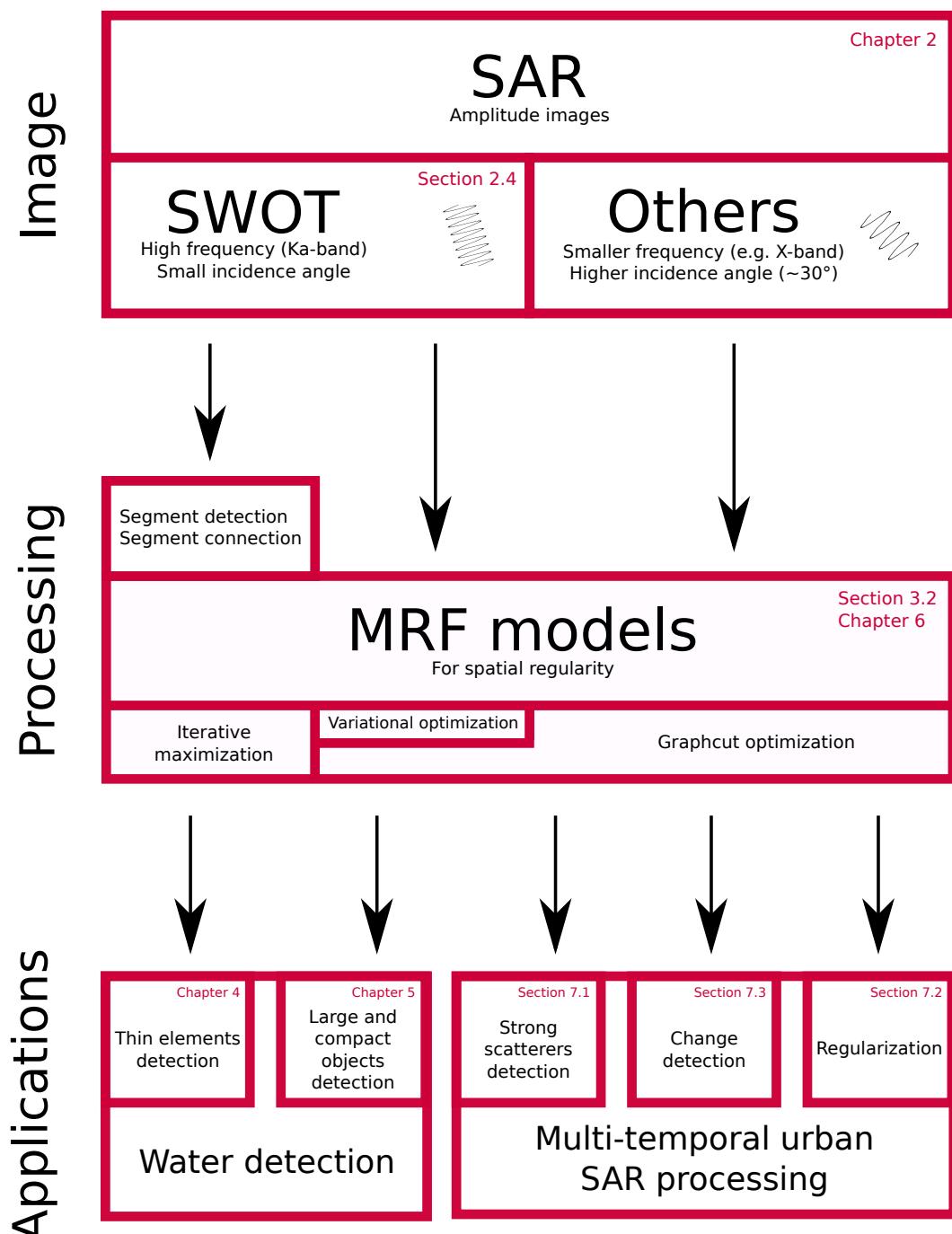


Figure 1.1 – Schematic view of the work achieved during this PhD.

Part I

Image models

Chapter 2

Physics and statistics of SAR images

2.1 Introduction

SAR images are corrupted by a phenomenon called speckle which is often regarded as an undesirable multiplicative noise. While speckle can be an obstacle for both visual interpretation and automatic image analysis, its distribution is precisely modeled thanks, in particular, to the pioneering work of (Goodman, 1976).

The distribution of the data can then be precisely taken into account in the processing methods, and it will be the case for the work we present in Part II and Part III. Therefore, it is important to review the acquisition process and the statistics of the observation that comes from it. It will be the topic of this chapter.

Organization of this chapter: in section 2.2 the acquisition process of SAR amplitude images is briefly recalled. As Part II is dedicated to the SWOT mission, we review in section 2.4 the specific characteristics of the Ka-band Radar Interferometer (KaRIn), SWOT's main instrument. Finally, we present the origin of speckle and the distribution of the data that will be used in the rest of this document in section 2.3.

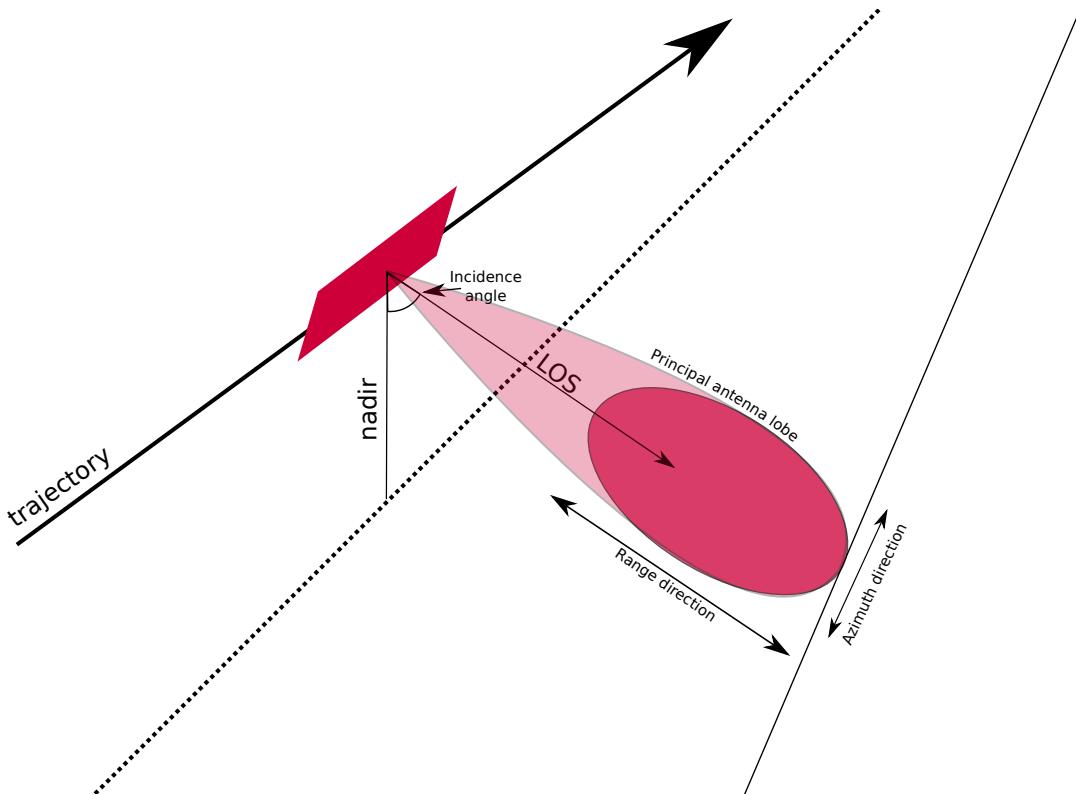


Figure 2.1 – Schematic view of the main geometrical parameters of a side-looking **RADAR** system.

2.2 Acquisition

The basic principle of **RADAR** imagery is to emit an electromagnetic wave with a given central frequency f_0 in a given direction called the **Line of Sight (LOS)** and measure the echo of the backscattered wave. The **LOS** is defined by two angles: the squint angle (the angle formed by the **LOS** and the perpendicular to the trajectory) and the incidence angle θ (the angle formed by the **LOS** and the nadir). When an electromagnetic wave is emitted, it illuminates the area on the ground located within the antenna lobes. The sensor then records the backscattered waves, along with the time since the emission. The principal antenna lobe, along with the incidence angle and the **LOS** are represented in Figure 2.1. The time delay between the emission and the reception of the wave gives the localization of the object on the ground in the range direction. By recording all the echoes coming from a given emission, it is possible to construct one line of the image in the range direction. After a given time step, the system can send a new electromagnetic wave and acquire a new line in the range direction. By repeating this process, the image is constructed.

Range resolution: The spatial resolution in the range direction along the **LOS** (provided that the signal is a chirp of central frequency f_0 and bounded by $[f_0 - \frac{B}{2}; f_0 + \frac{B}{2}]$) is given by:

$$\delta_r = \frac{c}{2B}, \quad (2.1)$$

where c is the speed of light. This resolution can be projected on the ground with respect to the incidence angle θ (making the hypothesis that the ground is flat):

$$\delta_{gr} = \frac{c}{2B \sin(\theta)} \quad (2.2)$$

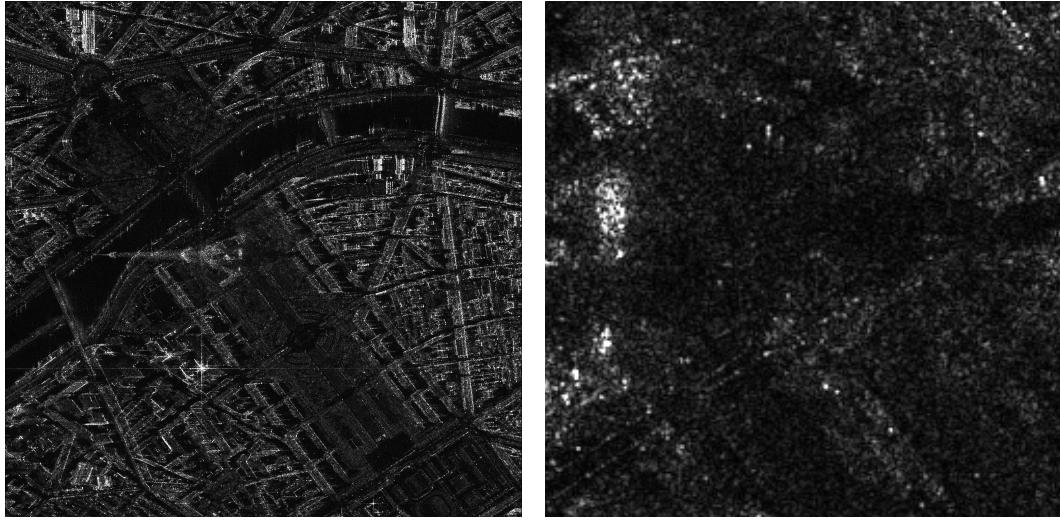
Azimuth resolution: For a real aperture imaging RADAR system, the resolution in azimuth depends on the wavelength λ_0 , the distance between the antenna and the object to be imaged R and the length of the antenna L :

$$\delta_{az} \propto \frac{\lambda_0 R}{L}, \quad (2.3)$$

Therefore, the larger the antenna, the better the resolution is. However, to obtain a resolution in the order of 10m, it would require an antenna whose length is in the order of the kilometer!

To cope with this problem, it is possible to exploit the fact that a point on the ground backscatter for several pulses as the satellite go forward. SAR systems use this property in order to simulate a much longer antenna, yielding a much better azimuth resolution. The process of synthesizing the image goes beyond the scope of this document, and we refer the reader to ([Deledalle et al., 2018](#)) for details about the technique. However, it is important to understand that the resolution will depend on the time period a point is present within the antenna pattern. It will also depend on the squint angle: this explains why a single sensor can provide data with different resolutions: in some acquisition modes, the squint angle is fixed, while in others it moves to obtain a trade-off between the spatial coverage of the image and the resolution.

A SAR image is complex with an amplitude and a phase for each pixel. Both the amplitude or phase differences between images can be exploited for different applications. In this work, we will deal with amplitude images. An example of such an image can be seen in [Figure 2.2](#). In this case, the satellite acquired the image from the left, the horizontal axis is the range direction and the vertical axis is the azimuth direction.



(a) Image of Paris

(b) Crop of the center of the image

Figure 2.2 – Amplitude **SAR** image of Paris acquired by TerraSAR-X in spotlight HR mode (resolution: $1m \times 1m$) between 01/24/2009 and 04/09/2010. Notice the strong variations due to the speckle.

2.3 Statistics of SAR data

Single-look complex image The speckle that can be observed in a **SAR** image (see [Figure 2.2](#)) is due to the coherent summing of the response of the multiple elementary targets in the resolution cell. This is illustrated in [Figure 2.3](#). To model the distribution of the backscattering vector resulting from the coherent summing of the contributions, ([Goodman, 1976](#)) proposed the fully-developed speckle model when the following hypotheses are met:

- the number of elementary targets in the resolution cell is large;
- the phase and the magnitude of each target are independent random variables;
- the phase of each target is uniformly distributed on $[-\pi; \pi]$.

In a stationary area, the large number of elementary targets makes it possible to apply the central limit theorem. Therefore, the resulting response z of a resolution cell follows a zero-mean complex circular Gaussian distribution of variance σ^2 which is a function of the normalized backscattering coefficient σ_0 :

$$p(z|\sigma^2) = \frac{1}{\pi\sigma^2} e^{-\frac{|z|^2}{\sigma^2}}. \quad (2.4)$$

The amplitude $v_A = |z|$ then follows a Rayleigh distribution of parameter $\mu_A = \sigma$:

$$p(v_A|\mu_A) = \frac{2v_A}{\mu_A^2} e^{-\left(\frac{v_A}{\mu_A}\right)^2}, \quad (2.5)$$

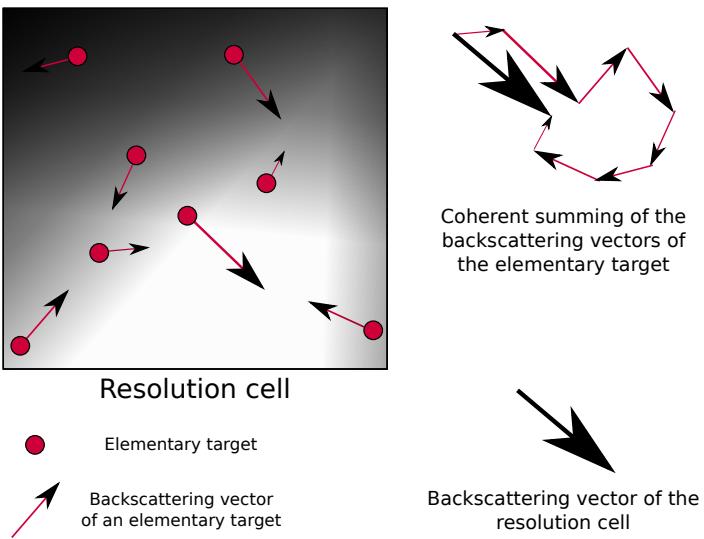


Figure 2.3 – Multiple elementary targets are generally present within a resolution cell, and each of them backscatter the signal in a different way. The resulting backscattered vector is the coherent sum of each of these individual responses.

and the intensity $v_I = |z|^2$ follows an exponential distribution of parameter $\mu_I = \sigma^2$. Since $v_I = v_A^2$, an amplitude image and an intensity image carry exactly the same information. In this document, we will use amplitude data. As explained in [chapter 3](#), many image models use the negative log-likelihood of an observation v_A , which gives:

$$-\log p(v_A|\mu_A) = 2\log(\mu_A) - \log(2v_A) + \left(\frac{v_A}{\mu_A}\right)^2. \quad (2.6)$$

An equivalent formulation, highlighting the multiplicative nature of speckle is to write:

$$v_A = \mu_A \times \xi, \quad (2.7)$$

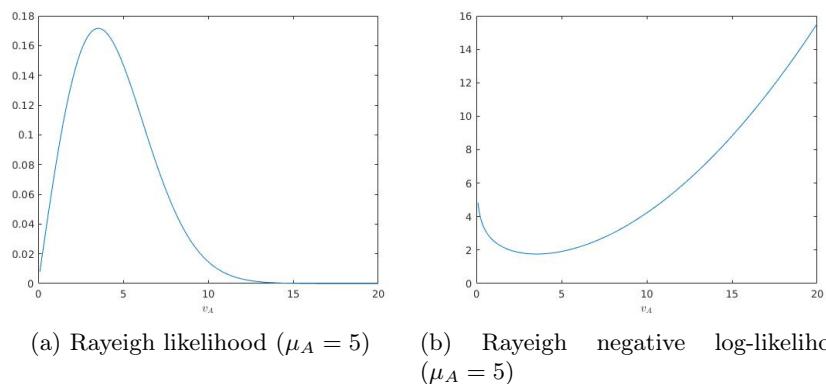


Figure 2.4 – Rayleigh likelihood and log-likelihood distributions. Note that the mode of the distribution is not μ_A .

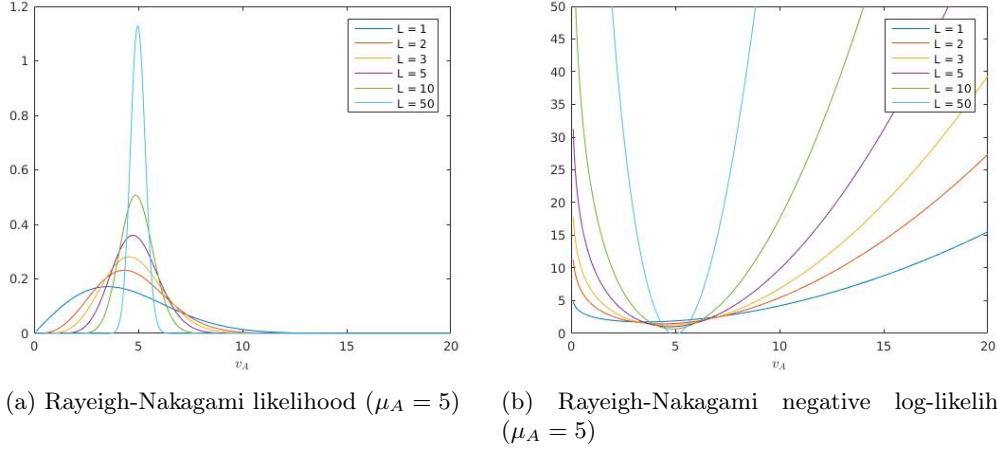


Figure 2.5 – Likelihood and log-likelihood distributions of Rayleigh-Nakagami. Note that the mode of the distribution is closer to μ_A when L is larger.

where ξ is a Rayleigh distributed random variable of parameter $\mu = 1$.

We show the distributions of the likelihood (Equation 2.5) and the negative log-likelihood (Equation 2.6) for an amplitude image with $\mu_A = 5$ in Figure 2.4.

Multi-looking It is possible to obtain a "multi-look" version of an image by averaging intensity samples in the spatial or temporal domain. For a homogeneous area, assuming fully developed speckle, the resulting intensity image then follows a Gamma distribution. When converting this image to amplitude, it follows a Rayleigh-Nakagami distribution:

$$p(v_A|\mu_A) = \frac{2\sqrt{L}}{\Gamma(L)\mu_A} \left(\frac{v_A\sqrt{L}}{\mu_A} \right)^{2L-1} e^{-\left(\frac{v_A\sqrt{L}}{\mu_A}\right)^2}. \quad (2.8)$$

Note that when $L = 1$ (that is, no multi-looking), Rayleigh-Nakagami is equivalent to the Rayleigh distribution. By taking the negative log-likelihood of this distribution, we get:

$$\begin{aligned} -\log p(v_A|\mu_A) &= -\log(2) - 2L \log(\sqrt{L}) + \log(\Gamma(L)) - (2L-1) \log(v_A) \\ &\quad + 2L \log(\mu_A) + L \left(\frac{v_A}{\mu_A} \right)^2. \end{aligned} \quad (2.9)$$

The multi-looking operation reduces the effect of the speckle as it can be seen in Figure 2.5 at a price of a coarser resolution.

One remarkable property of the Rayleigh-Nakagami distribution is that its coefficient of variation γ (which is the ratio between the standard deviation and the mean) only depends on the number of look (Jean-Marie Nicolas, 2012):

$$\gamma = \sqrt{\frac{\Gamma(L)\Gamma(L+1)}{\Gamma(L+\frac{1}{2})^2} - 1} \approx \frac{1}{2\sqrt{L}} \quad (2.10)$$

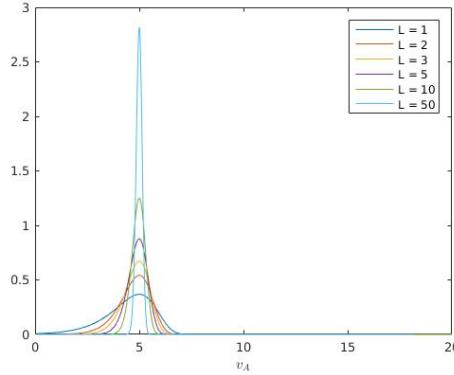


Figure 2.6 – Likelihood of the Fisher Tippett distribution for $\tilde{\mu}_A = 5$.

This reflects the multiplicative nature of the speckle: the variations will be more present in areas with a higher mean radiometry.

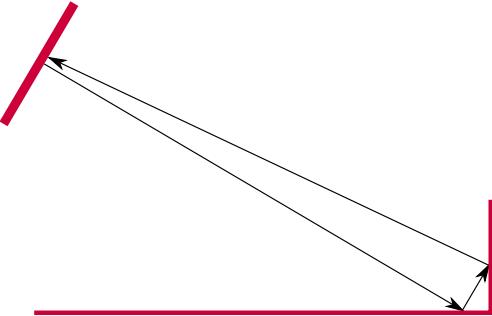
Logarithm of an image The speckle can be transformed to an additive contribution by applying a logarithm to the signal. The logarithm of the intensity $\tilde{v}_I = \log(v_I)$ then follows a Fisher-Tippett distribution of parameter $\tilde{\mu}_I = \log(\mu_I)$:

$$p(\tilde{v}_I | \tilde{\mu}_I) = \frac{L^L}{\Gamma(L)} e^{L(\tilde{v}_I - \tilde{\mu}_I)} \exp(-Le^{\tilde{v}_I - \tilde{\mu}_I}). \quad (2.11)$$

The distribution for different values of L is shown in Figure 2.6. The shape of this distribution features a left heavy-tail, as opposed to the Rayleigh-Nakagami which has a right heavy-tail. It can be considered close to a Gaussian distribution when the number of looks is large. It is then possible to apply methods developed for additive gaussian noise. However, a debiasing step is then required (Xie *et al.*, 2002a).

Strong scatterers We have seen that the backscattered vector for a resolution cell is the result of a coherent summing of the multiple targets, as illustrated in Figure 2.3. However, in some cases an elementary target has a backscattering value much higher than the others in the resolution cell. These situations typically arise where corner-type geometrical structures are present. In these cases, most of the emitted signal is backscattered towards the antenna. This is illustrated in Figure 2.7. Figure 2.7(a) shows a schematic view in the range direction and Figure 2.7(b) shows a corner reflector (which is a canonical target) on the field. As "near-perfect" corners are typically not present naturally, these strong scatterers are mostly present in urban areas, for man-made structures. In pixels where a strong scatterer is present, the statistics presented in this section are not the most relevant. While the backscattered vector is still the result of a coherent sum, the elementary target corresponding to the corner has a value much higher and is generally stable. Therefore, the resulting vector only fluctuates a little over time compared to pixels where no strong scatterer is present.

In the cases where a strong scatterer of cross section s is present in a homogeneous



(a) Schematic view of a corner resulting in a bright scatterer



(b) Corner reflector on the Eiffel Tower (Flora Weissgerber, ONERA) (Weissgerber, 2016)

Figure 2.7 – Corner reflector: most of the signal is reflected toward the antenna, resulting in a bright point.

region of mean intensity μ_I , it is recommended to use the Nakagami-Rice distribution (Tison and others, 2004):

$$p(v_I|\mu_I) = \frac{1}{\mu_I} \exp\left(-\frac{v_I + s^2}{\mu_I}\right) I_0\left(2\frac{\sqrt{v_I s^2}}{\mu_I}\right), \quad (2.12)$$

where I_0 is the first-kind modified Bessel function.

Summary: Statistics

In this section, we presented the statistics of the different SAR data that will be used in this document. The distributions are the following:

- Single-Look Complex (SLC): complex circular Gaussian
- amplitude of the SLC: Rayleigh;
- amplitude (multi-looked): Rayleigh-Nakagami;
- logarithm of the multi-looked amplitude: Fisher-Tippett;
- presence of strong scatterer s : Nakagami-Rice.

Name	Band	f_0 (in GHz)	wavelength (in cm)	incidence angle (in °)
SWOT	Ka	35.75	0.86	0.6 - 3.9
TerraSAR-X	X	9.65	3.11	15 - 60
Sentinel-1	C	5.4	5.55	20 - 46
Biomass	P	0.435	68.92	23+

Table 2.1 – Frequency and corresponding wavelength for popular SAR systems.

2.4 Particularities of SWOT

Besides its interferometric capacity, **SWOT**'s main instrument **KaRIn** has two main particularities: a high frequency ($f_0 = 35.75\text{GHz}$)/short wavelength (8.6mm) and a near-nadir incidence angle (from 0.6° to 3.9°). **SAR** systems generally use lower frequencies and higher incidence angle as it can be seen in [Table 2.1](#).

SWOT provides two modes: the **Low Rate (LR)** mode, dedicated to oceanography and the **High Rate (HR)** mode dedicated to hydrology. The **LR** mode features onboard unfocused **SAR** processing steps and a multi-looking to obtain a 1km^2 grid and reduce the output data rate. However, in this work, we are interested in the **HR** mode dedicated to hydrology. In this mode there is only an onboard pre-summing of 2 to keep a good resolution ($\sim 5\text{m}$ in azimuth, 70m to 10m in range) at the price of a high output data rate: $\sim 300\text{Mbps}$ ([Fjørtoft et al., 2010](#)). The difference in the range resolution between near-range (70m) and far-range (10m) is not surprising considering the ratio between the near-range and far-range incidence angle and [Equation 2.2](#).

Using Ka-band has the following implications:

- sensitivity to roughness at finer scales (which is useful to have signal from water surfaces);
- meteorological conditions will have a higher influence than for lower frequency systems;
- it features a weaker penetration into vegetation and snow.

With acquisition parameters close to those of **SWOT**, experiments have shown that the normalized backscattering coefficient σ_0 of the water is generally higher than the one of land at small incidence angles. ([Fjørtoft and others, 2014](#)) provides a review of such studies, along with new experiments. According to it, we can expect a water/land contrast of 10 to 20 dB in the range covered by **SWOT**, except for very low wind speeds. It should be noted that the **SNR** is very low in **SWOT** images due to power constraints as the system operates all the time. Therefore the signal from land surfaces is often at or below the thermal noise floor. The so-called noise-equivalent σ_0 is around 0dB for **SWOT** (in the center of the swath) whereas it is in the order of -20dB to -30dB for most space borne **SAR** systems. The σ_0 of water drops faster with incidence than the σ_0 of land. However, water is generally quite flat, whereas land surfaces may have strong topography.

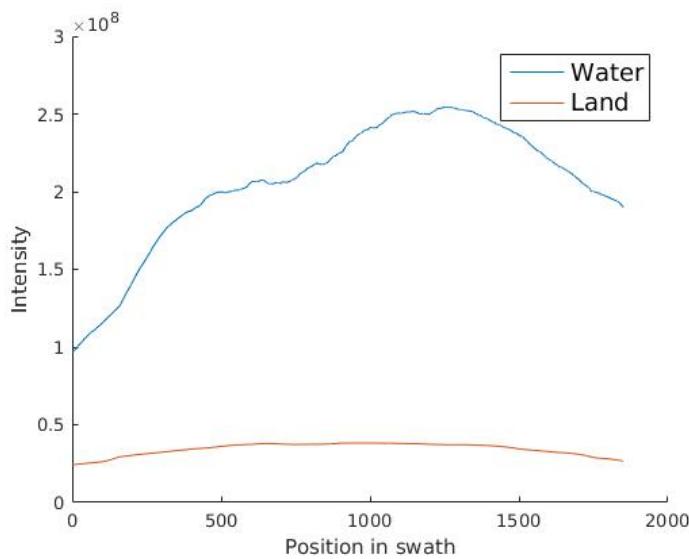


Figure 2.8 – Variations of the water and land parameters with respect to the range direction in a simulated SWOT image. Note that the variations are much stronger in the water class than in the land class.

Another type of variation in the observed radiometry comes from the antenna pattern. In all SAR systems, the radiometry change in the range direction according to a known pattern. This can be usually corrected (see (Bachmann and others, 2010) for the case of TerraSAR-X). However, in the case of SWOT, the land class is dominated by thermal noise that is not shaped by the antenna pattern. This evolution of the radiometry in the range direction is therefore not the same for both classes, as it can be seen in Figure 2.8. A global correction of the antenna pattern would yield variations in the land class. Therefore, the antenna pattern effect should not be corrected before land/water classification, but taken into account by the classification method.

The low incidence angle directly impacts the presence of layover. This phenomenon occurs as soon as the terrain slope exceeds the incidence angle. As the incidence angle used in SWOT is very low, we can expect a lot of layovers in the images. This is shown in Figure 2.9.

The effect of land/water layover on water detection is expected to be small, as the weak land backscattering adds to the strong water backscattering, making water even brighter. However land/land layover can result in areas with a strong radiometry, which can lead to false water detection.

The backscattering coefficient of the water surfaces also presents variations with respect to wind speed (as it has an effect on the surface roughness). If the wind speed is close to zero, there is little or no signal due to the specular reflection. From (Fjørtoft and others, 2014), it appears that a wind speed of $2.5m/s$ is the most favorable condition with a σ_0 approaching 20dB. For higher wind speeds, the σ_0 of water falls of

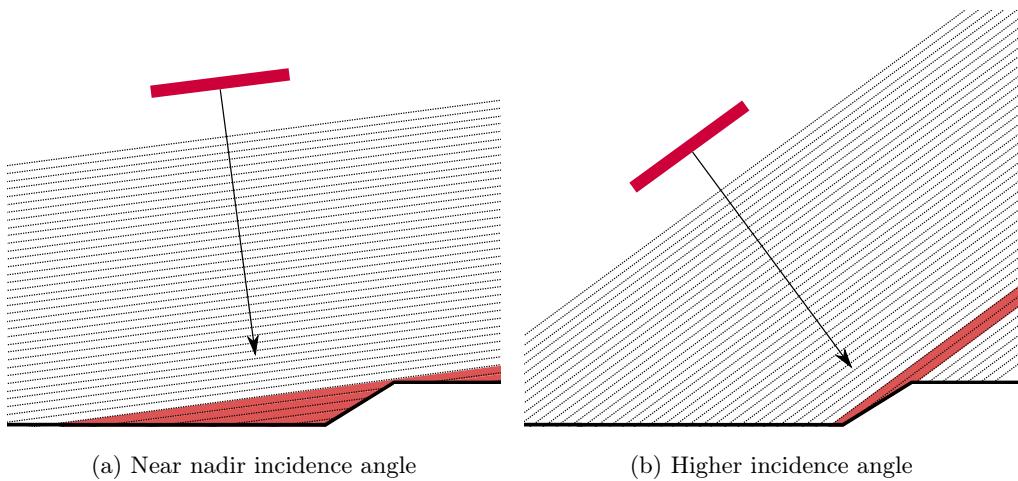


Figure 2.9 – Effect of the incidence angle on the extent of layover.

(in the order of 10dB at 10m/s).

To sum up, the variations in the radiometries will mainly depend on three elements:

1. the incidence angle, antenna pattern and **SNR**;
 2. the land topography (layover);
 3. the water roughness that is closely related to wind speed.

There exists *a priori* knowledge on the first two points. The incidence angle, antenna pattern and **SNR** are well known. Also, the topography can be approximately known from a **DEM**. However, the wind speed close to the water surface has strong local variations and cannot be precisely predicted from available meteorological data. The variations due to the incidence angle, antenna pattern and the topography are encoded in the so-called Xfactor. It links the theoretical σ_0 of the classes to the received signal using (in dB):

$$P_{\text{received}} = P_{\text{signal}} + P_{\text{noise}} \quad (2.13)$$

$$= \text{Xfactor} \times \sigma_0 + P_{\text{noise}}. \quad (2.14)$$

It should be noted that in the case of **SWOT**, water detection is actually not carried out on one single amplitude image, but rather on a combination of the images that are acquired simultaneously by the two antennas, mainly to reduce the thermal noise floor, and thereby increase the effective water/land contrast. The most straightforward way is to use the interferometric product, but it is slightly better to use the so-called coherent power, which is obtained by first phase-flattening the two **SLC** images with respect to a reference **DEM** (to approximately align their phases), then average them coherently, and take the power (square of the absolute value). For the latter, the gain in water/land contrast can be up to 3 dB (if the land σ_0 is at least 3 dB below the noise-equivalent σ_0 , and the phase-alignment is close to perfect). Taking the square root of

the interferometric product or the coherent power, we can assume the same marginal distributions as described for amplitude images in [section 2.3](#), and in what follows we will for simplicity refer to any of these images as amplitude images.

Summary: Particular characteristics of SWOT

As the objective of **SWOT** is to be able to do **SAR** interferometry on water (where there is usually almost no signal in conventional **SAR** images), it features unusual characteristics: a very low incidence angle (0.6° to 3.9°) and a high frequency (Ka band). Due to power constraints, it has a low **SNR**, especially over land. These characteristics have to be taken into account in the processing methods.

Chapter 3

Image models

3.1 Bayesian modeling

In this PhD, we consider an image formed by a finite set of sites $\mathcal{S} \subset \mathbb{Z}^d$, with d the dimension of the image (typically, $d \in \{2, 3\}$). To each site $s \in \mathcal{S}$ is associated a random variable U_s that has a value in E . E can be a set of labels (e.g. $E = \{1, \dots, q_{\max}\}$ when considering classification), or graylevels (e.g. $E = \{0, \dots, 255\}$ when considering discrete regularization, $E \in \mathbb{R}$ when considering continuous regularization). Note that the domain of E is independent of the one of the values v_s of the observed image itself, which will always be \mathbb{R}^+ in this document (as we are considering amplitude images with positive values).

An image processing task can generally be described as finding the "best" realization $\hat{\mathbf{u}} = \{\hat{u}_s\}$ of a random field \mathbf{U} . This is expressed as $\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} p(\mathbf{U} = \mathbf{u})$. In a similar way, we can define an observation $\mathbf{v} = \{v_s\}$ as the realization of a random field $\mathbf{V} = \{V_s\}$.

If we take into account an observation \mathbf{V} , the problem can be expressed as:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} p(\mathbf{U} = \mathbf{u} | \mathbf{V} = \mathbf{v}). \quad (3.1)$$

In Bayesian statistics, $p(\mathbf{U} = \mathbf{u} | \mathbf{V} = \mathbf{v})$ is called the posterior probability and can be linked to the likelihood $p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u})$ and a prior $p(\mathbf{U} = \mathbf{u})$:

$$p(\mathbf{U} = \mathbf{u} | \mathbf{V} = \mathbf{v}) = \frac{p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u})p(\mathbf{U} = \mathbf{u})}{p(\mathbf{V} = \mathbf{v})}. \quad (3.2)$$

The Maximum a posteriori (MAP) estimator of a problem can then be expressed as:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} p(\mathbf{U} = \mathbf{u} | \mathbf{V} = \mathbf{v}) = \arg \max_{\mathbf{u}} \frac{p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u})p(\mathbf{U} = \mathbf{u})}{p(\mathbf{V} = \mathbf{v})} \quad (3.3)$$

$$= \arg \max_{\mathbf{u}} p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u})p(\mathbf{U} = \mathbf{u}). \quad (3.4)$$

Note that when selecting the maximum value with respect to \mathbf{u} , $p(\mathbf{V} = \mathbf{v})$ is constant.

The model is then defined by the choice of the likelihood, which is related to the statistics of the data and the prior. The prior can either:

- be uniform; in this case, the solution is a **Maximum Likelihood Estimate (MLE)**;
- take into account statistics of the solution independently for each pixel;
- take into account neighbors at the pixel level and model spatial properties;
- describe the general shape of the solution.

Practical example of the **MLE** and the statistic prior and a review of different methods for the description of the general shape of the solution are given in [section 4.2](#). Taking into account the neighbors can be described by **MRF** which is explained in [section 3.2](#). We will see that while spatial interactions are only explicitly expressed at the pixel level, higher-level interactions are possible thanks to the Hammersley-Clifford theorem.

3.2 Markov Random Fields

Since the seminal work of Geman and Geman ([Geman and Geman, 1984](#)), **MRF** have become a classical way of modeling problems including spatial information. This formalism has proven to be efficient for image processing tasks that need to take into account spatial relationships (e.g. regularization, inpainting or classification ([Li, 2009](#))).

In this section, we go through the basis of **MRF**, and we review some optimization techniques.

3.2.1 Notation and definitions

To include spatial information in a model, the neighborhood of a site has to be defined:

$$\mathcal{N}_s = \{t\} \text{ with } \begin{cases} s \notin \mathcal{N}_s \\ t \in \mathcal{N}_s \implies s \in \mathcal{N}_t \end{cases} \quad (3.5)$$

We can now form the cliques: a clique is either a singleton in \mathcal{S}_i or a set of sites all neighbors with each others. Usual neighborhoods in 2D signals are illustrated in [Figure 3.1](#). In the following, the set of cliques is noted \mathcal{C} .

Local interactions between sites in a clique are modeled using clique potentials \mathcal{P}_c depending on the state of the sites. This allows for the definition of the global and local energy of the system for a given configuration:

$$\text{Global energy: } \mathcal{P}(\mathbf{u}) = \sum_{c \in \mathcal{C}} \mathcal{P}_c(u_s, s \in c) \quad (3.6)$$

$$\text{Local energy on site } s: \mathcal{P}_s(\mathbf{u}) = \sum_{c \in \mathcal{C}/s \in c} \mathcal{P}_c(u_s) \quad (3.7)$$

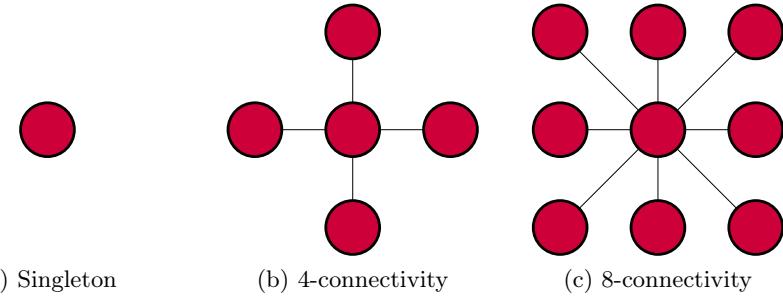


Figure 3.1 – Illustration of the usual neighborhoods for a single node (the central one) in the 2D case.

We can define the random process $\mathbf{U} = \{U_s\}$ as a MRF with $\mathbf{u} = \{u_s\}$ a realization of this process if and only if:

- $\forall \mathbf{u} \in \Omega$, $p(\mathbf{U} = \mathbf{u}) > 0$ (no forbidden configuration);
 - $\forall s \in \mathcal{S}, \forall \mathbf{u} \in \Omega$, $p(U_s = u_s | \{U_t = u_t, t \neq s\}) = p(U_s = u_s | \{U_t = u_t, t \in \mathcal{N}_s\})$ (contextual information is contained only in the neighbors).

This means that \mathcal{U} is a MRF if and only if the local conditional probability in one given site only depends on the configuration of its neighbors.

Using the Hammersley-Clifford theorem, we have an equivalence between a Gibbs field (the probability of a given configuration depends of the sum of cliques potentials) and a MRF. This allows to write:

$$p(U_s = u_s | \{U_t = u_t, t \neq s\}) = \frac{\exp(-\mathcal{P}_s(u_s | \{u_t, t \in \mathcal{N}_s\}))}{\sum_{e \in E} \exp(-\mathcal{P}_s(e | \{u_t, t \in \mathcal{N}_s\}))}. \quad (3.8)$$

In the image processing tasks that are considered in this document (classification, regularization), we start from an observation \mathbf{v} (the image to be classified, the image to be regularized) that we can model as a realization of a MRF \mathbf{V} and we want to estimate a result \mathbf{u} (the classification, the regularized image) that we will model as a realization of another MRF \mathbf{U} .

In order to link these two processes, we saw in section 3.1 that one can use the MAP criterion with Bayes theorem:

$$p(\mathbf{U} = \mathbf{u} | \mathbf{V} = \mathbf{v}) \propto p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u}) p(\mathbf{U} = \mathbf{u}). \quad (3.9)$$

As U is supposed to be a MRF, we have:

$$p(\mathbf{U} = \mathbf{u}) = \frac{\exp(-\mathcal{P}(\mathbf{u}))}{Z}, \quad (3.10)$$

with Z a normalization factor. From Equation 3.9:

$$p(\mathbf{U} = \mathbf{u} | \mathbf{V} = \mathbf{v}) = \frac{\exp(-(-\log p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u}) + \mathcal{P}(\mathbf{u})))}{K} = \frac{\exp(-\mathcal{E}(\mathbf{u}))}{K}, \quad (3.11)$$

where $\mathcal{E}(\mathbf{u}) = -\log p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u}) + \mathcal{P}(\mathbf{u})$ is the energy of the system. It is a sum of the negative log-likelihood of the data $-\log p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u})$ and a prior $\mathcal{P}(\mathbf{u}) = -\log(p(\mathbf{U} = \mathbf{u}))$. The negative log-likelihood is also called the data fidelity term, and will also be noted $DT(\mathbf{v}, \mathbf{u})$ in the following. From [Equation 3.11](#) we can see that maximizing $p(\mathbf{U} = \mathbf{u} | \mathbf{V} = \mathbf{v})$ is equivalent to minimizing the energy $\mathcal{E}(\mathbf{u})$.

Summary: Modelizing a problem using MRF

To sum up, using [MRF](#) to solve image processing problems boils down to the following energy minimization:

$$\mathcal{E}(\mathbf{u}) = DT(\mathbf{v}, \mathbf{u}) + \mathcal{P}(\mathbf{u}). \quad (3.12)$$

The term $DT(\mathbf{v}, \mathbf{u})$ describes the acquisition process which depends on the imaging system and the term $\mathcal{P}(\mathbf{u})$ is a prior on the result depending on the solution we are seeking (i.e., this term is application dependent). The prior at one site s should only involve s and its neighbors.

In the following, the notation of $p(\mathbf{V} = \mathbf{v} | \mathbf{U} = \mathbf{u})$ and $p(\mathbf{U} = \mathbf{u})$ will be shortened to $p(\mathbf{v} | \mathbf{u})$ and $p(\mathbf{u})$ respectively. The next step is to minimize the energy $\mathcal{E}(\mathbf{u})$. In the next subsection, we review methods dedicated to this task.

3.2.2 Optimization

Optimization of \mathcal{E} is not straightforward. The energy may be non-convex, and even when the problem is discretized, the number of solutions is $|E|^{|S_i|}$ preventing the use of exhaustive search approaches on any real problem.

Two techniques are widely used to optimize [MRF](#): [Iterated Conditional Modes \(ICM\)](#) and [Simulated Annealing \(SA\)](#). We also present methods based on graph-cuts that work for certain types of [MRF](#).

Iterated Conditional Modes (ICM) In this method, a solution is found by iteratively selecting the values that give the lowest energy at each pixel based on the current configuration of the image. This method is described in [algorithm 1](#).

Algorithm 1: Iterated Conditional Modes

```

Data: Observed image  $v$ 
Result: Model  $u$ 
 $u = \text{initialize}();$ 
changed = true;
while changed do
    changed = false;
    foreach  $u_i$  in  $u$  do
        prev =  $u_i$ ;
        /* Select  $u_i = q \in E$  s.t.  $\mathcal{E}(u)$  is the smallest. */
         $u_i = \text{select\_optimal\_at\_i}(u, v);$ 
        if prev  $\neq u_i$  then
            changed = true;

```

While this method is simple, it only gives a local minimum, highly dependent of the initialization.

Simulated annealing (SA) This method, inspired by the annealing process in metallurgy involving a phase of heating followed by a controlled cooling (Geman and Geman, 1984), uses a temperature that decreases during the algorithm. At each step, a new solution is proposed. But instead of choosing the solution that will minimize the energy as for the ICM, it will be chosen according to a sampler (either Gibbs or Metropolis) taking into account the temperature. Their behavior is close to an uniform sampling when the temperature is high while a low temperature will lead to the selection of the solution minimizing the local energy. Doing so makes it possible to escape from local minima: during the first iterations, it explores the space search, only to converge toward a solution when the temperature decreases. The general principle of the algorithm is described in [algorithm 2](#).

Algorithm 2: Simulated Annealing

```

Data: Observed image  $v$ 
Result: Model  $u$ 
 $u = \text{initialize}();$ 
changed = true;
temp = initial_temp;
while changed do
    changed = false;
    foreach  $u_i$  in  $u$  do
        prev =  $u$ ;
         $u_i = \text{select\_element\_in}(E, temp);$ 
        /* This step can decrease the energy */
        temp = decrease_temp(temp);

```

While there is a theoretical guarantee to find the global minimum, it is only for

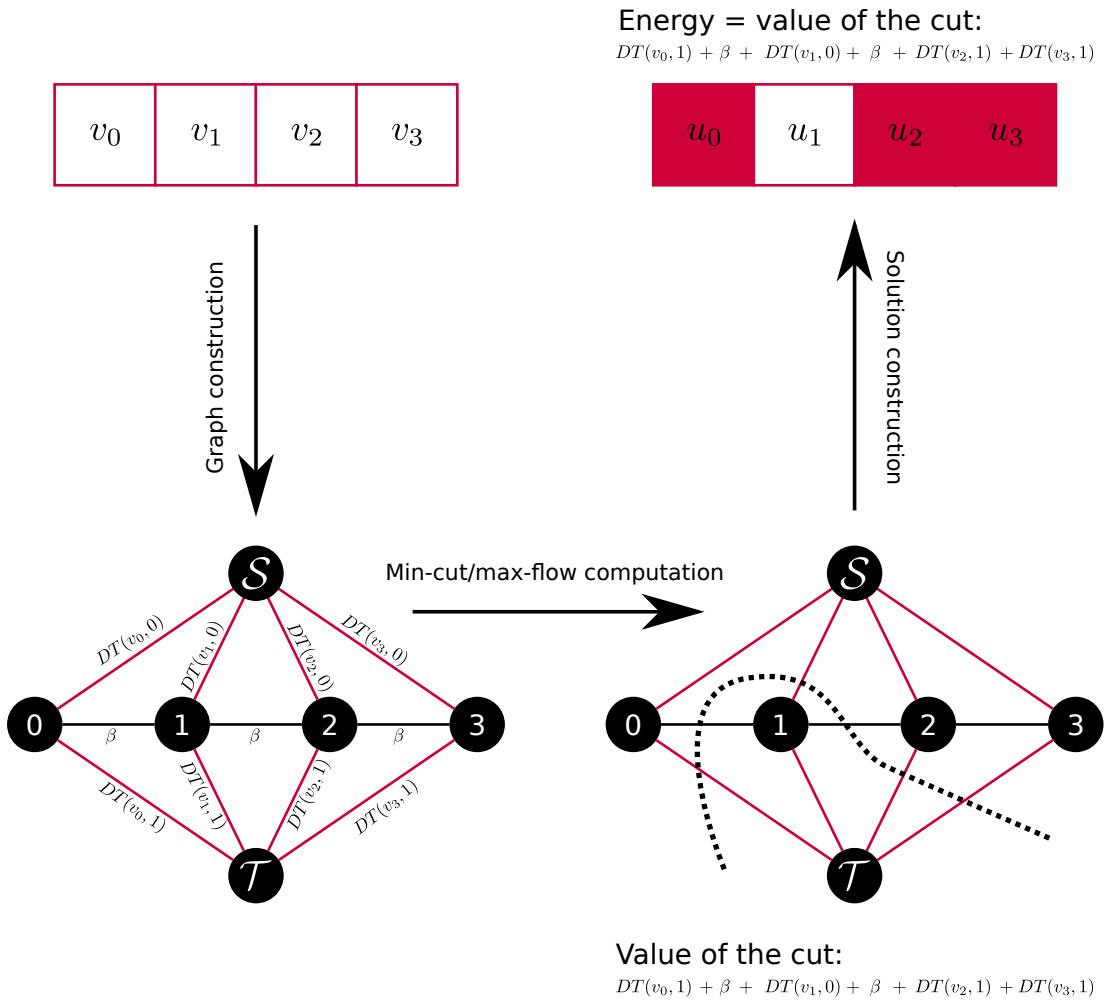


Figure 3.2 – Illustration of the correspondence between the cut and the energy in a simple 1D case with the binary Ising model.

an infinitely slow cooling scheme. In practice, we do not have a guarantee to find the global optimum. Furthermore, depending on the cooling scheme, it can take a long time to find a solution.

Graph-cuts In practice, we do not have the guarantee neither with **ICM** nor with **SA** to obtain the global minimum. To address this problem, several methods using graph-cuts have been developed. The general principle is to build flow networks (that is a graph with two terminal nodes: the source \mathcal{S} and the sink \mathcal{T} such that there is a bijection between the set of \mathcal{S} - \mathcal{T} possible cuts in the graph and the set of solutions of our problem. An \mathcal{S} - \mathcal{T} cut in a flow graph can be viewed as a partitioning of the nodes in the graph in two subsets: one which contains the source, and one which contains the sink. The weights on the edges are set so that the value of a cut corresponds to the energy of the solution for the original problem. This correspondence makes it possible to find the global optimum for some energies using a min-cut/max-flow algorithm. The global process is illustrated in Figure 3.2 for a simple example and the Ising model

defined in [Equation 3.14](#). In the following we will only consider priors on 2 sites, that we will note $\psi(x, y)$ i.e.:

$$-\log(p(\mathbf{U} = \mathbf{u})) = \beta \sum_{\{s,t\} \in \mathcal{C}} \psi(u_s, u_t) \quad (3.13)$$

Binary graphcuts The first approach, introduced in ([Greig and others, 1989](#)), finds the global optimum of a [MRF](#) with binary output $|E| = 2$ given that the prior is the Ising model:

$$\mathcal{E}(\mathbf{u}) = \sum_s \text{DT}(v_s, u_s) + \sum_{(s,t)} \beta \psi(u_s, u_t) \quad (3.14)$$

with:

$$\psi(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases} \quad (3.15)$$

Any \mathcal{S} - \mathcal{T} cut in the graph presented in [Figure 3.2](#) gives a solution u_0 with the value of the cut corresponding to its energy $\mathcal{E}(u_0)$. Therefore, finding the min-cut in this graph will give the exact solution to our minimization problem. It is then possible to use one of the several algorithms developed to find min-cut in graphs.

In ([Kolmogorov and Zabin, 2004](#)), an extension of ([Greig and others, 1989](#)) to any sub-modular prior ($\psi(0, 0) + \psi(1, 1) \leq \psi(0, 1) + \psi(1, 0)$) has been proposed. It uses the additive theorem saying that the sum of two graph-representable functions is graph-representable.

Graphcuts for multi-label problems Depending on the prior, several graphcut-based methods have been developed:

- if the prior is a semi-metric: $\alpha - \beta$ -swap ([Boykov et al., 2001](#)), giving an approximate solution.
- if the prior is a metric: α -expansion ([Boykov et al., 2001](#)), which has a lower complexity than $\alpha - \beta$ -swap, but still gives an approximate solution.
- if the prior is convex: ([Ishikawa, 2003](#)) gives the exact solution provided that E is finite.

In the following, we describe the method introduced by ([Ishikawa, 2003](#)) in the case of the anisotropic Total Variation prior. Introduced by ([Rudin et al., 1992](#)), [TV](#) corresponds to $\psi(u_s, u_t) = |u_s - u_t|$ and penalizes variations between the values of neighboring pixels while allowing for sharp edges. It will be used in [chapter 6](#).

The graph construction is represented in figure [3.3](#). The graph is composed of several layers, each of them having one node for each of the m pixels of the image. Each layer represents a possible value for the regularized image. Neighboring nodes are connected by pairs of arcs. The source (\mathcal{S}) and the sink (\mathcal{T}) are also connected to the first and

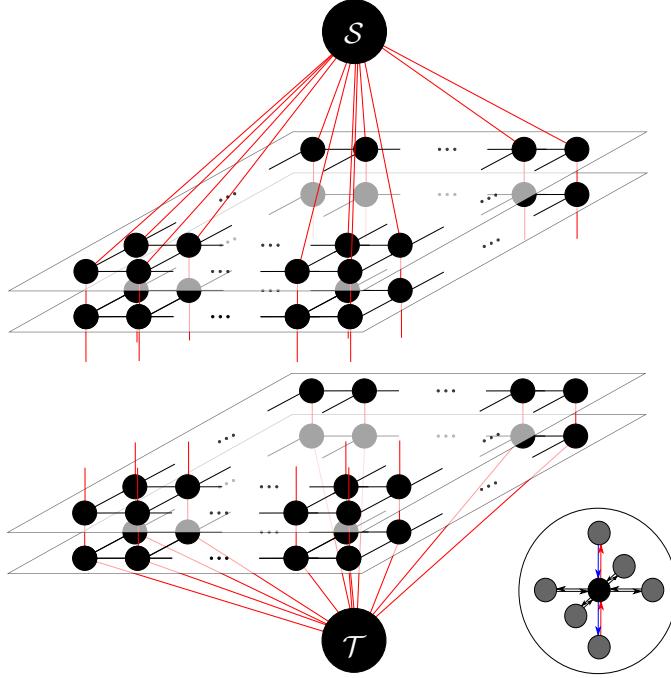


Figure 3.3 – Graph construction based on (Ishikawa, 2003).

last layers. Each node is connected with a weight of $(q_{n+1} - q_n)\beta$ for horizontal (black) arcs (see zoom in (3.3)), $DT(v_i, q_n)$ for top-down (blue) arcs (with v_i the value of the pixel corresponding to the node being defined), and ∞ for bottom-up (red) ones, where q_1, \dots, q_{max} are the possible values for u_s .

An \mathcal{S} - \mathcal{T} cut in the graph corresponds to a unique solution: each pixel will have exactly one vertical edge in the cut, giving the corresponding inferior value to this pixel. Furthermore it can be shown that the value of a cut corresponds to the energy of the associated solution. The likelihood in one pixel is given by the vertical edge, while the prior penalization is ensured by horizontal ones. Therefore, finding the \mathcal{S} - \mathcal{T} cut of minimal cost in this graph gives the optimal solution.

Min-cut algorithms At the core of all graphcut-based optimization techniques is the computation of the min-cut. Indeed, the graphcut construction can be seen as a simple reformulation of the original problem, with the optimization to be done. While out of the scope of this thesis, we present the general principle and a few algorithms that can be used for the min-cut computation. An extend review of classical algorithms solving this task can be found in (Boykov and Kolmogorov, 2004) These algorithms are often called min-cut/max-flow algorithms as the max-flow min-cut theorem states that the value of the minimum cut in a flow graph is also the value of the maximum flow that can go from the source to the sink. The flow of a path from \mathcal{S} to \mathcal{T} is a number which

is equal to the minimum weight among the edges of the path. The maximum flow is then the maximum value for all the paths from the source to the sink.

Two main classical algorithms try to find the maximum flow:

- In (Ford and Fulkerson, 1956). It is based on augmenting paths, that is paths with available capacity on all edges of the path (or not saturated). It iteratively selects augmenting paths and saturates them. The crucial step is to find augmented paths which is done with a residual graph. The computational cost of the classically used variant of (Edmonds and Karp, 1972) is in $O(VE^2)$ where V is the number of vertices and E the number of edges.
- In (Goldberg and Tarjan, 1988), another type of algorithm (named "push-relabel") is proposed. It also works on a residual graph, but instead of finding paths with available capacity, it only works on one vertex at a time by looking at its neighbors. Its computational cost is in $O(V^2E)$ for the basic version making it more efficient than the Ford–Fulkerson algorithm. Efficient selection rule can lead to a $O(V^2E^{\frac{1}{2}})$ complexity.

Most max flow algorithms are based on one of these algorithms, which is also the case of (Boykov and Kolmogorov, 2004) (based on augmented paths). It has a complexity of $O(EV^2|C|)$ where $|C|$ is the value of the minimum cut. While it is theoretically not as efficient as previous approaches, it is in practice faster for image processing applications. Indeed, it exploits the particular shape of the graph built for an image, which is very regular. Several improvements to this version have been proposed; for instance, (Liu and Sun, 2010) propose a parallel implementation.

In the following, results requiring the computation of a minimum flow have been computed using the implementation provided by (Boykov and Kolmogorov, 2004).

Summary: Optimizing MRF

The optimization of MRF is a widely studied problem. Along with classical methods such as ICM and SA, more recent works propose to use graph theory to this end.

When using classical models for the prior (such as Ising in the binary case or TV in multi-label problems), these methods lead to a global optimum for quantized values.

Part II

Water/land classification

For hydrological applications, the main objective of the **SWOT** mission is to estimate water elevation. It uses single-pass **SAR** interferometry (thanks to its two antennas) for this purpose. However, estimating water elevation from the interferometric data requires to know where the water is. Water detection in **SWOT** images is the object of this part.

Our goal is to use a **SWOT** amplitude image \mathbf{v} to obtain a map of water areas \mathbf{u} where:

$$u_i = \begin{cases} 1 & \text{if } i \text{ should be classified as water and} \\ 0 & \text{if } i \text{ should be classified as land.} \end{cases}$$

This task of binary classification of an amplitude image has been widely studied and a state of the art is presented in [section 4.2](#). However, in the case of **SWOT**, there are two specific problems that need to be addressed:

- The water and land classes parameters are not constant throughout the image. While most classification algorithms assume constant parameters, this hypothesis does not hold in the case of **SWOT**. We propose an adaptation of classical algorithms in order to take into account non-constant parameters and several methods to estimate them in [chapter 4](#).
- Many classification algorithms consider spatial regularity as a prior on the result, in order to limit the effect of speckle that is present in **SAR** images. However in the case of hydrological applications we have to detect rivers which are narrow curvilinear objects, violating the hypothesis of spatial compactness. While [chapter 4](#) presents general methods for classification that are adapted to the detection of lakes and other large water bodies, the detection of narrow rivers is studied in [chapter 5](#).

Chapter 4

Water/land classification

4.1 Introduction

Binary classification, which is closely related to detection, is a widely-studied problem and it will be the topic of [section 4.2](#) where a survey of methods dedicated to this task is presented with the exception of the [MRF](#).

Most classification algorithms assume that a class can be represented by a single distribution with one or several parameters. In our case (amplitude data), a class would be represented by a Rayleigh-Nakagami distribution (see [section 2.3](#)) parametrized by its mean reflectivity μ and its number of looks L . In this chapter, we will consider L to be given by the acquisition process and constant through the image (which is typically the case in synthesized but unprocessed data). On the other hand, μ depends on the scattering properties of the class, and the assumption that it should be considered constant through the image does not hold in the case of [SWOT](#) because of the antenna pattern, the low [SNR](#) and local variations in water roughness.

Our goal is therefore to have a spatially varying parameter for each class at each pixel of the image. This can be represented by as many images as the number of classes. In the case of [SWOT](#), we have two classes, and we will therefore have two images μ_0 and μ_1 taking into account these spatial variations instead of two scalars μ_0 and μ_1 in usual algorithms. While using these non-constant parameters is straightforward (we just have to replace the constant value in the likelihoods by the indexed value of μ_0 or μ_1), their estimation is not trivial and will be covered in [section 4.4](#).

Results of the different proposed methods are shown and discussed in [section 4.5](#).

4.2 Binary classification

In this section, we present existing methods that can be used for the classification of SAR images. We will cover widely used methods from the literature, with the exception of Markovian approaches which will be covered in [section 4.3](#).

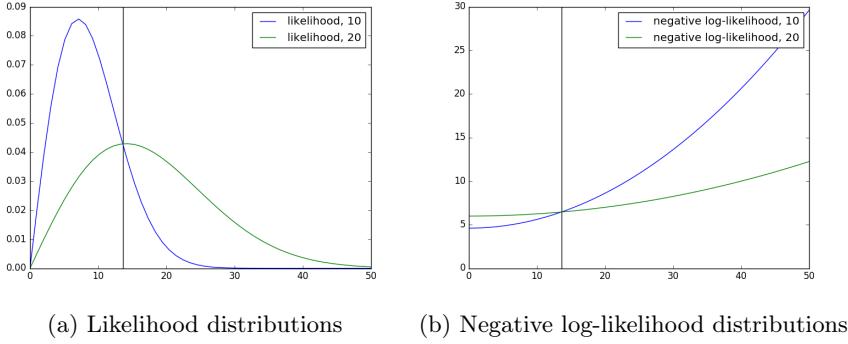


Figure 4.1 – Distribution of the likelihoods and the negative log-likelihoods for Rayleigh-Nakagami distributed variables with $L = 1$ and $\mu_0 = 10$, $\mu_1 = 20$.

Maximum likelihood estimation In a Bayesian framework, the simplest solution to find a binary map \mathbf{u} indicating water from an observation \mathbf{v} is to use the maximum likelihood estimate, that is:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} p(\mathbf{v}|\mathbf{u}). \quad (4.1)$$

This is equivalent to:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} -\log(p(\mathbf{v}|\mathbf{u})), \quad (4.2)$$

which becomes, when considering the data separable:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \sum_i -\log(p(v_i|u_i)). \quad (4.3)$$

In this case, we only need to define the negative log-likelihood. In [section 2.3](#), we have seen that in the case of amplitude images and assuming locally homogeneous radar reflectivity, the data follows a Rayleigh-Nakagami distribution. Hence, we can easily compute for each pixel $-\log(p(v_i|0))$ and $-\log(p(v_i|1))$ (for our specific case of binary classification) and choose the solution giving the minimum value. As it can be seen in [Figure 4.1](#), this test results in a simple thresholding (where the threshold is set where the two distributions are crossing). In this case, the threshold v_{th} can be computed (using the simplification of the likelihood introduced in [section 2.3](#)):

$$2L \log(\mu_0) + L \left(\frac{v_{th}}{\mu_0} \right)^2 = 2L \log(\mu_1) + L \left(\frac{v_{th}}{\mu_1} \right)^2 \quad (4.4)$$

$$v_{th} = \sqrt{\frac{2 \log \left(\frac{\mu_0}{\mu_1} \right)}{\frac{1}{\mu_1^2} - \frac{1}{\mu_0^2}}} \quad (4.5)$$

We show a simple example using this method in [Figure 4.2](#). We can see that this method gives poor results in the case of amplitude [SAR](#) data. The strong fluctuations of speckle for low L values prohibits the use of direct maximum likelihood estimations.

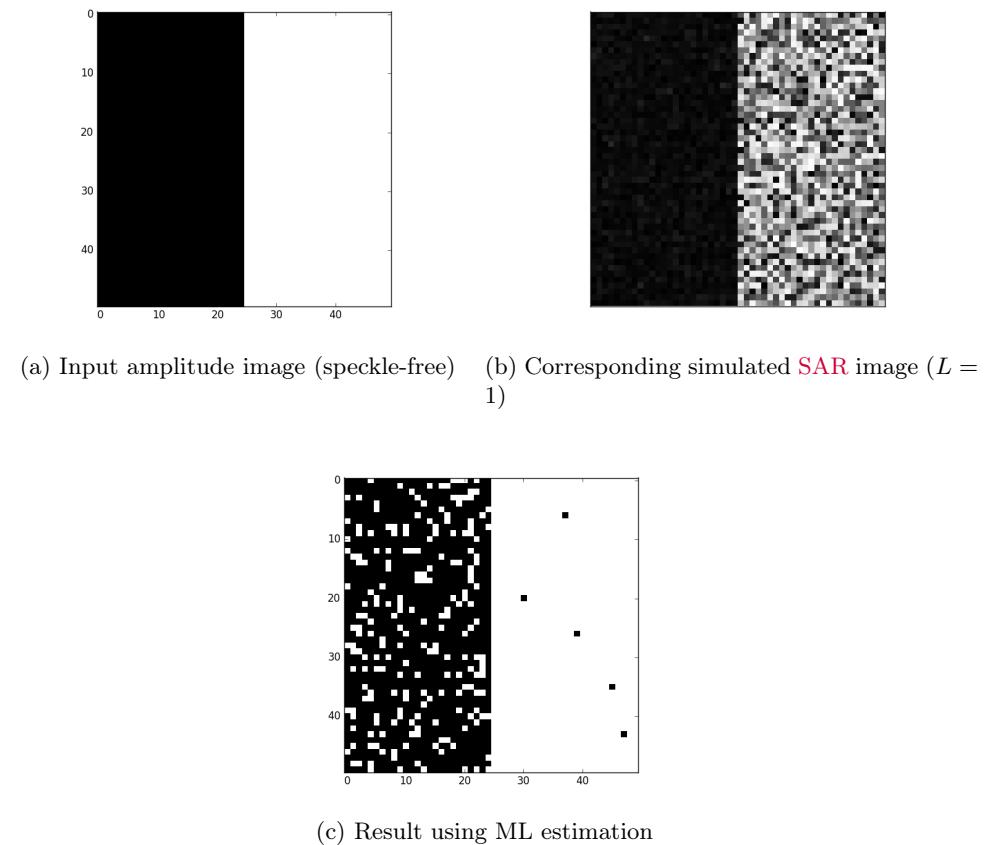


Figure 4.2 – Example on a simple simulated **SAR** image. The two classes are Rayleigh-Nakagami distributed with parameters $\mu_0 = 10$ and $\mu_1 = 20$ and $L = 1$. Using the maximum likelihood estimation gives a noisy result.

MAP To cope with the fluctuations due to speckle, classification methods generally use prior information on the distribution of the data. This information can either be spatially independent (e.g. statistics on the distribution of the data), or take into account the spatial distribution of the data. Integrating prior in the decision is well expressed by the **MAP** formulation. Trying to find \mathbf{u} using \mathbf{v} can be expressed as trying to find the model maximizing the posterior:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} p(\mathbf{u}|\mathbf{v}) \quad (4.6)$$

Using Bayes' theorem, this gives:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} \frac{p(\mathbf{v}|\mathbf{u})p(\mathbf{u})}{p(\mathbf{v})} \quad (4.7)$$

$$= \arg \max_{\mathbf{u}} p(\mathbf{v}|\mathbf{u})p(\mathbf{u}) \quad (4.8)$$

Where $p(\mathbf{u})$ is a prior on the desired solution. This is known as the **MAP** estimation. Note that when no *a priori* information is available, $p(\mathbf{u})$ is a constant distribution and [Equation 4.8](#) is then equivalent to the **MLE** introduced in [Equation 4.1](#). However, as the **MLE** generally gives noisy results in the case of SAR images, it is generally interesting to add a prior. Two mains types of priors can be used:

- spatial priors when information about the spatial distribution of the classes is available. In this case, the most common method is to use **MRF** introduced in [section 3.2](#). This approach will be further described in [section 4.3](#).
- statistical priors which will be covered in this paragraph: if we have some knowledge on the prior probabilities of the classes, then the threshold can be improved.

We explain how statistical priors can be encoded in the **MAP** framework for the application we are interested in in this chapter: water/land classification. It is estimated that the continental water represents about 2.5% of the total continental surface of Earth. Therefore, it can be interesting to use this statistic when detecting water. Assuming the distribution of $p(u)$ can be considered separable, this can be expressed at each pixel as

$$p(u_i = 0) = 0.975 \quad (4.9)$$

$$p(u_i = 1) = 0.025. \quad (4.10)$$

When taking the negative log likelihood of [Equation 4.8](#) and assuming that the likelihood distribution is separable we have:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} - \sum_i (\log(p(v_i|u_i)) + \log(p(u_i))), \quad (4.11)$$

which can be solved pixel by pixel by comparing the values of $-(\log(p(v_i|u_i = 0)) + \log(p(u_i = 0)))$ and $-(\log(p(v_i|u_i = 1)) + \log(p(u_i = 1)))$. As for

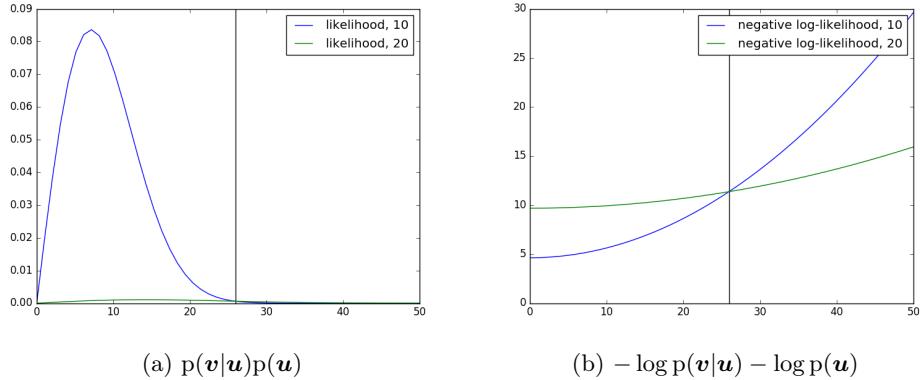


Figure 4.3 – Distribution of the energies using the statistical priors of the water distribution.

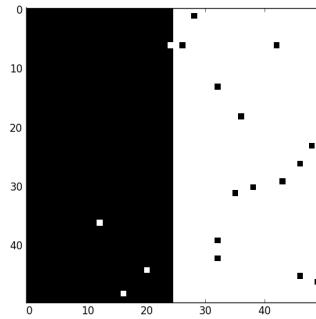


Figure 4.4 – Result on the dataset presented in Figure 4.2 for the statistical MAP estimation.

the MLE seen in this section, this can be expressed as a simple threshold:

$$t_{\text{stat}} = \sqrt{\frac{\log\left(\frac{\mu_2^2 p(1)}{\mu_1^2 p(0)}\right)}{\frac{1}{\mu_1^2} - \frac{1}{\mu_0^2}}}, \quad (4.12)$$

which can be interpreted as a shift compared to the distributions shown in Figure 4.2. This is shown in Figure 4.3.

In practice, this means that we are less likely to wrongly classify a "land" pixel as water, but it will be more likely to miss "water" pixels. Using the threshold presented in [Equation 4.12](#), result on the dataset presented in [Figure 4.2](#) can be seen in [Figure 4.4](#). It shows that indeed the number of "land" pixels classified as "water" pixels are less numerous, while the number of "water" pixels classified as "land" pixels has increased. This minimizes the total error rate (as there is much more land than water), but it is dangerous when water is the principal object of interest. In an extreme case, a classifier which would label every pixel as land would obtain an overall error rate of only 2.5%.

We have seen that pixel-based classification, with or without the use of priors, can not deal efficiently with the fluctuations of SAR images due to speckle. To this end it

is necessary either to limit the effect of the speckle with a first step of denoising, or to have priors on the spatial distribution of the classes. In the following paragraphs, we present methods using a pre-processing step of denoising or segmentation.

Denoising as a first step As pixel-based classification can not efficiently deal with the variations induced by speckle, a straightforward way to solve the problem is to apply a method to reduce these variations as a pre-processing. Pixel-based classification can then be applied. In (Liu and Jezek, 2004), a Lee filter (Lee, 1981) is first applied to reduce the speckle followed by an anisotropic diffusion algorithm (Perona and Malik, 1990). The image is then thresholded assuming that the data follows a Gaussian distribution. A similar method is used by (Cazals *et al.*, 2016) which uses the Perona-Malik filter of (Perona and Malik, 1990) followed by an hysteresis thresholding to detect flooding on Sentinel-1A data. While the objective is not only to remove the effects of speckle, (Cao *et al.*, 2011) use a multi-scale approach to detect water in simulated SWOT images with a technique similar to the one presented in chapter 5 for the detection of narrow rivers in SWOT images. It is based on the assumption that any land water element can be considered as thin at a sufficiently large scale. Note that state of the art denoising methods for SAR images such as the one presented in (Deledalle *et al.*, 2015) could be used and provide the advantage, in addition to generally better speckle removal, to preserve the statistics of the signal (as it will multilook the signal, the distribution will only change on the number of looks L , see section 2.3). This property can be used for the classification afterwards, which will be done in subsection 5.2.2.

Segmentation as a first step Another way to address the challenge of speckle for pixel-based classification, is to perform a segmentation as a first step. In this case, the critical step is generally the segmentation, as it should neither output regions containing both classes, nor too over-segment. Once a good segmentation has been found (that is, sufficiently large regions containing only one class so that statistics can be reliably estimated) the actual classification can be made based on these statistics. In a first approach, this problem can be solved using edge detection. Generally, edge detection algorithms are based on the gradients of the images, but the multiplicative nature of the speckle present in SAR images prevents from having good performances using differences. (Touzi *et al.*, 1988) and (Fjørtoft *et al.*, 1998) tackle this problem by using a ratio operator which is adapted to the presence of speckle. When using this kind of method, it is required to apply a method extracting closed and skeleton edges, defining a segmentation, for instance using the watershed algorithm.

In (Silveira and others, 2009) the classical formulation of (Chan and Vese, 2001) is adapted to SAR images for water segmentation using log-normal distributions to model the statistics. Chan-Vese method belongs to a family of methods called "active contour", where the goal is to modify the shape of the segmentation based on an energy taking into account its quality (based on the statistics of the images computed region-wise or along the boundaries) and its fidelity to a prior model on its shape. In the case

of (Chan and Vese, 2001), its formulation is a sum of terms enforcing three properties:

- the length of the curve is penalized (to enforce regularity);
- the area inside the curve is controlled to obtain a desired size;
- the radiometry inside and outside the curve should be separated to two constant values (which are found during the optimization process).

It is based on a level set formulation. In the case of SAR images, this is not adapted as it assumes a Gaussian distribution of the noise. Level set models adapted to SAR images have been presented in (Ben Ayed *et al.*, 2005) in the case of intensity images, or in (Ben Ayed *et al.*, 2006) for polarimetric SAR images. An attempt to use the formulation of (Ben Ayed *et al.*, 2005) for water/land classification has been made in (Silveira and others, 2009) without success as the assumption that both classes are homogeneous in intensity does not hold. This is indeed a problem that we encounter in the case of SWOT images, and which is tackled in this chapter.

Another family of active contour models, snake-based methods, explicitly define the limits of the regions by a set of points which are moved to fit a model. They have also been applied to SAR in (Chesnaud *et al.*, 1999). In this approach, the authors consider the statistics of the region delimited by the curves instead of those of the curve itself. The motivation behind this is that when using SAR images, boundaries are not well defined. Once again, this kind of approach is not well suited when each region can not be considered homogeneous. To this effect, (Jung *et al.*, 2012) proposed to use non-local information for the computation of the statistics of the region (pixels are not compared to their spatial neighbors, but to pixels that have similar neighbors, while still inside the same region). This idea is extended to SAR data by (Xia *et al.*, 2016) using the adapted similarity measures along with a multi-scale processing.

Other segmentation techniques applied to SAR are based on the minimum description length (MDL) introduced in the field of information theory by (Rissanen, 1978). The goal of this technique is to find a representation of the data which is well fitted while having a low complexity (i.e. few parameters). In term of image segmentation, this translates to finding a good segmentation and a good statistical description of the radiometry of each region. In (Galland *et al.*, 2003), this idea is applied to SAR data and it introduces the concept of the *active grid*, i.e. a set of nodes which can be joint by a segment to delimit the regions. It starts with a regular grid, and applies iteratively three operations: a merge (adjacent sets of nodes can be merged), a move (to move some nodes in the grid) and a remove (to remove unnecessary nodes for the representation).

Such methods could be useful for binary classification. However, we have seen that in the case of SWOT, the data have variations in the range resolution (making geometrical description of the result harder to use in practice) and both global (from the antenna pattern) and local (from the variations in the surface roughness) radiometry variations.

In the following, we present methods based on the MRF formulation to tackle these problems.

4.3 Markovian approaches

In section 3.2 we have seen that in order to define a random field, one has to define the clique potentials. Then the solution will be given by:

$$\begin{aligned}\widehat{\mathbf{u}} &= \arg \min_{\mathbf{u}} \mathcal{E}(\mathbf{u}) \\ &= -\log p(\mathbf{v}|\mathbf{u}) - \log p(\mathbf{u}),\end{aligned}\quad (4.13)$$

with the data term $-\log p(\mathbf{v}|\mathbf{u})$ depending on the acquisition of the data and the prior $-\log p(\mathbf{u})$ depending on the task.

A suitable and widely-used prior for binary classification is the Ising model that will penalize neighbor pixels belonging to different classes:

$$\psi(x, y) = \begin{cases} 0 & \text{if } x = y \text{ or,} \\ 1 & \text{if } x \neq y. \end{cases} \quad (4.14)$$

This prior will, independently of the data term, favor spatially homogeneous outputs. In the case of water bodies, this is particularly adapted to lakes. In the following, we will therefore use:

$$-\log p(\mathbf{u}) = \beta \sum_{i \sim j} \psi(u_i, u_j) + k, \quad (4.15)$$

where β is a positive weight to give more or less importance to the prior and k is a constant.

The data fidelity term ($-\log p(\mathbf{v}|\mathbf{u})$) describes how well the data \mathbf{v} is explained by the model \mathbf{u} . In the case of amplitude images, we have seen in chapter 2 that the data follows a Rayleigh-Nakagami distribution. Making the assumption that the distribution can be considered separable and that the parameter μ is spatially varying (denoted by $\mu_{u_i,i}$ for class u_i at pixel i), it gives:

$$-\log p(\mathbf{v}|\mathbf{u}) = -\sum_i \log \left(\frac{2}{\Gamma(L)} \frac{\sqrt{L}}{\mu_{u_i,i}} \left(\frac{\sqrt{L}v_i}{\mu_{u_i,i}} \right)^{2L-1} e^{-\left(\frac{\sqrt{L}v_i}{\mu_{u_i,i}} \right)^2} \right), \quad (4.16)$$

with L the number of looks. As we want to allow for non constant class parameters, the parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are images (which appear indexed in Equation 4.16). As for each pixel we will optimize on u_i , Equation 4.16 can be simplified by removing class independent terms (see section 2.3):

$$-\log p(\mathbf{v}|\mathbf{u}) = \sum_i \left(2L \log(\mu_{u_i,i}) + L \left(\frac{v_i}{\mu_{u_i,i}} \right)^2 \right) \quad (4.17)$$

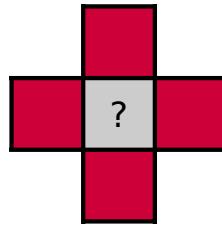


Figure 4.5 – We want to classify the pixel in the center to either class "red" of parameter μ_r or "white" of parameter μ_w . Given the configuration where the pixel to be classified is surrounded by "red" class, the regularization would favor labeling it as "red". We propose a method that answers the following question: for a given amplitude, what is the value of β from which the prior term will be higher than the data-term ?

While the number of looks L is usually a known image parameter, the distribution parameters μ_{u_i} (in our case u_i can either be 0 or 1, so we have two parameters) depend both on the acquisition parameters (mainly the wavelength and incidence angle) and the surface to be imaged (its roughness and topography). Their computation will be covered in [section 4.4](#).

The energy defined in [Equation 4.13](#) can then be minimized using the binary graphcut method of ([Greig and others, 1989](#)) presented in [subsection 3.2.2](#).

Setting of β In this framework, the β parameter balances the data term and the prior. It can not be set fully automatically as it depends on the application and desired result. While it can be set by an exhaustive search (trying many values for β , computing the results with respect to a ground truth and selecting the best value, provided that such a ground truth is available on a representative test area), it can also be set using the qualitative boxes method described by ([Azencott, 1992](#)); in the configuration shown in [Figure 4.5](#), we can compute the local energy for both classes (which differs by the class parameter used and the term 4β):

$$\begin{array}{ccc} \text{?} & \longrightarrow & \square \\ \text{?} & \longrightarrow & \square \end{array} \quad 2L \log(\mu_w) + L \left(\frac{v}{\mu_w} \right)^2 + 4\beta \quad (4.18)$$

$$\begin{array}{ccc} \text{?} & \longrightarrow & \blacksquare \\ \text{?} & \longrightarrow & \blacksquare \end{array} \quad 2L \log(\mu_r) + L \left(\frac{v}{\mu_r} \right)^2 . \quad (4.19)$$

Setting [Equation 4.18](#) equal to [Equation 4.19](#) gives:

$$\beta = \frac{L}{2} \log \left(\frac{\mu_w}{\mu_r} \right) + \frac{Lv^2}{4} \left(\frac{1}{\mu_w^2} - \frac{1}{\mu_r^2} \right) . \quad (4.20)$$

Once μ_r and μ_w have been computed, we can obtain a curve such as the one shown in [Figure 4.6](#), allowing for a simpler way to set β . In order to use this curve to set β , it should be read as follow: on the considered configuration of [Figure 4.5](#) (where the most likely class should be red), what is the minimum value of the amplitude of the pixel that should be classified as white? Setting v to this value in [Equation 4.20](#) gives the β

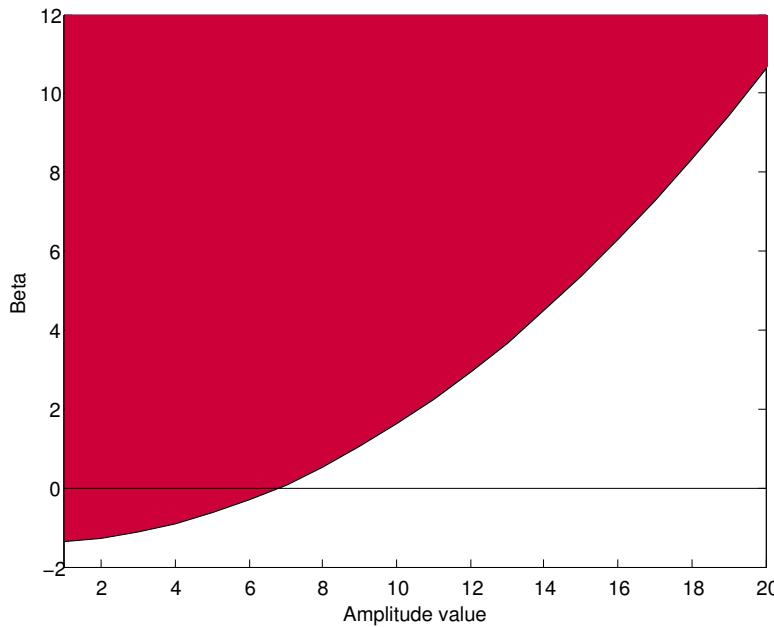


Figure 4.6 – Tipping point using $\mu_r = 5$ and $\mu_w = 10$, $L = 4$ and the configuration presented in Figure 4.5. For a given amplitude value, the assigned class will be "red" if β is higher than the curve, or "white" otherwise.

that should be used.

Summary: Markov Random Fields for water/land classification

While this is not a new method, MRF is a simple model that enforces spatial regularity in the classification. Using an Ising prior allows for a spatial regularization to limit the effect of the speckle.

A regularization parameter is needed to balance the data term and the prior. We presented a method to easily set this parameter.

The pseudo-code of this method is presented in [algorithm 3](#).

Algorithm 3: MRF for water/land classification

Data: Observed image v

Result: Model u

```

Set class parameters  $\mu_0, \mu_1$  /* Either known a priori or computed from a
first unsupervised classification (e.g. K-Means) */
 $\beta = \text{choose\_from\_qualitative\_boxes}(\mu_0, \mu_1, v_{ref});$ 
/* Optimization, see subsection 3.2.2 */
graph = construct( $v, \mu_0, \mu_1, \beta$ ) /* Using method from (Greig and others,
1989)
gc = graphcut(graph);
 $\forall i, u_i = \begin{cases} 0 & \text{if } gc(i) \in \text{source} \\ 1 & \text{otherwise.} \end{cases}$ 
```

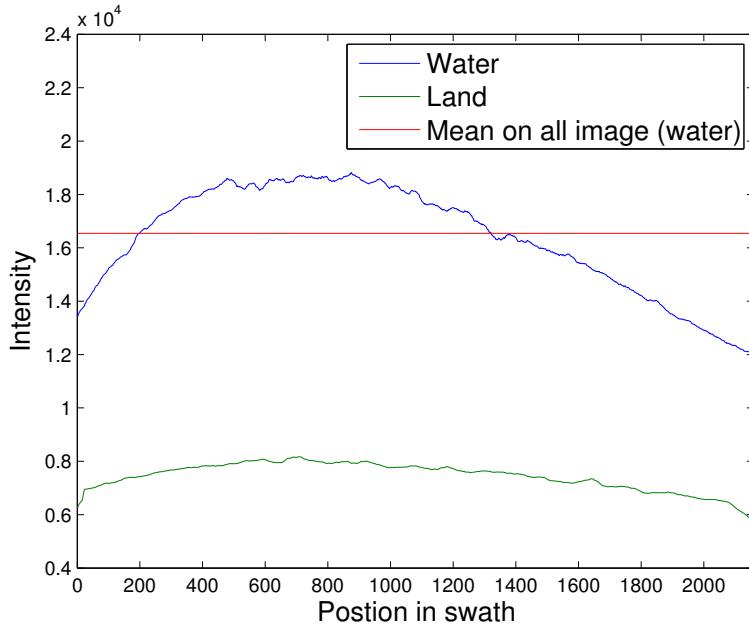


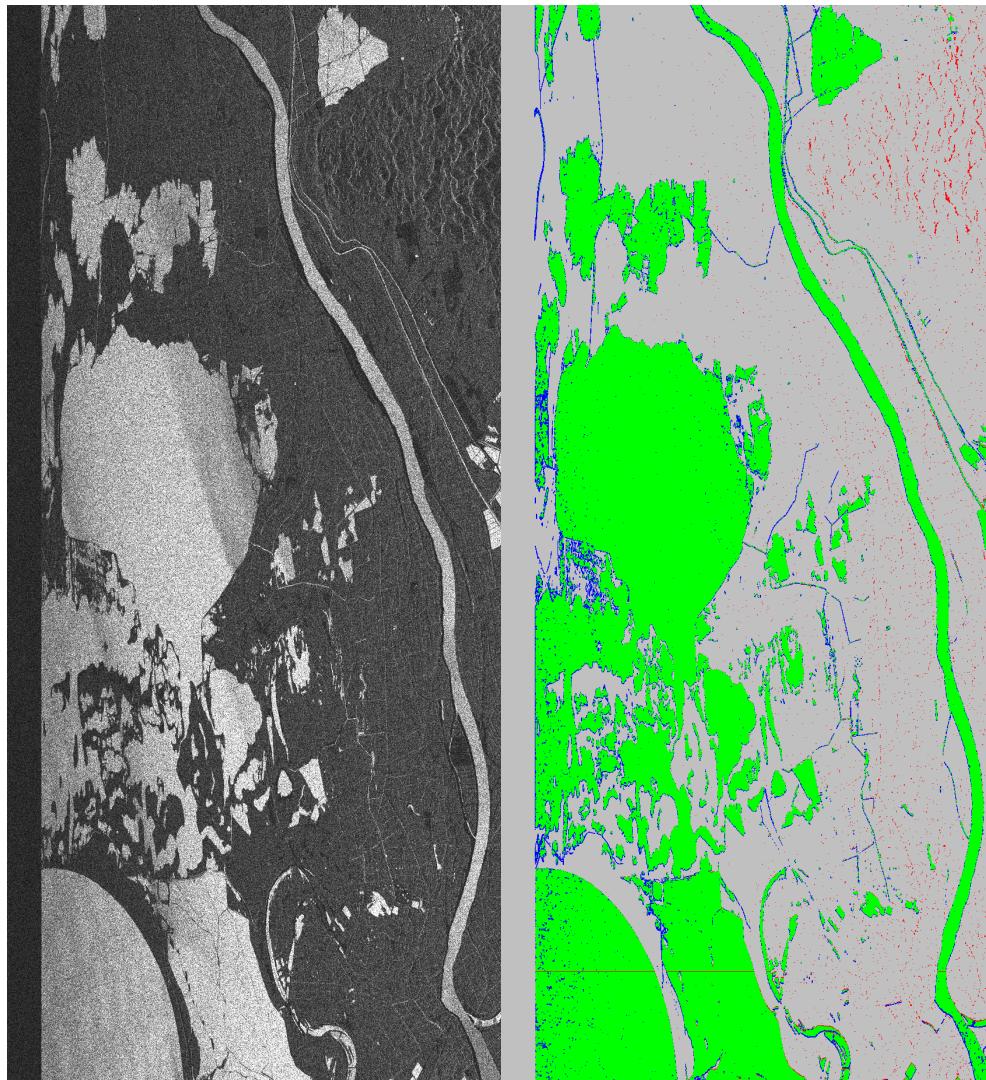
Figure 4.7 – Average parameter evolution in the swath. The green curve shows the parameters for land class, blue for water and red show the MLE when assuming constant water parameter. The water curve reflects the antenna pattern. The land profile is flatter, as land is dominated by thermal noise.

4.4 Estimation and spatial regularization of class parameters

In this section, we propose to adapt the algorithm of section 4.2 to the particular properties of SWOT images. We have seen in section 2.4 that the antenna pattern of SWOT amplitude images cannot be properly corrected. Also, because of the wind-related water roughness variations and land topological ones, there may be strong variations in the backscattered signal for both classes. These differences make popular classification algorithms presented in section 4.1 not suited to SWOT amplitude images. To demonstrate this, we show in Figure 4.8 a result obtained using classical MRF with the exact optimization method presented in subsection 3.2.2 proposed by (Greig and others, 1989). This result can be easily explained, even when only considering the variations introduced by the antenna pattern. We show in Figure 4.7 the parameters for both land and water classes when computed at each column of the image (range direction), and the MLE for the water parameter if we consider a single parameter. It can be easily seen that a scalar parameter (red curve) does not well represent the class at the beginning and end or at the middle of the swath.

In this section, we will use the framework provided by MRF as defined in section 3.2 and adapt this classical algorithms to the case of non-constant classes parameters.

Our goal is to estimate parameters images μ_0 and μ_1 so we can use them in a MRF framework as seen in section 4.3. Parameter estimation can be done either once, prior to the classification, or come in an iterative process, such as proposed in (Fjørtoft *et*



(a) Input amplitude simulated image of the
Camargue area (b) Result using a scalar parameter for land
and water

Legend: True positive True negative False positive False negative

Figure 4.8 – Extract of a result using **MRF** with constant parameters on a simulated **SWOT** image. Note that we have more false negative at the beginning of the swath than on the middle

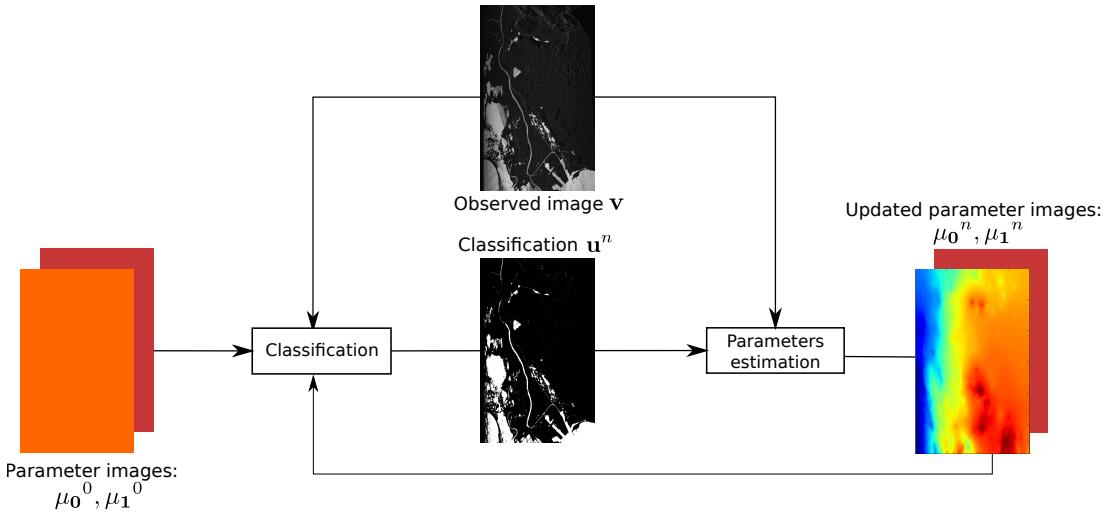


Figure 4.9 – Iterative parameters estimation: from a first estimation of the parameters, a classification is computed and used to reestimate a new set of parameters. This process is then iterated.

al., 2003). In this case, a first estimation $\{\mu_0^0, \mu_1^0\}$ is made. This estimation is then used by the MRF to obtain a classification u^0 , which is then used to estimate a new set of parameters $\{\mu_0^1, \mu_1^1\}$. This process is repeated, with a refined set of parameters and classification at each iteration. The general process is sketched in Figure 4.9. In the following, we will present 4 methods for the estimation of the parameters, ordered by complexity:

1. Constant parameters, with no re-estimation;
2. Non-constant parameters with no re-estimation;
3. Region-based parameters estimation;
4. Dense estimation.

While the first and second ones only use information known *a priori*, the two other methods use the iterative process pictured in Figure 4.9.

4.4.1 Constant parameters

A simple method is to consider that both classes have the same physical properties everywhere in the image and to estimate the two parameters using the MLE (which corresponds to an averaging in intensity) based on a first classification or a prior mask u^0 :

$$\forall j, \mu_{0,j} = \sqrt{\frac{\sum_i (1 - u_i^0) v_i^2}{\sum_i (1 - u_i^0)}} \quad \forall j, \mu_{1,j} = \sqrt{\frac{\sum_i u_i^0 v_i^2}{\sum_i u_i^0}}. \quad (4.21)$$

Another method is to use a simple model for the expected physical properties of the water and land areas, and use them to obtain the values of μ_0 and μ_1 .

We will see in the results presented in [section 4.5](#) that in the case of SWOT, using this simple method for the parameters definition gives poor results. Therefore, estimating the parameters tailored to the observed data is the main problem we seek to solve in this chapter.

4.4.2 Non-constant parameters with no re-estimation

In [section 2.4](#), we saw that:

- Water and land's σ_0 do not behave the same way with respect to the incidence angle;
- We have a low [SNR](#) (especially for the land class). This means that land signal is dominated by thermal noise which is not affected by the antenna pattern.

Considering these two reasons, the antenna pattern of the system cannot be corrected as it is usually done in SAR systems. This is one of the two sources of variations in the backscattering properties across the image that we presented in the introduction of [section 4.4](#). While local variations depends on the imaged scene, variations due to the antenna pattern ("global variations") are always present in the system.

Obtaining the antenna pattern is a widely studied problem in SAR as it is often necessary to correct the effect of the antenna before exploiting the images. For instance, in ([Bachmann and others, 2010](#)), a model for the antenna of TerraSAR-X is proposed. An other method, used in the case of ENVISAT, is to use observations of homogeneously distributed targets to obtain a precise model of the antenna pattern after the launch of the satellite. In the case of [SWOT](#), the antenna pattern will be very precisely known.

In addition to the effects of the antenna pattern, the backscattered signal for land areas highly depends on the topography (see [section 2.4](#)). For instance, it can result in layover with a strong signal. This can be predicted from the [DEM](#) when it is available. In this section we present two different simple methods for the definition of the class parameters.

Estimation from the Xfactor As seen in [section 2.4](#), the Xfactor is a by-product of the acquisitions that links the received signal to the σ_0 of both classes. We recall [Equation 2.14](#) (in dB):

$$P_{\text{received}} = P_{\text{signal}} + P_{\text{noise}} \quad (4.22)$$

$$= \text{Xfactor} \times \sigma_0 + P_{\text{noise}}. \quad (4.23)$$

Considering that we also know (from the calibration phase) the mean noise level and the typical σ_0 for both classes, we can obtain the class parameters. While the Xfactor would normally provide good classes parameters, it can be inaccurate in many ways:

- We make hypotheses about the expected σ_0 of water and land. However, natural variations (e.g. wind speed and topography) can result in variations of the σ_0 .

- The Xfactor is computed based on a reference DEM which is not perfect.

For these reasons, it is sometimes better to discard the Xfactor for the parameter estimation, and use the following method.

Estimation based on a first classification In this paragraph, we propose to estimate the parameter images based on the observed images and a first classification. This first classification can either be obtained using the constant MRF as defined in section 4.3 (using the constant parameter estimation presented in subsection 4.4.1; or the Xfactor presented in the previous paragraph) or be a prior mask (either from a priori data, or from a previous pass over the same site).

A straightforward way to obtain it is to use an averaging window; we use a window with the same height as the image, and a width of few columns (depending on the distribution of both classes). For the image of Figure 4.10, we have chosen a width of 30 pixels. Sliding this window in the range direction gives a curve for each class, such as the ones presented in Figure 4.7. This method assumes that the variations in the class parameters only come from the antenna pattern, which is not true (as explained in the introduction of section 4.4). Even in the case of perfectly constant class parameters beside the effect of the antenna pattern, we can expect to have variations when only a limited number of samples of a class are present in a window.

To limit the effect of varying parameters and estimation errors, we fit a second order polynomial using the least-square method to the parameter curve. The resulting second order polynomials are noted p_0 and p_1 in the following. Note that when a more accurate antenna pattern model is available, it can also be expressed as a 1D signal, and used in the same way in the following. The process is illustrated in Figure 4.10.

Using the 1D signal representing the antenna pattern (estimated or known a priori), obtaining the parameters images is straightforward; the 1D signal is repeated along the azimuth direction in order to obtain a 2D signal:

$$\forall (i, j) \in [1, N_{\text{Azimuth}}] \times [1, N_{\text{Range}}] \quad \mu_{x,(i,j)}^0 = p_{x,j}, \quad (4.24)$$

where N_{Azimuth} is the number of pixels in the azimuth direction and N_{Range} is the number of pixels in the range direction and x can either be 0 (land class) or 1 (water).

Comparison of the two methods These two methods only take into account prior information and can be easily compared. Such a comparison is shown in Figure 4.11 using the ground truth mask for the window averaging method, and the theoretical values for the σ_0 of land and water. Based on the image size ($\approx 3500 \times 3500$ data pixels), the estimation is made on a window of size 21×3500 , making the MLE accurate. We can see that we have a small bias (that comes from inaccurate values for σ_0). However, we can see that the shapes of the curves are mostly the same.

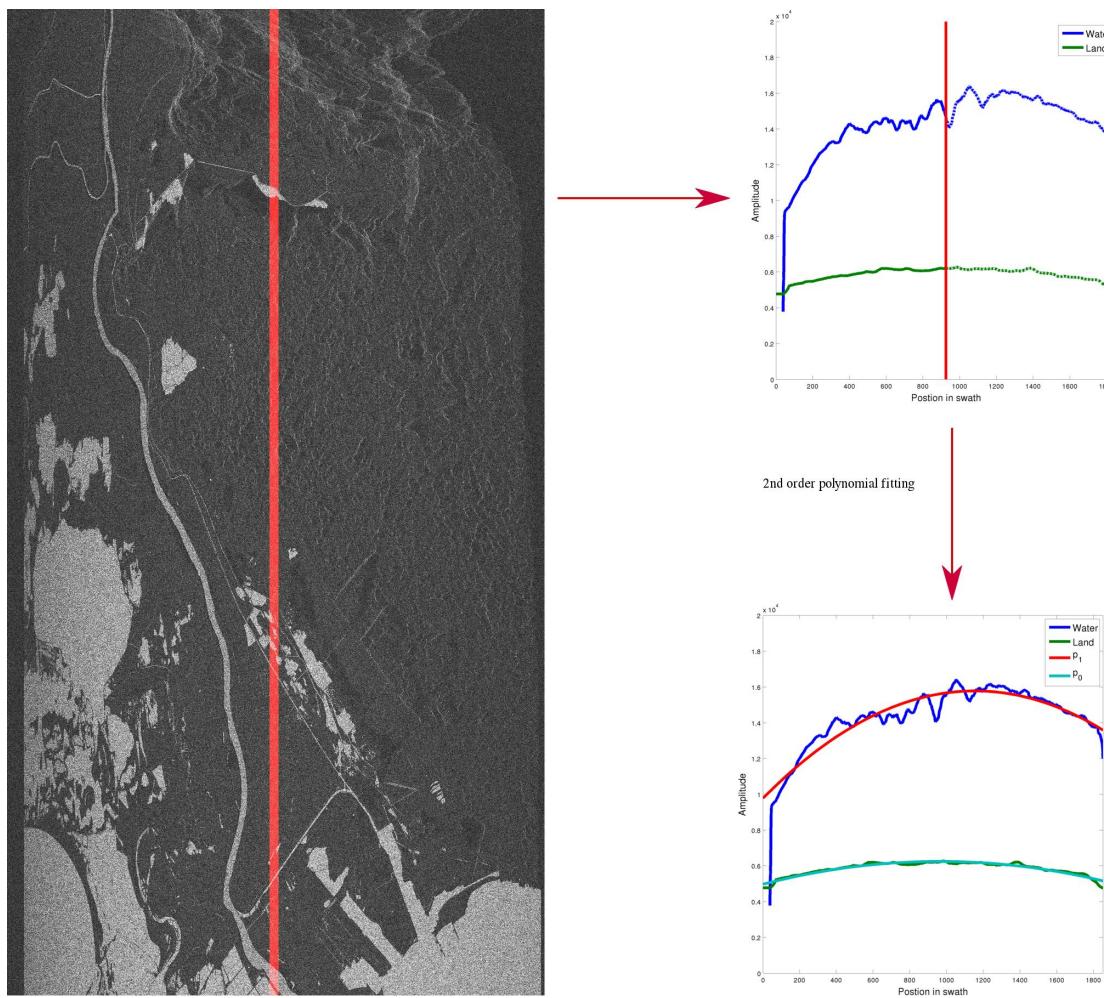


Figure 4.10 – To compute the antenna pattern, a sliding averaging window (in red on the amplitude image) is applied. The resulting curves are then fit to a second order polynomial. See www.sylvainlobry.com/phd for the animated image.

Summary: Non-constant parameters estimation from antenna pattern

We present two simple methods to estimate spatially variable class parameters. In the first method, we use the theoretical antenna pattern and a DEM. The second method only uses a reference mask to compute variable parameters in the range direction. The pseudo-code for this algorithm is given in [algorithm 4](#).

We will see in [section 4.5](#) that while these methods do not give very good results, they provide a good initialization for the after-mentioned methods.

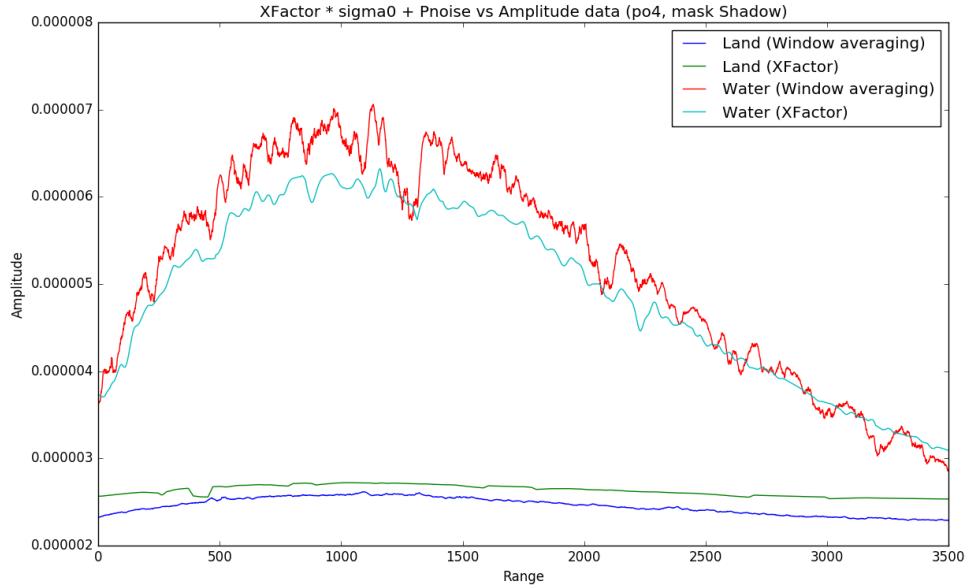


Figure 4.11 – Comparison of the two methods for the estimation of the parameters based on prior information averaged (window width = 21) on the range direction.

Algorithm 4: MRF for water/land classification with antenna pattern

Data: Observed image \mathbf{v}

Result: Model \mathbf{u} , antenna pattern A

```

 $\mu_0, \mu_1$  = compute_parameters_from_antenna_pattern( $A$ );
 $\beta$  = choose_from_qualitative_boxes( $\mu_0, \mu_1, v_{ref}$ );
/* Optimization, see subsection 3.2.2 */ 
graph = construct( $\mathbf{v}, \mu_0, \mu_1, \beta$ ) /* Using method from (Greig and others,
1989) */
gc = graphcut(graph);
 $\forall i, u_i = \begin{cases} 0 & \text{if } gc(i) \in \text{source} \\ 1 & \text{otherwise.} \end{cases}$ 

```

4.4.3 Region-based parameters estimation

While the method presented in subsection 4.4.2 can use non-uniform parameters, it is only able to take into account systematic variations coming from the antenna pattern. In this section, we describe an iterative method based on a quadtree partitioning of the image to estimate parameters that can capture local variations. Our goal is to obtain a partition $P = \{r_1, \dots, r_n\}$ where the parameters can be considered constant within each region r_i . Hence, the class parameters can be estimated in each region independently. To obtain a satisfying non-uniform estimation of the parameters, the partition should fulfill two requirements:

- \mathcal{R}_1 : each region of the partition should be large enough and contain enough samples of each class so the MLE can be applied reliably.
- \mathcal{R}_2 : each region of the partition should be small enough so we can capture local variations.

Partitioning process The partitioning process is made using a quadtree approach (Samet, 1984). Quadtrees have been extensively used in image processing; application includes image coding (e.g. (Sullivan and Baker, 1994)) and image segmentation (e.g. (Bouman *et al.*, 1994)). In our approach, the partition is updated at the "parameter estimation box" in the general process presented in Figure 4.9. The partitioning is made as follow:

1. We start with a partition in one region ($P = \{r_0\}$) and a first classification \mathbf{u}^0 . Note that in this case, \mathcal{R}_2 is not fulfilled as r_0 can not be considered small. Parameter images $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are either computed by the MLE on the entire image (using \mathbf{u}^0) and are then constant, or computed from the antenna pattern (see subsection 4.4.2).
2. For each region of the partition, we seek a new partition:
 - (a) First, the region is divided into 4 regions of equal size.
 - (b) If one of the newly-created regions break \mathcal{R}_1 (verified with the current classification), the proposed partitioning is canceled, and a new partition in 2 regions (with a vertical cut) is tried.
 - (c) If this partition breaks \mathcal{R}_1 , we cancel the proposed partitioning and we try a partition with only an horizontal cut.
 - (d) If this also breaks \mathcal{R}_1 , the proposed partitioning is canceled and the region is left unchanged.
3. Parameters are computed using the MLE region-wise to obtain the new parameter images $\boldsymbol{\mu}_0^n$ and $\boldsymbol{\mu}_1^n$.
4. We perform the regularization described in the next paragraph on $\boldsymbol{\mu}_0^n$ and $\boldsymbol{\mu}_1^n$.

5. A new classification \mathbf{u}^n is obtained using μ_0^n and μ_1^n with the method described in [section 4.3](#).
6. We repeat the process from step 2 until no region can be partitioned.

The partitioning process for each region (item 2.) is illustrated in a simple case in [Figure 4.12](#)

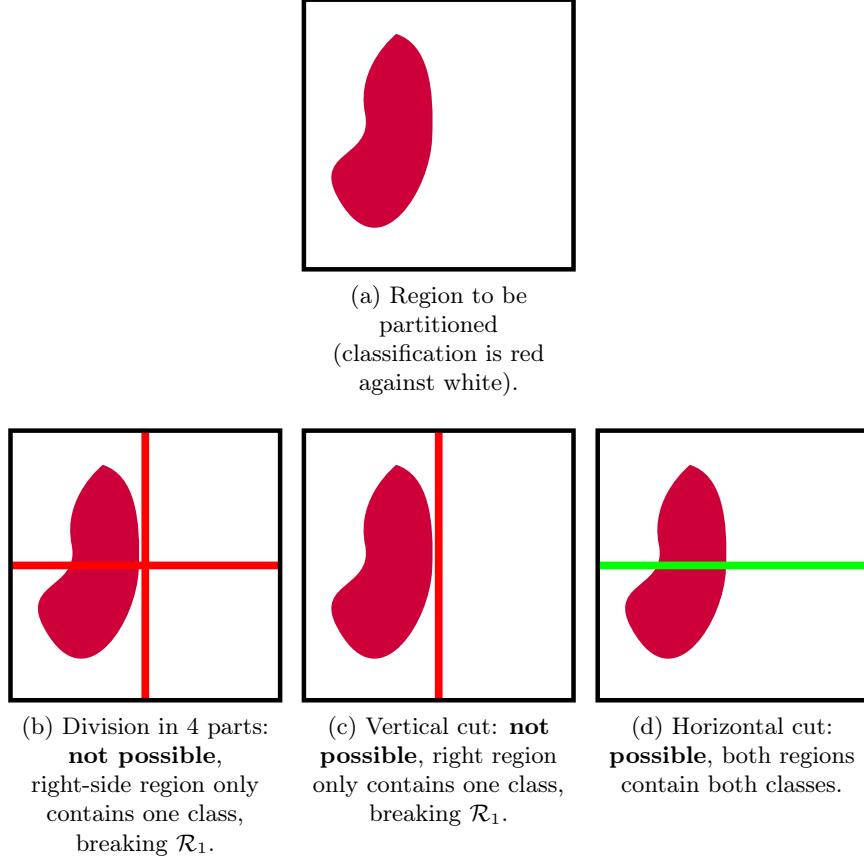


Figure 4.12 – Partitionning process for a given region. At first a division in 4 regions is attempted. If it is not possible, a division in 2 regions is attempted.

In practice, the partitioning process can only break \mathcal{R}_1 (as it is trying to create smaller region). Therefore, we only need to check for \mathcal{R}_1 during the process. \mathcal{R}_2 is then fulfilled as well as possible when we break \mathcal{R}_1 .

Post-estimation regularization The objective sought in \mathcal{R}_1 is that we have enough points to reliably use the [MLE](#) for both classes. Using the current classification \mathbf{u}^i we check at each step i that we have enough representatives of each class in each region to be created. The problem with this method is that we do not know the exact classification, and \mathbf{u}^i may contain wrong information (this is especially true during the first steps). In this case, we may create regions where only one class is actually present. When we try to do binary classification on images containing only one class, we have a degenerate case which can lead to a wrong estimation of the parameters. An illustration of this phenomenon on a simple example is presented in [Figure 4.13](#).

12	5	4
4	10	16
13	11	10

(a) Input image to be classified, white parameter is 10, red is 20.

12	5	4
4	10	16
13	11	10

(b) First classification.

12	5	4
4	10	16
13	11	10

(c) Second classification.

12	5	4
4	10	16
13	11	10

(d) Third classification

12	5	4
4	10	16
13	11	10

(e) Fourth classification.

12	5	4
4	10	16
13	11	10

(f) Fifth classification.

Figure 4.13 – Simple example; we want to classify the input image (let's assume Gaussian noise, for this illustration only). This image should be clearly classified as white. However, we can note that due to the iterative estimation of parameters, it ends up classified as mostly red, with a red parameter close to the initial white parameter.

To cope with this behavior, we propose to add a global regularization step. While water reflectivity can have local variations that are not known *a priori*, global variations due to the antenna pattern are generally known. For the land class, we can also expect variations from the land topography, that can be computed from a DEM when it is available. After each iteration of the parameters estimation, we check that the estimated parameters are not too far from their theoretical value according to the Xfactor ("too far" being controlled by an additional parameter β_{reg}). Recall that the theoretical values come from the prior information contained in the first parameter images μ_0 and μ_1 . When an estimated parameter is too far from its theoretical value, it is assumed to be diverging. In this case, it is set to the theoretical value prior to the re-estimation of the parameter:

$$\forall i \quad \mu_{x,i}^n = \begin{cases} \mu_{x,i}^n & \text{if } |\mu_{x,i}^n - \mu_{x,i}^0| < \beta_{\text{reg}}, \\ \mu_{x,i}^0 & \text{otherwise.} \end{cases} \quad (4.25)$$

In practice, this prevents from diverging points, as it is shown in [section 4.5](#).

Summary: Region-based estimation of the class parameters

In the proposed method, classification and partitioning of the image are alternatively performed. Classes parameters are estimated region-wise, allowing for variations. An additional step of regularization is performed after each estimation to avoid degenerate cases. In this method, we have three parameters:

- The minimum size of a region.
- The minimum proportion of each class in a region.
- The maximal difference between the theoretical value (given by the antenna pattern) and the estimated value.

The pseudo-code for this algorithm is given in [algorithm 5](#).

Algorithm 5: Water/land classification and quadtree parameter estimation

Data: Observed image v , min_size_region, min_portion_of_each_class, β_{reg}

Result: Model u , antenna pattern A

```

 $\mu_0^0, \mu_1^0 = \text{compute\_parameters\_from\_antenna\_pattern}(A);$ 
 $\beta = \text{choose\_from\_qualitative\_boxes}(\mu_0, \mu_1, v_{\text{ref}});$ 
graph = construct( $v, \mu_0, \mu_1, \beta$ ) /* Using method from (Greig and others, 1989) */
gc = graphcut(graph);
 $\forall i, u_i^0 = \begin{cases} 0 & \text{if } gc(i) \in \text{source} \\ 1 & \text{otherwise.} \end{cases};$ 
 $P_0 = \text{init\_partition}();$ 
while  $\mathcal{R}_1$  do
    /* Parameter estimation */  

     $P^n = \text{partition}(P^{n-1});$ 
     $\mu_0^n, \mu_1^n = \text{compute\_parameters\_from\_}(P^n, u_{n-1}, v);$ 
     $\mu_0^n, \mu_1^n = \text{post\_estimate\_regul}(\mu_0^n, \mu_1^n, \beta_{\text{reg}})/* \text{ See Equation 4.25 */}$ 
    /* Optimization, see subsection 3.2.2 */  

    graph = construct( $v, \mu_0^n, \mu_1^n, \beta$ ) /* Using method from (Greig and others, 1989) */
    gc = graphcut(graph);
     $\forall i, u_i^n = \begin{cases} 0 & \text{if } gc(i) \in \text{source} \\ 1 & \text{otherwise.} \end{cases};$ 

```

4.4.4 MRF estimation of the parameter images

While the region-based parameter estimation allows for some variations in the parameters, we have seen that the two requirements on the partition are contradictory; the first one requires large regions so the **MLE** can be used efficiently, the second one requires small regions so that we can have local variations. In practice, this prevents the model from capturing variations at small scales. It also tends to create sharp transitions rather than gradual evolution in the parameters

To cope with this problem, we propose a model that is still based on the general process presented in [Figure 4.9](#) with the difference that the parameters images (μ_0, μ_1) are now allowed to take different values at each pixel (i.e. no curve-based or region-based constraint). To obtain such parameters images, we are going to stay in the **MRF** framework; in addition to the classification step of [Figure 4.9](#) being a **MRF**, we will also use such a model for the estimation part defining one **MRF** for each parameter map μ_0 and μ_1 .

We have seen in [chapter 2](#) that speckle in amplitude SAR images can be described by a multiplicative model on the square root of the reflectivity μ :

$$v_i = \mu_i \times \xi_i . \quad (4.26)$$

To obtain an energy function easier to optimize, we consider the logarithm of the images. When considering the logarithm of the image, the image is corrupted by an additive term $\tilde{\xi}_i$. This term can be considered in a first approximation to be Gaussian distributed (see [chapter 2](#)):

$$\tilde{v}_i = \tilde{\mu}_i + \tilde{\xi}_i \approx \tilde{\mu}_i + \eta_i , \quad (4.27)$$

where $\tilde{x} = \log x$ and η_i is Gaussian distributed.

In the following, we define two **MRF**; one for each $\tilde{\mu}_x, x \in \{0, 1\}$. We can estimate $\tilde{\mu}_x$ in a Bayesian framework, based on the knowledge of \tilde{v} and u (dropping the constant values and assuming that $p(\tilde{\mu}_x|u) = p(\tilde{\mu}_x)$):

$$-\log(p(\tilde{\mu}_x|\tilde{v}, u)) = -\log(p(\tilde{v}|\tilde{\mu}_x, u)) - \log(p(\tilde{\mu}_x)) \quad (4.28)$$

In the **MRF** of $\tilde{\mu}_x$, we want to define the data-term such that it takes into account the observation at pixels i where the current classification is $u_i^n = x$. As we assumed the noise is Gaussian distributed, the data term is simply an L2 distance between the model and the observation. In the cases where $u_i^n \neq x$, the data-term is set to a constant value (0 in our case):

$$-\log(p(\tilde{v}|\tilde{\mu}_x, u)) = \sum_i \begin{cases} (\tilde{v}_i - \tilde{\mu}_{x,i})^2 & \text{if } u_i^n = x , \\ 0 & \text{otherwise .} \end{cases} \quad (4.29)$$

$$= \sum_i \bar{\delta}(u_i^n = x)(\tilde{v}_i - \tilde{\mu}_{x,i})^2 , \quad (4.30)$$

with:

$$\bar{\delta}(x) = \begin{cases} 1 & \text{if } x \text{ is true,} \\ 0 & \text{otherwise.} \end{cases} \quad (4.31)$$

The regularization term enforces spatial regularization. As more systematic variations in the range direction are expected (because of the antenna pattern), the spatial regularization is expressed as the sum of two terms: one for the range direction, the other for the azimuth. For the same reasons as in [subsection 4.4.3](#), we also add a regularization term taking into account the theoretical value from the Xfactor and the σ_0 hypotheses. It yields the following formula for $x \in \{0, 1\}$:

$$-\log(p(\tilde{\boldsymbol{\mu}}_{\mathbf{x}})) = \beta_{\text{az}} \sum_{\substack{i \sim j \\ \text{az}}} (\tilde{\mu}_{x,i} - \tilde{\mu}_{x,j})^2 + \beta_{\text{rg}} \sum_{\substack{i \sim j \\ \text{rg}}} (\tilde{\mu}_{x,i} - \tilde{\mu}_{x,j})^2 + \beta_{\text{th}} \sum_i (\tilde{\mu}_{x,i} - \tilde{\mu}_{x,i}^0)^2. \quad (4.32)$$

Another option is to take into account the expected variations coming from the Xfactor directly in the first two terms:

$$-\log(p(\tilde{\boldsymbol{\mu}}_{\mathbf{x}})) = \beta_{\text{az}} \sum_{\substack{i \sim j \\ \text{az}}} (\tilde{\mu}_{x,i} - \tilde{\mu}_{x,j} - \Delta(i,j))^2 + \beta_{\text{rg}} \sum_{\substack{i \sim j \\ \text{rg}}} (\tilde{\mu}_{x,i} - \tilde{\mu}_{x,j} - \Delta(i,j))^2, \quad (4.33)$$

where $\Delta(i,j) = \tilde{\mu}_{x,i}^0 - \tilde{\mu}_{x,j}^0$. Note that in the case of $-\log(p(\tilde{\boldsymbol{\mu}}_1))$, we do not have variations in the azimuth direction (i.e. $\forall i \sim j, \Delta(i,j) = 0$). While [Equation 4.33](#) remove β_{th} making it easier to set, [Equation 4.32](#) allows for a finer tuning, expressing how confident one is in the prior.

In the following, we will use [Equation 4.33](#), giving the following two MRF:

$$\begin{aligned} \boldsymbol{\mu}_0^{n+1} &= \arg \min_{\boldsymbol{\mu}} \mathcal{E}_0(\boldsymbol{\mu}) \\ &= \arg \min_{\boldsymbol{\mu}} \sum_i (1 - u_i^n)(\tilde{v}_i - \tilde{\mu}_i)^2 + \beta_{\text{az}} \sum_{\substack{i \sim j \\ \text{az}}} (\tilde{\mu}_i - \tilde{\mu}_j)^2 \\ &\quad + \beta_{\text{rg}} \sum_{\substack{i \sim j \\ \text{rg}}} (\tilde{\mu}_i - \tilde{\mu}_j)^2 + \beta_{\text{th}} \sum_i (\tilde{\mu}_i - \tilde{\mu}_{0,i}^0)^2 \end{aligned} \quad (4.34)$$

$$\begin{aligned} \boldsymbol{\mu}_1^{n+1} &= \arg \min_{\boldsymbol{\mu}} \mathcal{E}_1(\boldsymbol{\mu}) \\ &= \arg \min_{\boldsymbol{\mu}} \sum_i u_i^n (\tilde{v}_i - \tilde{\mu}_i)^2 + \beta_{\text{az}} \sum_{\substack{i \sim j \\ \text{az}}} (\tilde{\mu}_i - \tilde{\mu}_j)^2 \\ &\quad + \beta_{\text{rg}} \sum_{\substack{i \sim j \\ \text{rg}}} (\tilde{\mu}_i - \tilde{\mu}_j)^2 + \beta_{\text{th}} \sum_i (\tilde{\mu}_i - \tilde{\mu}_{1,i}^0)^2 \end{aligned} \quad (4.35)$$

As \mathcal{E}_0 and \mathcal{E}_1 are continuous and quadratic problems, they can be optimized using the conjugate gradient method.

Summary: Markovian estimation of the parameter images

In this method, parameter images are estimated using a markovian framework to regularize the parameters. The regularization is both spatial and based on the initial parameters to prevent over-fitting. This method uses three parameters:

- two weights for the spatial regularization in both directions;
- a weight for the regularization considering prior information such as the Xfactor.

The pseudo-code of this algorithm is given in [algorithm 6](#).

Algorithm 6: Water/land classification and MRF parameters estimation

Data: Observed image $v, \beta_{az}, \beta_{rg}, \beta_{th}$

Result: Model u , antenna pattern A

$\mu_0^0, \mu_1^0 = \text{compute_parameters_from_Xfactor}(A);$

$\beta = \text{choose_from_qualitative_boxes}(\mu_0, \mu_1, v_{ref});$

$\text{graph} = \text{construct}(v, \mu_0, \mu_1, \beta) /* \text{ Using method from (Greig and others, 1989)}$

*/

$\text{gc} = \text{graphcut}(\text{graph});$

$$\forall i, u_i^0 = \begin{cases} 0 & \text{if } \text{gc}(i) \in \text{source} \\ 1 & \text{otherwise.} \end{cases};$$

while $\mu^n \neq \mu^{n-1}$ **do**

/* Parameter estimation

*/

$\mu_0^n = \text{compute_parameters}(v, u^{n-1}, \beta_{az}, \beta_{rg}, \beta_{th});$

$\mu_1^n = \text{compute_parameters}(v, u^{n-1}, \beta_{az}, \beta_{rg}, \beta_{th});$

/* Optimization, see subsection 3.2.2

*/

$\text{graph} = \text{construct}(v, \mu_0^n, \mu_1^n, \beta) /* \text{ Using method from (Greig and others, 1989)}$

*/

$\text{gc} = \text{graphcut}(\text{graph});$

$$\forall i, u_i^n = \begin{cases} 0 & \text{if } \text{gc}(i) \in \text{source} \\ 1 & \text{otherwise.} \end{cases};$$

4.5 Results

In this section we present the results obtained using the following methods (presented in this chapter):

- the pixel-based **MLE** and **MAP** classification algorithms as well as the **MRF** classifier with:
 - constant parameters;
 - parameters accounting for the antenna pattern;
 - parameters based on the Xfactor;
- the region-based **MRF**;
- and the **MRF** classification with the Markovian parameters estimation.

The results are presented on four images:

- an extract of a simulated **SWOT** image (using **JPL**'s HR science simulator of the Garonne river (in France);
- an extract of a simulated **SWOT** image (using **JPL**'s HR science simulator of the Po river (in Italy);
- a simulated **SWOT** image (using **CNES** HR simulator) of the Camargue area (in France);
- and a real image of the Kaw area (in France) acquired by **Office National d'Études et de Recherches Aérospatiales (ONERA)** SETHI, a P-band airborne sensor, during the TropiSAR campaign .

Data: The main characteristics of the data are presented in [Table 4.1](#) and the images are shown in [Figure 4.14](#). We only use extracts of the Garonne and Po images as the **JPL** simulator outputs "no data" pixels (generally when the input **DEM** tile is smaller than the swath). These pixels can be difficult to handle in the graphcut optimization of the **MRF** and in the markovian estimation of the parameters. While we have developed (and implemented in the **CNES** processing chain) methods to take them into account, their presence can slightly decrease the performances of the markovian estimation of the parameters. Furthermore, "no data" pixels will be very limited in real **SWOT** images. Therefore, we extracted portions of the images without such pixels.

The data acquired by SETHI has two advantages for the evaluation of the methods:

1. the antenna pattern is not corrected. Therefore, it shows similar behavior to what is expected in **SWOT** data;
2. it allows us to evaluate our models on real data, in addition to simulations.

	Garonne	Po
Origin	JPL SWOT simulator	
Band	Ka (35.75GHz)	
Incidence angle	0.6 - 3.9°	
Resolution (azimuth × range)	70m-10m × 5m	
Size (azimuth × range)	2915 × 1530	1517 × 3108
Ground truth	from simulator	from simulator (with manual corrections)
Mixed pixels	Land	Land

	Camargue	Kaw
Origin	CNES SWOT simulator	ONERA SETHI
Band	Ka (35.75GHz)	P (0.44GHz)
Incidence angle	0.6 - 3.9°	24 - 62°
Resolution (azimuth × range)	70m-10m × 5m	1.5m × 1.2m
Size (azimuth × range)	2979 × 1839	5788 × 4000
Ground truth	from simulator	manual
Mixed pixels	Water	N/A

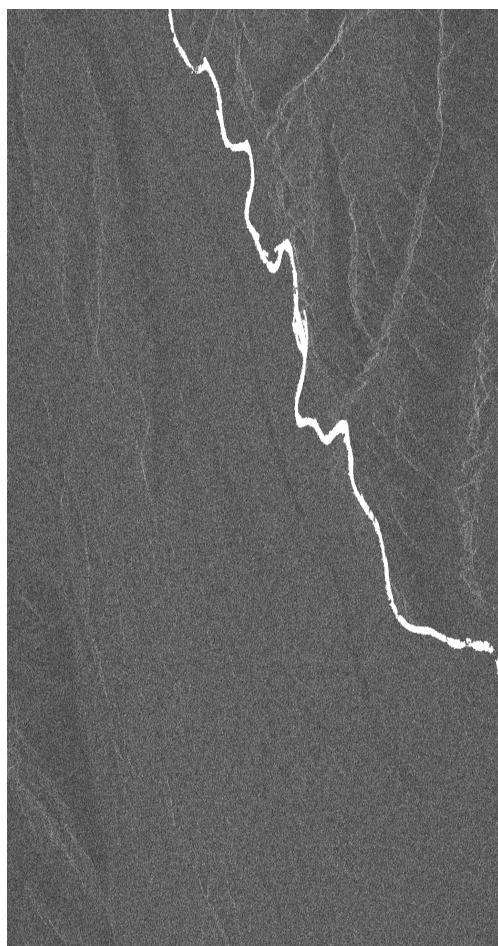
Table 4.1 – Characteristics of the data used for the evaluation of the methods

Note that in the case of SETHI, water is dark and land is brighter (which is the opposite of SWOT). This is not a problem for the proposed methods.

In the SWOT simulators (both for JPL and CNES), there is the option to add local variations in σ_0 (see Equation 2.14). Over land, there will also be variations due to the topography. As we can see in Figure 4.14, these local variations are not present in a same way in all images. In the Garonne image, there are almost no variations except in the right part of the image. On the opposite, in the Po image, there are cases of "dark water": when there is little or no wind on the surface (i.e. it is smooth), it acts as a mirror and no signal is backscattered towards the radar. In the Camargue image, some areas are a bit darker but there is no completely dark water. Also note that the water/land contrast is higher in the Garonne and Po images than in the Camargue image.

Methods: For the methods not iteratively estimating parameters (i.e. MLE, MAP with statistical distribution of the data and MRF), we use three set of parameters:

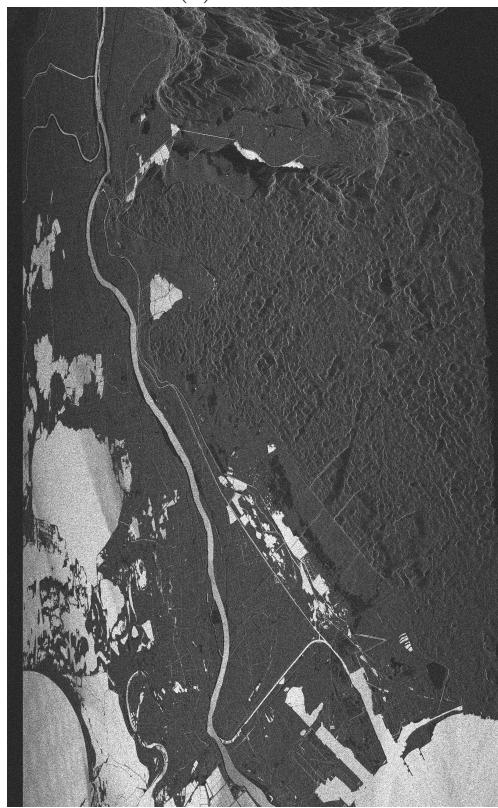
- constant parameters: to compare with the usual approach of binary classification (i.e. constant classes parameters), we compute a single parameter as the mean of the Xfactor parameter image.
- parameters accounting for the antenna pattern: to have a more realistic case, we fitted a second-order polynomial on the ideal parameters images in the range direction. It is the same principle as shown in Figure 4.10, but we directly use the parameters from the Xfactor instead of a first classification.



(a) Garonne



(b) Po (range direction on the vertical axis)



(c) Camargue



(d) Kaw

Figure 4.14 – Data used for the evaluation of the water/land classification methods

- parameters based on the Xfactor (see [Equation 2.14](#)): this represents an ideal scenario. Indeed, the Xfactor is computed from the [DEM](#), which will not be perfectly accurate for real images. Furthermore, the computation of the parameters from the Xfactor requires knowledge of the σ_0 of water and land. While measurements have been performed and the system will be calibrated (see [section 2.4](#)), there will be spatio-temporal variations in σ_0 .

We tested the classification methods jointly estimating the parameters (either region-based or using a Markovian model) using these three initializations. We discuss the influence of the initialization in [subsection 4.5.4](#).

Metrics: In the following, we compute the performances using widely-used metrics: the [True Positive Rate \(TPR\)](#), the [False Positive Rate \(FPR\)](#), the [Error Rate \(ER\)](#) and the [Matthews Correlation Coefficient \(MCC\)](#). They all can be expressed from:

- the number of water pixels correctly classified (TP);
- the number of land pixels correctly classified (TN);
- the number of water pixels wrongly classified (FN);
- and the number of land pixels wrongly classified (FP).

Then, the metrics are defined as follow:

$$\text{TPR} = \frac{TP}{TP + FN}, \quad (4.36)$$

$$\text{FPR} = \frac{FP}{FP + TN}, \quad (4.37)$$

$$\text{ER} = \frac{FP + FN}{TP + FN}, \quad (4.38)$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (4.39)$$

While the [TPR](#) and [FPR](#) are widely used measures, [ER](#) and the [MCC](#) is interesting for us as they are well adapted to problems where the number of elements belonging to one class is much larger than the number of elements belonging to the other class. They both give a single score which is useful for the comparison of the methods.

In the case of the [ER](#), a good classification has a score close to 0, while the [MCC](#) is a number between -1 and 1, where 0 is equivalent to a random classification and 1 is a perfect classification.

4.5.1 Results on Garonne

We show the numerical results for the Garonne image in [Table 4.2](#). The best classification, according to both [ER](#) and [MCC](#) is the one obtained using the [MAP](#) with the statistical prior and the parameters from the Xfactor, which is unexpected. Indeed, it

Method	Estimation	TPR	FPR	Error Rate (water detection)	MCC
MLE	Constant	95.87%	3.80%	418.79%	0.4153
	Polynomial	95.81%	3.91%	430.00%	0.4104
	Xfactor	95.77%	4.61%	506.42%	0.3816
MAP	Constant	90.27%	0.12%	22.32%	0.8890
	Polynomial	90.23%	0.11%	22.26%	0.8893
	Xfactor	90.15%	0.11%	22.13%	0.8897
MRF	Constant	97.08%	0.23%	28.93%	0.8738
	Polynomial	97.02%	0.24%	28.73%	0.8744
	Xfactor	96.98%	0.23%	28.24%	0.8761
MRF (Region-based)	Constant	94.72%	0.18%	24.8%	0.8851
	Polynomial	94.66%	0.18%	24.86%	0.8848
	Xfactor	94.66%	0.18%	24.85%	0.8848
MRF (Markovian)	Constant	97.09%	0.24%	28.94%	0.8738
	Polynomial	97.04%	0.24%	28.73%	0.8744
	Xfactor	97.00%	0.23%	28.25%	0.8760

Table 4.2 – Performance metrics for the Garonne image

does not use the spatial information. It is interesting to note that for this data, **MAP** has the smallest **TPR** (but also the smallest **FPR**). We show in [Figure 4.15](#) the results obtained by the **MAP** with parameters estimated from the Xfactor and the method giving the highest **TPR**. From these results, we can note two things:

- This is a simple case: we can see on the amplitude image that the water/land contrast is very high so spatial regularization to limit the effects of speckle is not crucial. This explains why all the proposed methods (except for the **MLE**) obtain equivalent classification results.
- The false positives are only located at the edges of the river. This comes from the fact that mixed pixels here are labeled as land in the ground truth (see [Table 4.1](#)). It explains why methods having the highest **TPR** also have a higher **FPR**.

Therefore, we can deduce that in favorable cases such as this one, all the presented methods (except for the **MLE**) are equivalent.

4.5.2 Results on Po

The Po dataset also presents a high water/land contrast, but with simulates variations in wind speed over the water surfaces, yielding in variations in backscattering, including regions with so-called dark water (i.e. weak or no signal). The results for this image are presented in [Table 4.3](#). It is interesting to note that the **MRF** with the region-based estimation of the parameters obtain the same results with the different initializations of the parameters maps. We can see that, as expected, the lowest **FPR** is obtained by **MAP** with parameters estimated from the Xfactor and the correct (as used in the simulation) hypotheses on the values of the σ_0 of land and water. Also, the highest

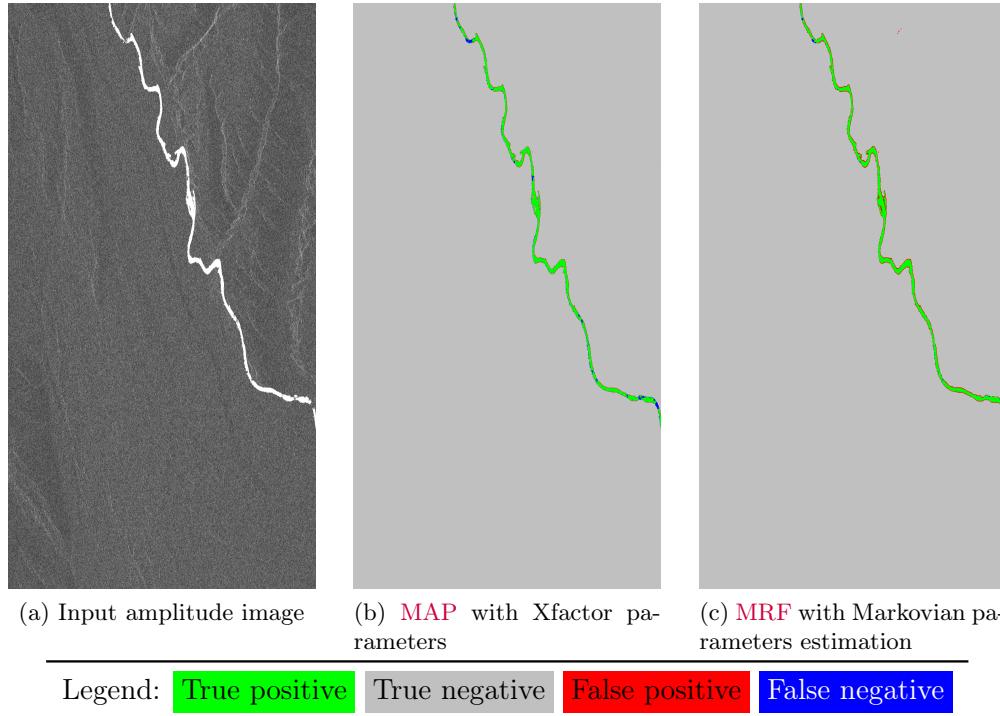


Figure 4.15 – Results on the Garonne image

TPR is obtained by the **MRF** with Markovian parameters estimation. This method also gets the highest **MCC**. We show the visual results of these methods in [Figure 4.16](#). When there is "dark water", it is not taken into account in the Xfactor, which explains why a method estimating the parameters on the image to be classified obtains better results. It appears clearly that the **MRF** with Markovian estimation of the parameters obtains better results than pixel-based methods when the contrast is locally low.

4.5.3 Results on Camargue

For the Camargue image, the Xfactor is not available. Therefore, we show the results of the different methods with constant parameters and the polynomial estimation of the antenna pattern (presented in [Figure 4.10](#)) in [Table 4.4](#). First we can note that as the water/land contrast is lower, using **MRF** as a classification method gives better results. In this example, it is interesting to compare the performances of the **MRF** with no re-estimation and the ones with the Markovian estimation. We can see that while obtaining similar performances, there is a slight improvement when using the Markovian estimation of the parameters with a constant initialization, while we obtain equivalent results when the polynomial initialization is used. Note that the method used to obtain the polynomial parameters can already be considered as an estimation of the parameters on the image with regularization. Visual comparison of the constant **MRF** and the **MRF** with a constant initialization and Markovian estimation of the parameters is presented in [Figure 4.17](#). We can note that we detect more water pixels in the near-nadir part of the image using the Markovian estimation. The land/land layover present at the top-

Method	Estimation	TPR	FPR	Error Rate (water detection)	MCC
MLE	Constant	69.31%	5.61%	119.26%	0.5165
	Polynomial	71.92%	7.03%	139.03%	0.4930
	Xfactor	72.02%	6.83%	135.69%	0.4991
MAP	Constant	53.67%	0.34%	51.60%	0.6859
	Polynomial	53.26%	0.33%	51.95%	0.6833
	Xfactor	53.36%	0.33%	51.86%	0.6839
MRF	Constant	71.41%	0.72%	39.95%	0.7726
	Polynomial	72.45%	0.75%	39.44%	0.7768
	Xfactor	72.93%	0.79%	39.47%	0.7774
MRF (Region-based)	Constant	67.77%	0.52%	40.44%	0.7657
	Polynomial	67.77%	0.52%	40.44%	0.7657
	Xfactor	67.77%	0.52%	40.44%	0.7657
MRF (Markovian)	Constant	72.62%	0.72%	38.75%	0.7806
	Polynomial	73.94%	0.77%	38.32%	0.7846
	Xfactor	74.24%	0.80%	38.40%	0.7847

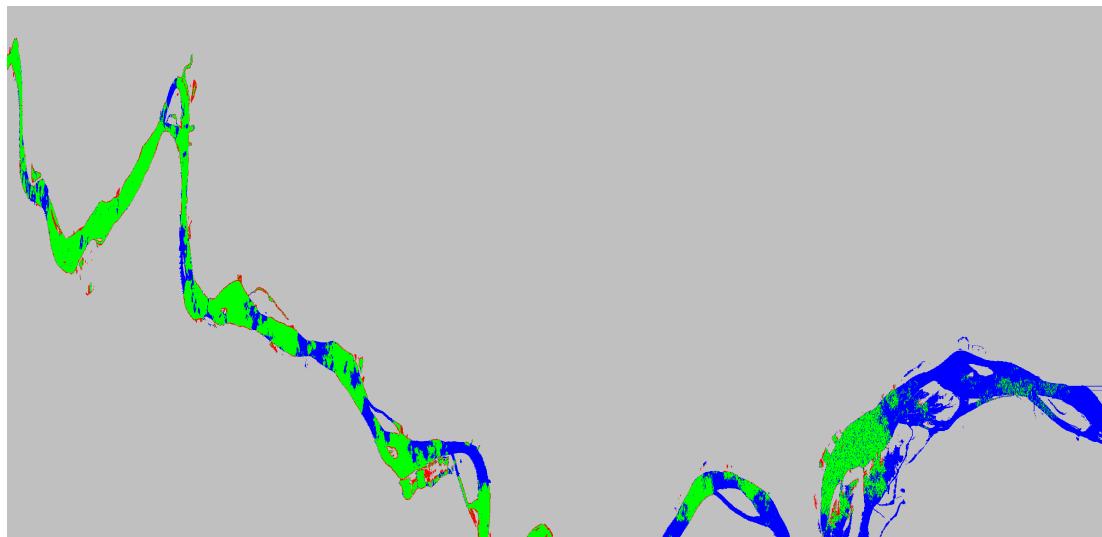
Table 4.3 – Results for the Po image

Method	Estimation	TPR	FPR	Error Rate (water detection)	MCC
MLE	Constant	83.26%	7.54%	54.85%	0.7023
	Polynomial	86.20%	6.38%	46.07%	0.7407
MAP	Constant	39.94%	0.32%	61.69%	0.5817
	Polynomial	46.71%	0.25%	54.56%	0.6387
MRF	Constant	91.27%	2.11%	19.41%	0.8847
	Polynomial	92.53%	1.59%	15.50%	0.9074
MRF (Region-based)	Constant	93.16%	2.95%	21.74%	0.8748
	Polynomial	93.16%	2.95%	21.74%	0.8748
MRF (Markovian)	Constant	91.80%	2.18%	19.22%	0.8863
	Polynomial	92.78%	1.64%	15.52%	0.9074

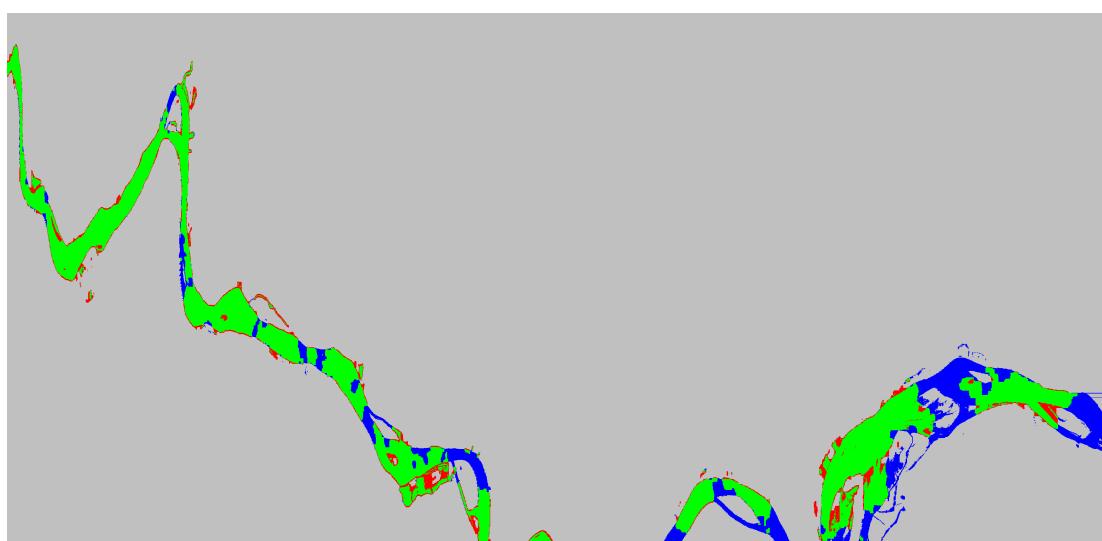
Table 4.4 – Performance metrics for the Camargue image



(a) Input amplitude image



(b) MAP with Xfactor parameters



(c) MRF with Markovian parameters estimation

Legend: True positive True negative False positive False negative

Figure 4.16 – Performance metrics on the Po image

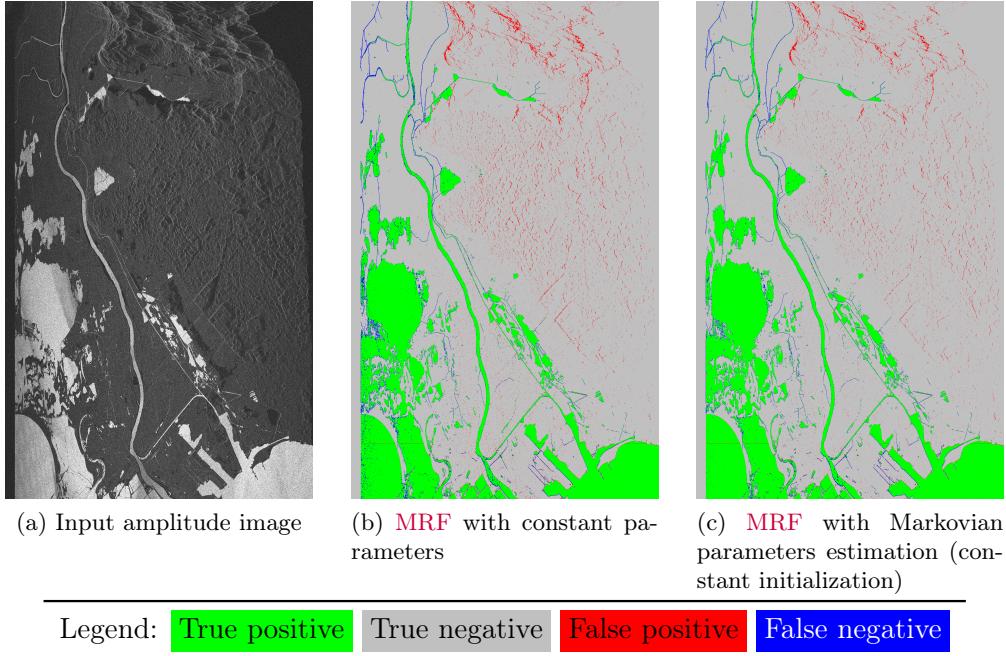


Figure 4.17 – Results on the Camargue image

right of the image accounts for the majority of the false detections, and is a problem for the Markovian estimation of the parameters. Indeed, the initial classification labels these pixels as water, and the re-estimation will lower the parameters of the water class in this region, giving even more false detections.

4.5.4 Results on Kaw (SETHI)

We show the results for the Kaw image in Figure 4.18 and the performance metrics in Table 4.5. Once again, the MRF classification obtain the best results. Here, the best result comes from the region-based estimation method. The fact that this time the region-based estimation has better result than the Markovian can be easily explained

Method	Estimation	TPR	FPR	Error Rate (land detection)	MCC
MLE	Constant	54.49%	7.06%	48.93%	0.4565
	Polynomial	77.28%	10.96%	28.05%	0.6263
MAP	Constant	27.99%	2.85%	73.39%	0.2963
	Polynomial	58.67%	3.39%	42.98%	0.5282
MRF	Constant	79.88%	7.78%	23.90%	0.6820
	Polynomial	98.51%	5.70%	4.26%	0.9346
MRF (Region-based)	Constant	93.28%	2.04%	7.71%	0.8888
	Polynomial	98.52%	2.01%	2.46%	0.9626
MRF (Markovian)	Constant	94.69%	10.43%	10.39%	0.8415
	Polynomial	99.00%	9.58%	5.66%	0.9132

Table 4.5 – Performance metrics for the Kaw image

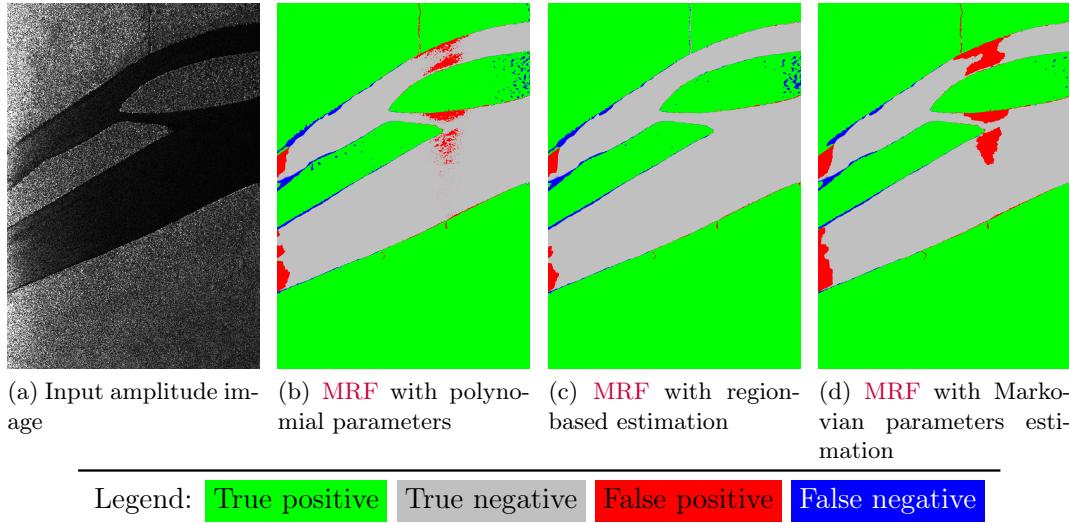


Figure 4.18 – Results on the Kaw image

by looking at the visual results presented in Figure 4.18. Indeed, the classification using the MRF with the polynomial parameters, shown in Figure 4.18(b), has some false positives in the middle of the swath. In the case of a wrong initial classification, the region-based parameter estimation is much more robust.

On this image, we can clearly see that having a good initial classification has a strong importance for the Markovian estimation of the parameters. Indeed, the data term defined in Equation 4.30 explicitly takes into account the classification. An erroneous classification could therefore propagate a wrongly classified region through the iterations. For the region-based method for the estimation of the parameters, it will only play a role during the first iteration. This also explains why the performances of the region-based method have similar results using different initialization methods on the other images.

4.6 Conclusion

We presented different methods for water/land classification in SWOT images. The classification can be based on prior statistical information or on a prior of spatial regularity such as enforced with the binary MRF.

Except in favorable cases where the water/land contrast is very strong, MRF yield better results than pixel-wise ML or MAP. The originality of the proposed methods relies on the use of non-uniform parameters. These parameters can either be set using prior information or estimated in an alternated optimization framework. We showed that using spatially variable parameters improves the results when the radiometries of the classes are non constant. In the case of SWOT images, we have non constant parameters mainly because of the antenna pattern (global variations)

or local variations in the roughness of the surface (local variation). These algorithms have been implemented for the SWOT mission. Our main contribution in this chapter is the Markovian estimation of the parameters presented in subsection 4.4.4.

Future work includes exploiting multi temporal series, adapting the results presented in Part III.

Chapter 5

Detection of narrow rivers

5.1 Introduction

While the methods presented in [chapter 4](#) are able to detect large water bodies, they make the hypothesis that water and land are compact objects. This is generally true when a fine resolution is available, but it does not actually hold in the case of SWOT's coarser resolution (from 60m to 10m in the range direction). This resolution means that many targeted rivers (that have a width greater than 50m) might have a width under one pixel. Hence, the hypothesis made by the Ising model used in the [MRF](#) framework in [chapter 4](#) is not fulfilled. To detect these rivers, a dedicated approach is presented in this chapter.

Detection of thin structures is a widely studied problem in image processing, as it leads to many applications. In ([Amhaz et al., 2016](#)), automatic road cracks detection is performed. It first selects pixels that are likely to belong to the crack network based on their intensities before connecting them. In ([Rossant and others, 2011](#)), a morphological toolchain is used for the vessels detection in eye fundus images. Thin objects detection has also been studied for SAR images in the framework of road detection, for mapping applications. In ([Tupin et al., 1998](#)), a two-step approach is used; first a linear structure detection is performed, then detected lines are connected using a high-level step. This method has been applied to water detection in ([Cao et al., 2011](#)) with the addition of a multiscale step, contrast inversion and a tolerance for curved line. A morphological approach has also been applied to simulated SWOT images in ([Grosdidier and others, 2012](#)).

In this chapter, we present a two-step approach. The outline of the proposed method is given in [section 5.2](#). We present in [subsection 5.2.1](#) the single-scale pixel-based detection, and the object-based processing that follows is presented in [subsection 5.2.2](#). Finally, results are presented in [section 5.3](#).

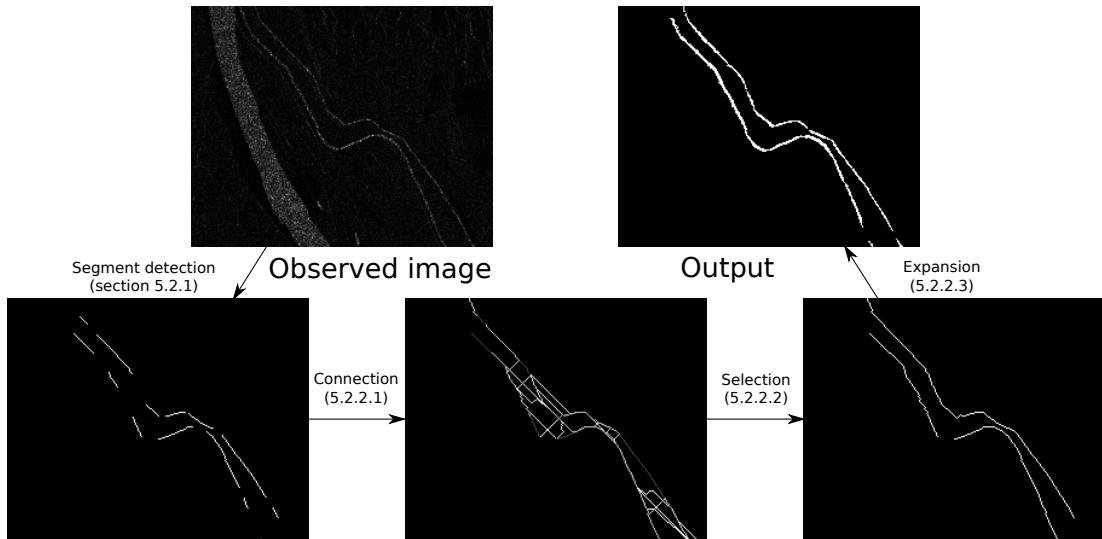


Figure 5.1 – Overview of the proposed toolchain

5.2 Two-step approach

Based on the literature presented in [section 5.1](#), we decided to propose a two step approach for the detection of thin elements in SWOT amplitude images:

1. Line segment detection (pixel-based step): in the first step, the goal is to detect straight portions of rivers based on the radiometric values of the pixels. We follow the approach proposed by ([Tupin *et al.*, 1998](#)).
2. Connection and selection step (object-based step): as only portions of the river network are detected during the first step (either because some parts are not straight lines or because the contrast is locally low), a second step is applied; candidate pairs of segments are connected, then the connections are chosen based on a [MRF](#) model enforcing basic geometric properties of rivers. Finally an expansion step is performed in order to go back from objects to pixels.

This approach is illustrated in [Figure 5.1](#) and the first step is recalled in [subsection 5.2.1](#), while the second step is described in [subsection 5.2.2](#).

5.2.1 Low-level step

Principle In this section, our objective is to find small segments that are likely to belong to a river. Because of speckle, we can only expect to retrieve portions of rivers while keeping a low false alarm rate when working on small regions. These segments will then be used as an initialization to the second step of the proposed method ([subsection 5.2.2](#)).

The general process is described in [Figure 5.2](#). We test several rectangles as segment candidates: at each pixel, N_d directions are tested, each time with N_w different rectangle widths. In an image of \mathcal{S}_i pixels, $\mathcal{S}_i \times N_w \times N_d$ rectangles are tested. A score related to the probability of belonging to a part of a river is then assigned to each rectangle.

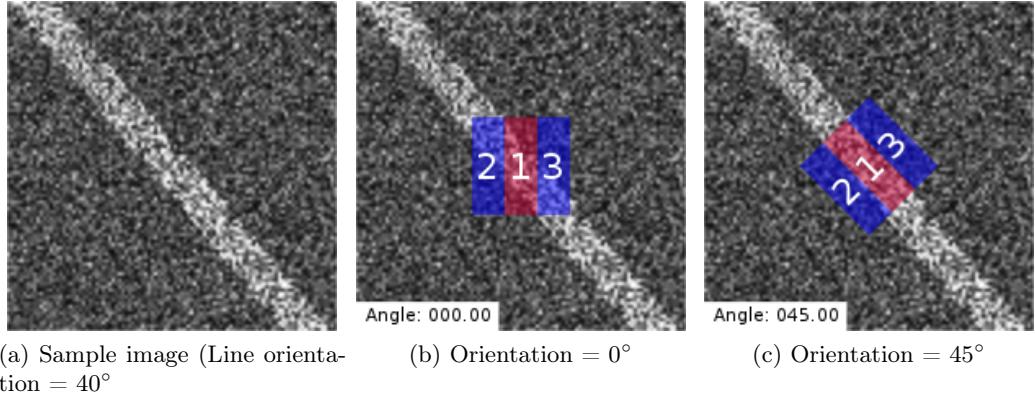


Figure 5.2 – Illustration of the line detector. To detect a line, a rectangular detector is applied at every pixel of the image. $N_d = 16$ directions and $N_w = 5$ widths are tested. In this example the 16 directions are tested with a fixed width. A higher score for the orientation at 45° than for the orientation at 0° will be obtained. See www.sylvainlobry.com/phd for the animated image.

The computation of this score is described in [subsubsection 5.2.1.1](#). For every pixel, the rectangle centered on it having the best score is selected. A thresholding is then applied to select pixels that are likely to belong to a river, and a post-processing step is applied. This part is described in [subsubsection 5.2.1.2](#).

5.2.1.1 Line detector

In this subsection, we briefly describe the computation of a score for a given rectangle (centered on pixel s of direction d and width w). This score is computed by comparing statistics computed in this rectangle (region 1 in [Figure 5.2](#) which will be named r_1 in the following) to those of the two rectangles surrounding it (regions 2 and 3 in [Figure 5.2](#) which will be named r_2 and r_3 respectively in the following).

D1 detector: D1 is a simple line detector based on a ratio edge detector (see ([Touzi et al., 1988](#))). Following the work of ([Tupin et al., 1998](#)), a normalized ratio edge detector between r_1 and r_2 can be defined as:

$$re(r_1, r_2) = 1 - \min \left(\frac{\mu_{r_1}}{\mu_{r_2}}, \frac{\mu_{r_2}}{\mu_{r_1}} \right), \quad (5.1)$$

where μ_{r_x} is the distribution parameter in the rectangle r_x (which is obtained using the [MLE](#) and is given by the mean for an intensity image, which is Gamma distributed). Based on re , the line detector $D1$ can be obtained:

$$D1(r_1) = \min (re(r_1, r_2), re(r_1, r_3)) . \quad (5.2)$$

This detector response will be close to 1 where both contrasts between r_1 and r_2 and between r_1 and r_3 are high. Note that it can be simplified to:

$$D1(r_1) = 1 - \max \left(\min \left(\frac{\mu_{r_1}}{\mu_{r_2}}, \frac{\mu_{r_2}}{\mu_{r_1}} \right), \min \left(\frac{\mu_{r_1}}{\mu_{r_3}}, \frac{\mu_{r_3}}{\mu_{r_1}} \right) \right) \quad (5.3)$$

D2 detector: D2 is a line detector based on the normalized cross correlation between r_1 and r_2, r_3 . In (Tupin *et al.*, 1998), it is defined as follows:

$$D2(r_1) = \min (cc(r_1, r_2), cc(r_1, r_3)) , \quad (5.4)$$

with cc the discrete normalized cross-correlation computed as:

$$cc(r_x, r_y) = \sqrt{\frac{n_x n_y (\mu_x - \mu_y)^2}{1 + (n_x + n_y)(n_x \sigma_x^2 + n_y \sigma_y^2)}}, \quad (5.5)$$

with n_x the number of pixels in r_x and σ_x^2 the variance of r_x . Compared to $D1$, $D2$ has the advantage of being resilient when a bright scatterer is present in one of the rectangles to be tested.

Fusion: To use both detectors, a symmetric associative sum is used (Bloch, 1996):

$$D1D2(r_1) = \frac{\overline{D1}(r_1)\overline{D2}(r_1)}{1 - \overline{D1}(r_1) - \overline{D2}(r_1) + 2\overline{D1}(r_1)\overline{D2}(r_1)}, \quad (5.6)$$

where $\bar{x}(r)$ is the score $x(r)$ centered between $[0, 1]$. This score is associated to one of the oriented rectangles tested for each pixel, but we want to have a single score at each pixel. This can be obtained by selecting the maximum score from the different orientations tested for a given pixel:

$$l_i = \max_{d_i \in \mathcal{D}_i} D1D2(d_i), \quad (5.7)$$

where \mathcal{D}_i is the set of orientations centered in i . The direction of the rectangle giving the maximum score at pixel i is also saved in d_i , giving an image of directions \mathbf{d} .

5.2.1.2 Post-processing:

While \mathbf{l} contains a score for each pixel reflecting the probability to be a line, our objective at this step is to obtain line segments. Therefore several operations are applied to this image:

- Thresholding of the score image \mathbf{l} to obtain a binary map;
- removal of isolated detections;
- segment selection.

Each of these steps is briefly explained below.

Thresholding: The symmetrical associative sum defined in Equation 5.6 has the following properties:

- conjunctive: $\max(\overline{D1}(r), \overline{D2}(r)) < 0.5 \Rightarrow D1D2(r) < \min(\overline{D1}(r), \overline{D2}(r)) < 0.5$;
- disjunctive: $\min(\overline{D1}(r), \overline{D2}(r)) > 0.5 \Rightarrow D1D2(r) > \max(\overline{D1}(r), \overline{D2}(r)) > 0.5$;
- compromise: $\overline{D1}(r) \leq 0.5 \leq \overline{D2}(r) \Rightarrow \overline{D1}(r) \leq D1D2(r) \leq \overline{D2}(r)$.

As a first approximation, using a threshold equals to 0.5 on D is adapted to our problem as it will guarantee that at least one of the two detectors $\overline{D1}(r)$ or $\overline{D2}(r)$ has a value larger than 0.5.

Removal of isolated detections: On the binary image representing segment detections, a first simple approach is applied to remove isolated detections; for pixel i , we explore the neighbors restricted to the direction d_i . If no detection with a similar (the same or ± 1 in the quantified space of the directions) associated direction is found, the pixel is considered as isolated and therefore removed from the binary detection image.

Furthermore, a local Hough transform is applied. On overlapping blocks of size 20×20 a simple Hough transform restricted to N_d directions is applied. Each detected pixel in the window vote for its associated direction. Only pixels belonging to the line getting the maximum number of votes are kept.

Obtaining segments: A simple pixel to segment algorithm is then applied: for each pixel i , we locally search pixels with similar directions in a neighborhood restricted to its direction d_i . These pixels are then labeled giving a first set of segments. Finally, small segments are removed, giving the final set of segments $\mathcal{S}eg$.

Summary: Low-level step

We have presented a method for the detection of segments at the pixel level mainly inspired from the works of (Tupin *et al.*, 1998) and (Cao *et al.*, 2011). It is based on two detectors (one using a ratio between regions and the other the cross correlation) that are fused and on a simple post-processing chain. Detected segments can then be used as input for object-based detection.

5.2.2 Connection and regularization

5.2.2.1 Connection

At this stage, we have a set of segments $\mathcal{S}eg$ where most of the elements represent river parts. Our goal is now to find a connection between pairs of segments in $\mathcal{S}eg$. When both segments belong to the same river, we seek a connection that will also belong to this river. Therefore, we will look for the shortest path between two segments with a distance that take into account the probability for the pixel to belong to the river. Whereas a geodesic distance was used in (Perciano and others, 2016), we use a distance

derived from the score obtained in the low-level step, allowing to take into account the spatial context at a larger scale. This gives the following distance definition for two neighbors:

$$\forall a \in \mathcal{N}_b, \quad D_r(a, b) = 1 - l_b \quad (5.8)$$

where \mathcal{N}_b is the set of pixels that are neighbors to b (in this case, each pixel is a neighbor for its 8 surrounding ones). Following graph theory, it can be generalized between any two pixels a and b (not necessarily neighbors) as the shortest path between a and b .

To obtain the shortest path, we propose to use Dijkstra's algorithm originally described in (Dijkstra, 1959). This algorithm finds the shortest path between a source node and all other nodes in a directed graph, where the distance to go from node a to b is the weight of the vertex going from a to b . In our case the nodes are the pixels of the image, and the weight from node a to b is $D_r(a, b)$ defined in Equation 5.8. Since the final distance is not normalized by the length of the connection, long connections are penalized. The algorithm, when using a min-priority queue implemented with a Fibonacci heap (Fredman and Tarjan, 1987), gives a complexity in $O(|E| + |V| \log |V|)$ where E is the set of edges and V the set of nodes. Note that in our case, the number of edges is constant with respect to the number of nodes: $|E| = 8|V|$. This gives a total complexity of:

$$O(|V| \log |V|). \quad (5.9)$$

As running this algorithm for each pair of segment end-points on the whole image is computationally expensive, we do the following steps (both illustrated in Figure 5.3):

- During the graph construction, we restrict the space search to a circle with a diameter equals to the distance between the two extremities of the segments to be connected;
- We restrict the connections to "compatible" pairs of segments. The compatibility between two segments is defined by the proximity of their extremities (in our experiments, we have found that a distance of 80 pixels works well).

In the following, the set of the connections between every pair of compatible segments is called \mathcal{C}_o .

5.2.2.2 Selection

While the union of $\mathcal{S}eg$ and \mathcal{C}_o is likely to contain most parts of the rivers, it is clear that \mathcal{C}_o may contain many unwanted connections (as it can be seen in Figure 5.1). In this step, we want to select meaningful connections with respect to their distance defined in Equation 5.8 and to their contributions to priors that can be made regarding a river network:

- a river network should have few end-points;
- it has few intersections;

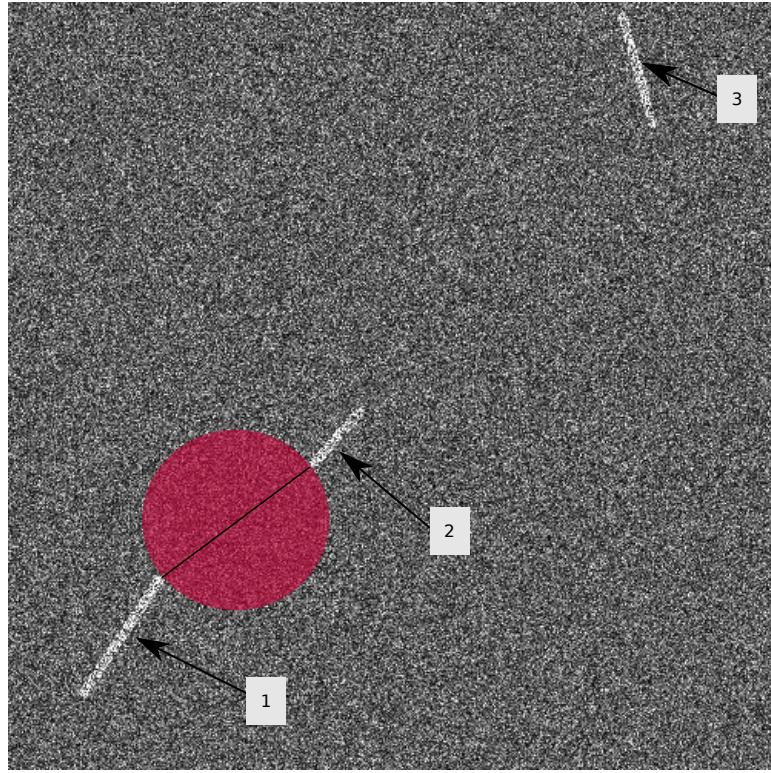


Figure 5.3 – Connection step. We look for a connection between segments 1 and 2, as they are close enough. On the contrary, segment 3 is too far from both 1 and 2. The space search in which Dijkstra’s algorithm is applied is here restricted to the red circle.

- in most cases, an end-point of a segment has only one connection to another segment;
- connections only fill gaps; therefore they should be short;
- long segments strongly indicate the presence of rivers so they should be part of the network;
- it is moderately curved.

To enforce these priors, we use the **MRF** framework. Our goal is to find a labeling \mathbf{x} of $\mathcal{C}o$ such that:

$$\forall co \in \mathcal{C}o, \quad x_{co} = \begin{cases} 1 & \text{if } co \text{ belongs to the network,} \\ 0 & \text{otherwise.} \end{cases} \quad (5.10)$$

Note that, the set of segments detected at the low-level step $\mathcal{S}eg$ is kept intact. We propose to define an energy of the form:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \mathcal{E}(\mathbf{x}) \quad (5.11)$$

$$= DT(\mathbf{l}, \mathbf{x}) - \log(p(\mathbf{x})). \quad (5.12)$$

For the dataterm, we will use once again the score from the low-level step. However, as the length of the connection is taken into account in the prior, we normalize the score of the connection (changing the meaning of the score from the probability for it to be a good connection to the general quality of the connection):

$$DT(\mathbf{l}, \mathbf{x}) = \sum_{co \in \mathcal{C}_o} DT(\mathbf{l}, x_{co}), \quad (5.13)$$

where:

$$DT(\mathbf{l}, x_{co}) = \begin{cases} \frac{1}{|co|} \sum_{i \in co} (1 - l_i) & \text{if } x_{co} = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (5.14)$$

The prior term $-\log(p(\mathbf{x}))$ needs to encode the six assumptions made about the shape of a river network. This will be done using six terms:

$$\begin{aligned} -\log(p(\mathbf{x})) = & \mathcal{K}_{\text{end}} \mathcal{P}_{\text{end}}(\mathbf{x}) + \mathcal{K}_{\text{int}} \mathcal{P}_{\text{int}}(\mathbf{x}) + \mathcal{K}_{\text{over}} \mathcal{P}_{\text{over}}(\mathbf{x}) + \mathcal{K}_{\text{L co}} \mathcal{P}_{\text{L co}}(\mathbf{x}) \\ & + \mathcal{K}_{\text{L seg}} \mathcal{P}_{\text{L seg}}(\mathbf{x}) + \mathcal{K}_{\text{align}} \mathcal{P}_{\text{align}}(\mathbf{x}). \end{aligned} \quad (5.15)$$

In the following paragraphs, we will explain for each of these terms how they enforce the priors we set. We use the following indicator functions:

$$\delta(x) = \begin{cases} 1 & \text{if } x = 0; \\ 0 & \text{otherwise.} \end{cases} \quad (5.16)$$

$$\bar{\delta}(x) = \begin{cases} 0 & \text{if } x = 0; \\ 1 & \text{otherwise.} \end{cases} \quad (5.17)$$

Also, we use the following convention for logical expressions: true is 1, false is 0. For the sake of simplicity, we will define the images \mathbf{I}_x and \mathbf{I}_{Seg} representing the connection and segments in the image domain:

$$I_{x,i} = \sum_{\substack{co \in \mathcal{C}_o \\ x_{co}=1}} \bar{\delta}(i \in co) \quad (5.18)$$

$$I_{\text{Seg},i} = \begin{cases} 1 & \text{if } \exists se \in \mathcal{S}eg \text{ s.t. } i \in se \\ 0 & \text{otherwise.} \end{cases} \quad (5.19)$$

Note that while the image representing segments \mathbf{I}_{Seg} is a binary image indicating the presence of a segment, the image representing connections \mathbf{I}_x indicates the number of connections including each pixel (this number is not limited to 1, as it can be seen in [Figure 5.1](#)). In the following, the terms will be defined globally. Note that it is straightforward to compute them locally from the global definition.

A river network should have few end-points ($\mathcal{K}_{\text{end}}/\mathcal{P}_{\text{end}}$) In a river network, endpoints generally correspond to sources or a connection with a compact region. With respect to the number of detected segments, end-points should be sparse. Therefore, a term \mathcal{P}_{end} is added to penalize end-points in the network. To compute this term, we count the numbers of segment extremities that are not linked to a connection in the current labeling \mathbf{x} :

$$\mathcal{P}_{\text{end}}(\mathbf{x}) = \sum_{se \in \mathcal{S}_{\text{eg}}} (\delta(I_{\mathbf{x}, ex1(se)}) + \delta(I_{\mathbf{x}, ex2(se)})) , \quad (5.20)$$

where $ex1(se)$ and $ex2(se)$ are functions returning the pixels of both extremities of se .

A river network should have few intersections ($\mathcal{K}_{\text{over}}/\mathcal{P}_{\text{over}}$) It should be rare to have intersections in a river network. To enforce this property, we use a term $\mathcal{P}_{\text{over}}$ that counts the number of pixels appearing in more than one selected connection or in a connection and a segment:

$$\mathcal{P}_{\text{over}}(\mathbf{x}) = \sum_i (\bar{\delta}(I_{\mathcal{S}_{\text{eg}}, i} \text{ and } I_{\mathbf{x}, i}) + \bar{\delta}(I_{\mathbf{x}, i} > 1) \times I_{\mathbf{x}, i}) , \quad (5.21)$$

where "and" is the pixel-wise logical operator.

In a river network, an end point of a segment has usually only one connection to another segment ($\mathcal{K}_{\text{int}}/\mathcal{P}_{\text{int}}$) Except for bifurcations, segments only have one connection to another segment. As this is a "rare" event in river networks, this is penalized:

$$\mathcal{P}_{\text{int}}(\mathbf{x}) = \sum_{se \in \mathcal{S}_{\text{eg}}} (\bar{\delta}(I_{\mathbf{x}, ex1(se)} > 1) \times I_{\mathbf{x}, ex1(se)} + \bar{\delta}(I_{\mathbf{x}, ex2(se)} > 1) \times I_{\mathbf{x}, ex2(se)}) \quad (5.22)$$

In a river network, connections only fill gaps; therefore they should be short ($\mathcal{K}_{\text{L co}}/\mathcal{P}_{\text{L co}}$) Connections are only meant to fill gaps between detected segments at the first step. Therefore, we want their cardinals to be small. This is enforced using:

$$\mathcal{P}_{\text{L co}}(\mathbf{x}) = \sum_{\substack{co \in \mathcal{C}_{\text{o}} \\ x_{co}=1}} |co| \quad (5.23)$$

In a river network, long segments strongly indicate the presence of rivers so they should be part of it ($\mathcal{K}_{\text{L seg}}/\mathcal{P}_{\text{L seg}}$) Inversely, it means that we want to avoid connecting small segments (that are more likely to be false detections). Each connected segment is then penalized by a term $\mathcal{P}_{\text{L seg}}$ decreasing with the length of the segment:

$$\mathcal{P}_{\text{L seg}}(\mathbf{x}) = \sum_{se \in \mathcal{S}_{\text{eg}}} \frac{\bar{\delta}(I_{\mathbf{x}, ex1(se)})}{|se|} + \frac{\bar{\delta}(I_{\mathbf{x}, ex2(se)})}{|se|} \quad (5.24)$$

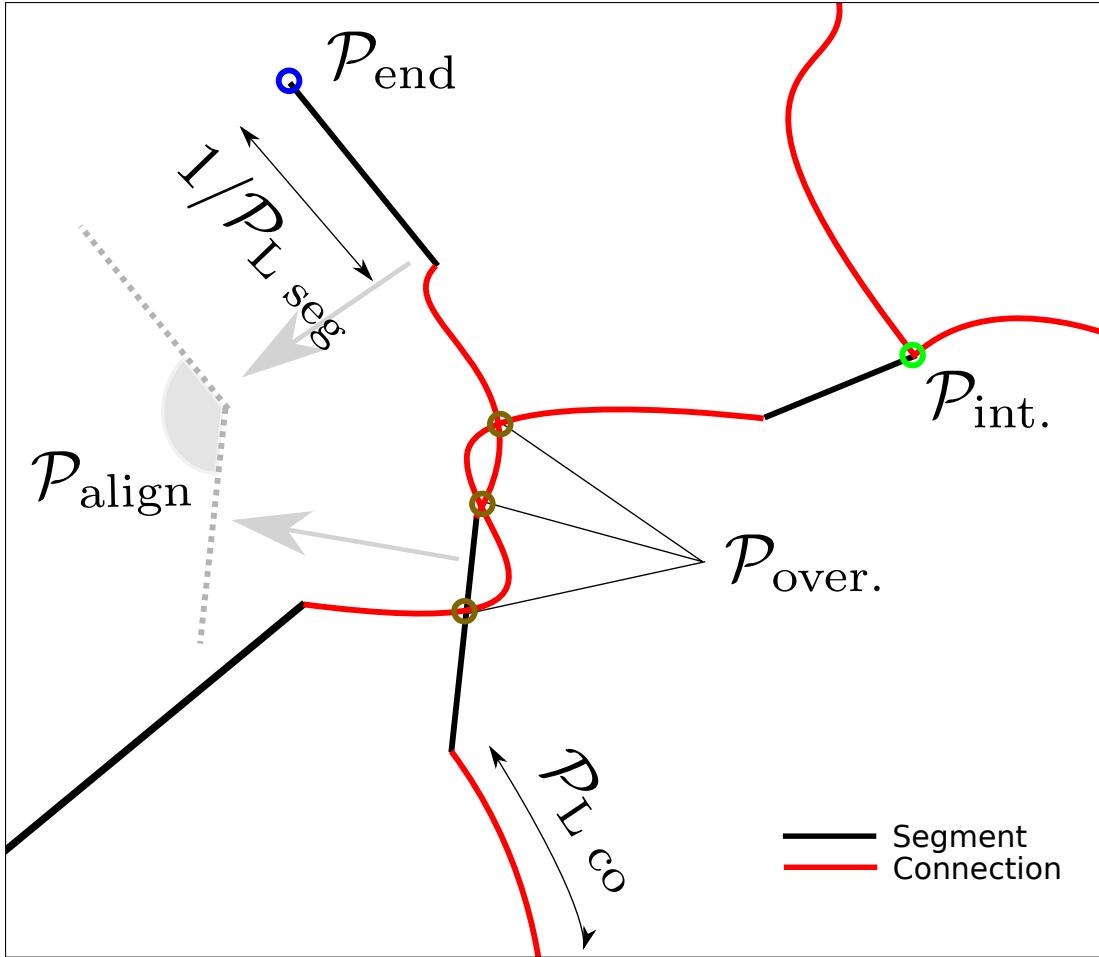


Figure 5.4 – Illustration of the prior terms introduced in Equation 5.15.

Rivers are moderately curved ($\mathcal{K}_{\text{align}}/\mathcal{P}_{\text{align}}$) While rivers can have high sinuosity at a large scale, between two segments the sinuosity index is expected to be low. To enforce this property, we define the term $\mathcal{P}_{\text{align}}$:

$$\mathcal{P}_{\text{align}}(\mathbf{x}) = \sum_{\substack{co \in \mathcal{C}_o \\ x_{co}=1}} \angle(\text{seg1}(co), \text{seg2}(co)), \quad (5.25)$$

Where functions $\text{seg1}(co)$ and $\text{seg2}(co)$ return the two segments connected by co .

These terms are illustrated on a simple example in Figure 5.4.

Optimization While being a binary labeling problem, the selection process cannot be easily optimized by a graph-cut algorithm as the neighborhood configuration is not regular. To allow for fast computation, we use the ICM algorithm presented in subsection 3.2.2. This method finds the closest local minimum from the initial solution provided. To initialize it, we use a simple threshold on $DT(\mathbf{l}, x_{co})$ defined in Equation 5.14.

At the end of this step, we note $\hat{\mathcal{C}}_o$ the set of connections such that $x_{co} = 1$.

5.2.2.3 Expansion

The union \mathcal{RN} of \mathcal{Seg} and $\hat{\mathcal{Co}}$ indicates the presence of rivers as chains of pixels with a width of 1 pixel. To recover a pixel-based classification and improve the positioning of the rivers, a local classification step is needed. Several approaches can be applied, notably level-sets based methods presented in section 4.2.

In our case, we decided to use a simple approach based on a denoising step, illustrated in Figure 5.5. This method has the advantage of being both simple, and does not require new parameters. It is based on the following steps:

1. Connected component labeling is performed on \mathcal{RN} . This can be done efficiently with a two-pass algorithm.
2. For each connected component, we extract a rectangular region, with a small padding (in our implementation, it is set to 20 pixels), see Figure 5.5a.
3. We apply a denoising method on the region. In our case, we apply NL-SAR (Deledalle *et al.*, 2015). After this step, we obtain the result shown in Figure 5.5c
4. The denoised region is thresholded based on the expected water radiometry (either obtained theoretically or from an iterative process, see section 4.4). This gives the result shown in Figure 5.5d.
5. A connected component labeling is performed on the binary image.
6. We select the components from the binary image that have large intersections with \mathcal{RN} (in our implementation, more than 10% of the size of the connected component). This is illustrated in Figure 5.5e.

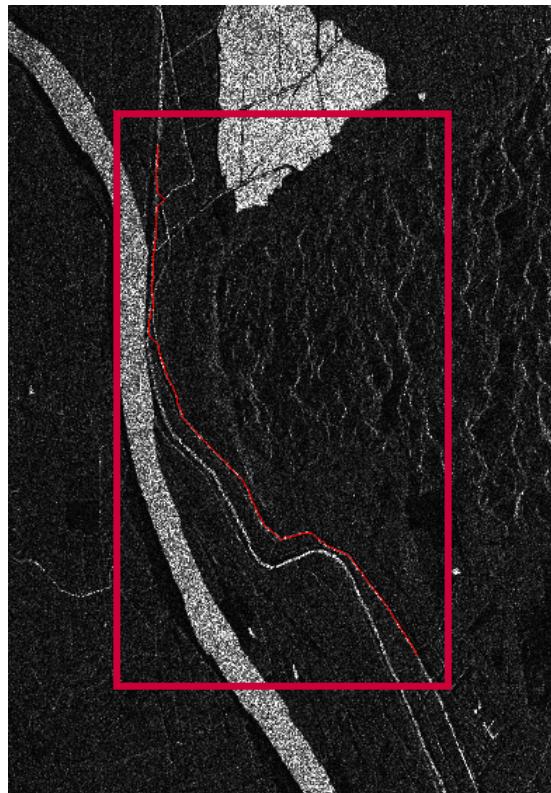
As we do not want to suppress previously selected connections, the response from this step is combined with the output from the selection step (using a simple logical or on the binary maps).

Summary: Object-based detection

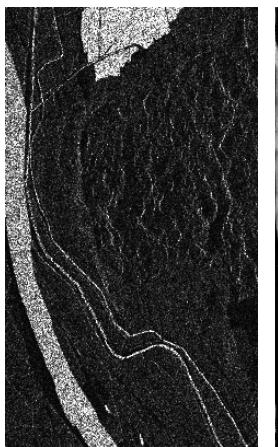
Detected segments from the pixel-based step only cover parts of the river network. To retrieve it, we use a toolchain taking as inputs the set of segments. First, connections between candidates pairs are performed using Dijkstra's algorithm. Then, these connections are selected based on their contribution to the global river network. This is done using a MRF defined on the connections. Finally, a simple expansion algorithm is applied to recover a pixel-based detection.

5.3 Results

We evaluate the proposed method on the Camargue image presented in Figure 4.14(c). This image is interesting for our algorithm as it contains many narrow rivers. The results are shown in Figure 5.6 and a zoom on a region of interest is shown in Figure 5.7.



(a) Input image



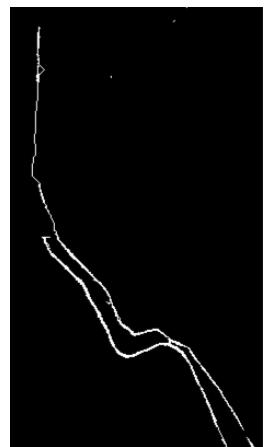
(b) Input image



(c) Denoised image



(d) Thresholded image



(e) Output

Figure 5.5 – Illustration of the expansion step. NL-SAR denoising is applied locally for each connected component of the river network $\mathcal{R}N$. This denoising is then thresholded and used to obtain the pixel-based classification.

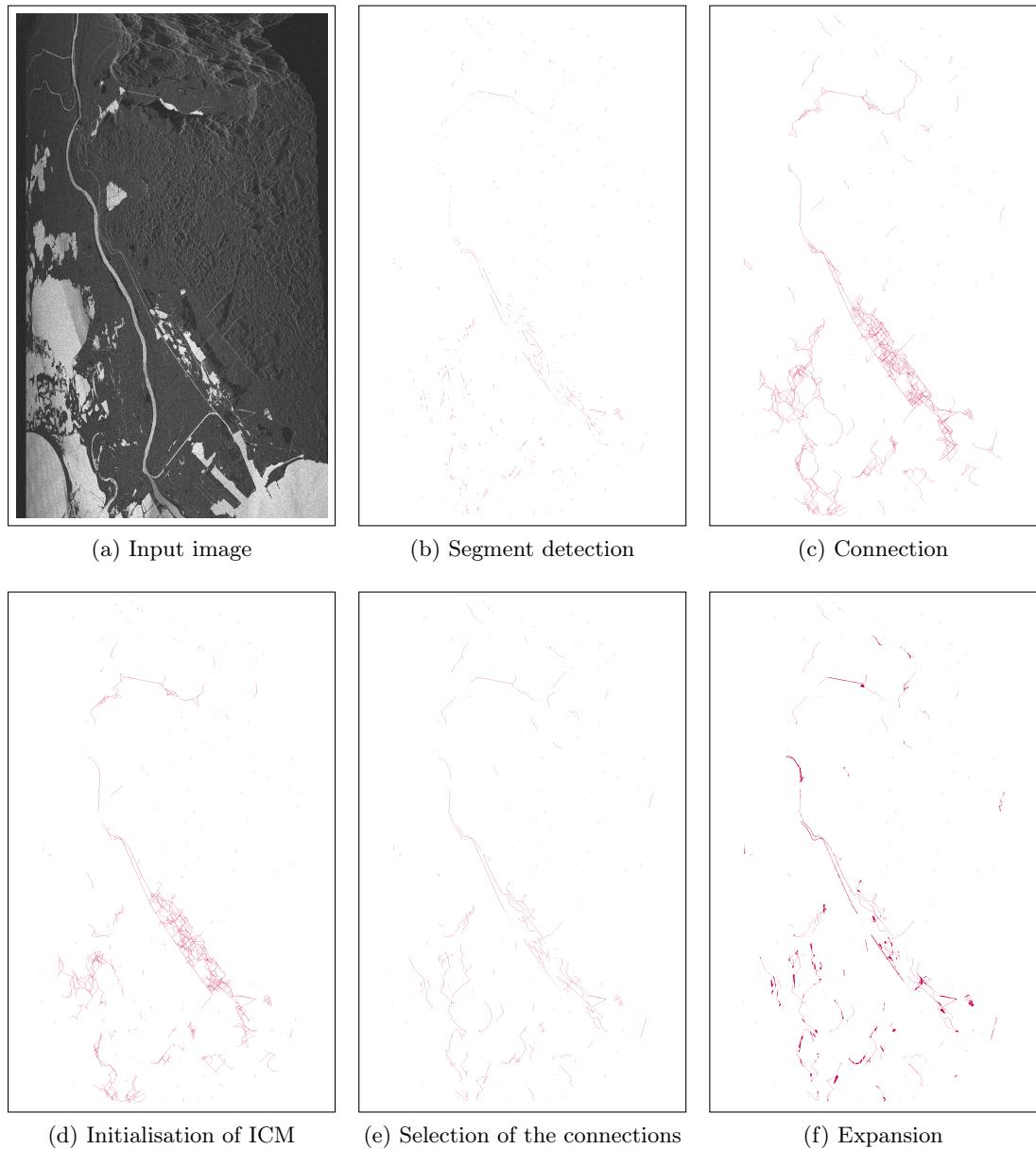


Figure 5.6 – Illustration of the different steps of the proposed approach on the Camargue image.

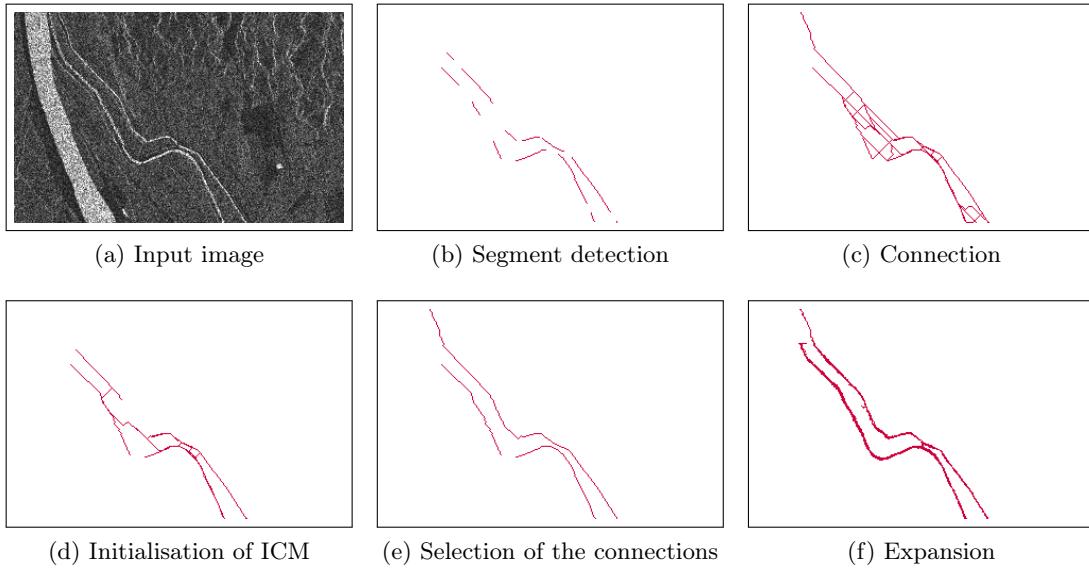


Figure 5.7 – Illustration of the different steps of the proposed approach on a zoom of the Camargue image.

Method	TPR	FPR	Error Rate (water detection)	MCC
MRF (Markovian)	92.78%	1.64%	15.52%	0.9074
MRF + Rivers	93.08%	1.69%	15.46%	0.9080

Table 5.1 – Results for the Camargue image

The segment detection has been tuned so that it has only a limited number of false detections. The connection step finds a link between all the close segments. The selection is based on an initialization made by a simple thresholding of the associate score of each connection. Finally, the expansion step recovers a pixel-based detection.

This algorithm is designed to detect only narrow elements. Therefore we evaluate it combined with the results obtained with the MRF with Markovian parameters estimation presented in chapter 4. To combine the results, we simply use a logical "or". The numerical results (using the metrics defined in section 4.5) are presented in Table 5.1 and visual results are presented in Figure 5.8. We can see that the global water error rate is only improved by a fraction (-0.06%) with the addition of the rivers, which is not surprising considering that narrow rivers only account for a small fraction of the pixels in the image.

5.4 Conclusion

We presented a method combining two main steps for the detection of thin elements in SAR images. The first step performs segment detection on small regions, while the second one connects these detections in a meaningful way with respect to the

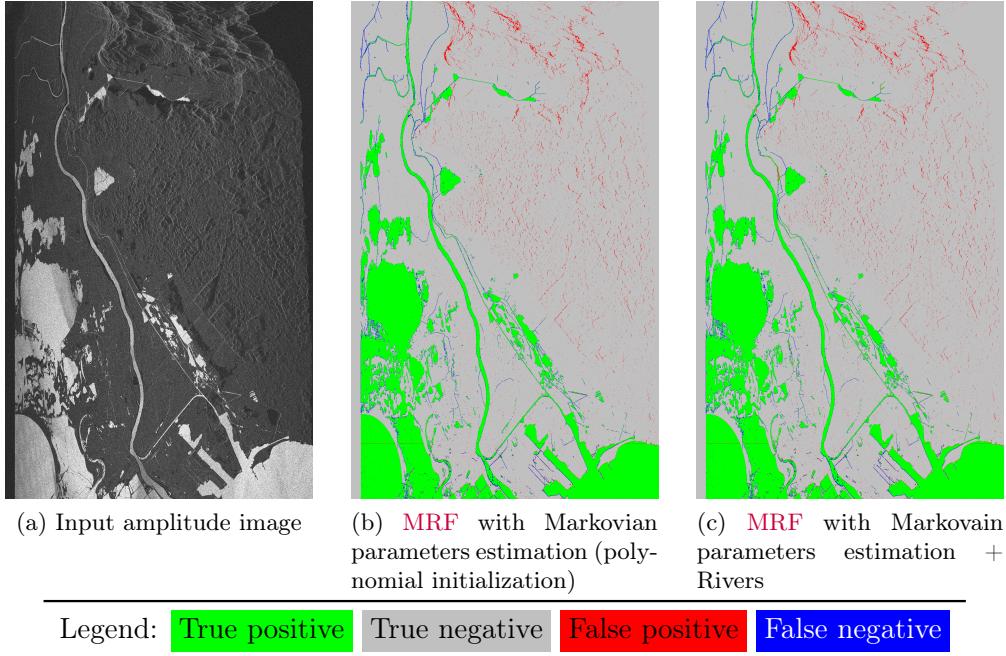


Figure 5.8 – Results on the Camargue image

geometrical properties of river networks. This algorithm has been designed for the detection of narrow rivers in SWOT images. We show some results on simulated SWOT images. For this chapter, our main contribution is the definition of the different steps of the object-based detection. While the connexion step is based on a widely used algorithm (Dijkstra), the distance definition has been defined from the low-level measures. The selection step uses original priors and is completed by a simple expansion step. From the results, it occurs that the algorithm should not necessarily be run systematically, but only when narrow rivers are of particular interest. Indeed, the improvement of the global error rate is, as expected, marginal.

While this algorithm detects narrow rivers that would not be retrieved with a classical contextual method (such as the ones presented in [chapter 4](#)), it is computationally expensive. Further work should be devoted to the simplification of the proposed method so it could be run systematically in the SWOT processing chain. Also, it requires an extensive and currently manual parameters tuning. Each step should be improved by either finding an alternative algorithm requiring less parameters, or finding a way to automatically tune them.

Part III

Processing of multi-temporal series of SAR images

Chapter 6

Decomposition models for multi-temporal series of SAR images

6.1 Introduction

While the work presented in [Part II](#) was dedicated to the simple case of binary classification in single amplitude images, we now tackle the case of multi-temporal SAR images of urban areas. In these images, we will target the following applications:

- strong scatterers detection;
- regularization;
- change detection.

We have seen in [section 2.3](#) that SAR images suffer from strong fluctuations related to the speckle phenomenon. While containing information about the sub-resolution texture of the scene, speckle is often regarded as an undesirable noise for image interpretation tasks. Speckle is then typically modeled as a multiplicative noise. Another peculiarity of SAR images of urban areas is their high dynamic range: man-made structures such as buildings, fences or transmission towers produce very strong back-scatterings, with intensities much larger than the surrounding areas. This is illustrated in [Figure 6.1](#). Such scatterers are especially numerous in urban areas. Therefore, it can be interesting to take into account the strong scatterers when regularizing images. In this chapter, we will propose a method which is able to jointly detect the strong scatterers and regularize the image.

The simplest way to reduce speckle fluctuations is spatial multi-looking which amounts to averaging pixel values within a sliding window. This speckle variance reduction is obtained at the cost of a resolution loss proportional to the size of the averaging

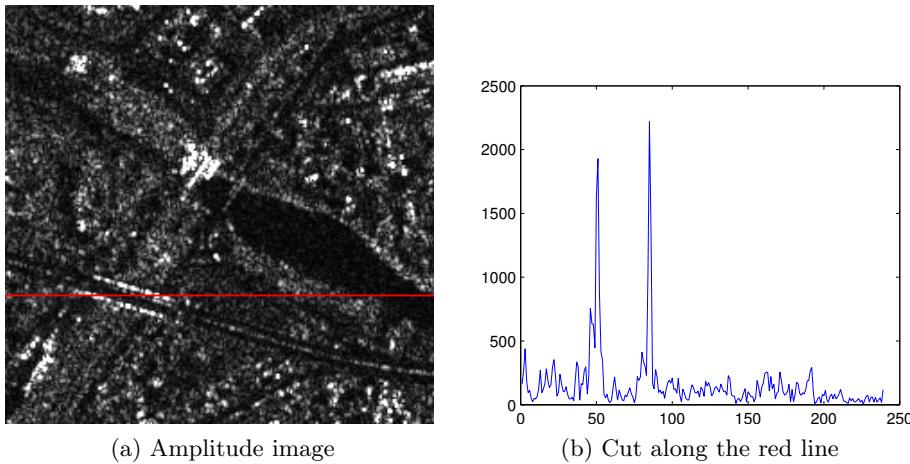


Figure 6.1 – Illustration of the speckle and scatterers phenomena on an amplitude image acquired by TerraSAR-X (Stripmap mode) in Saint-Gervais, France (GPS coordinates: 45.910877, 6.705635). Points corresponding to sides of the bridge have a radiometry of approximately 10 times those of surrounding points.

window. While it may produce satisfying results in homogeneous areas, it strongly blurs textured areas, edges between regions and bright scatterers.

Numerous approaches have been proposed to prevent the introduction of blur by mixing values from distinct regions. Lee *et al.* (Lee *et al.*, 2003) locally select the best window among a few oriented windows. The IDAN algorithm (Vasile *et al.*, 2006) builds an adaptive window by region growing. Several methods have been derived from the non-local means (NL-means) approach by Buades *et al.* (Buades *et al.*, 2005). These methods select similar pixels in an extended window based on patch-similarity (Deledalle *et al.*, 2009; Deledalle *et al.*, 2011; Chen *et al.*, 2011; Zhong *et al.*, 2011; Parrilli *et al.*, 2012; Cozzolino *et al.*, 2014; Deledalle *et al.*, 2015). Another family of methods reduce speckle by regularization, i.e., by computing the maximum a posteriori estimate under a given prior. Wavelet-based approaches model the distribution of wavelet coefficients (Achim *et al.*, 2003; Argenti *et al.*, 2006; Xie *et al.*, 2002b). Total variation (TV) regularization penalizes variations between neighboring pixels while preserving sharp edges (Rudin *et al.*, 1992). Total variation has been applied to the regularization of SAR amplitudes (Aujol *et al.*, 2003; Denis *et al.*, 2009; Palsson *et al.*, 2012), SAR intensities (Aubert and Aujol, 2008; Steidl and Teuber, 2010) and log-transformed intensities (Bioucas-Dias and Figueiredo, 2010) using different optimization strategies (discrete optimization by graph-cuts, gradient descent, Douglas-Rachford splitting or the alternating directions method of multipliers). Isolated strong scatterers are challenging for speckle reduction methods because they generally break the statistical assumptions made about the radar scene: repetition of similar patches within the search window (patch-based methods), sparse representation in the wavelets domain (wavelets-based methods), or piece-wise constant regions (TV minimization). An approach that would not specifically handle the strong scatterers separately is illustrated in Figure 6.2. We can see that it is necessary to identify and

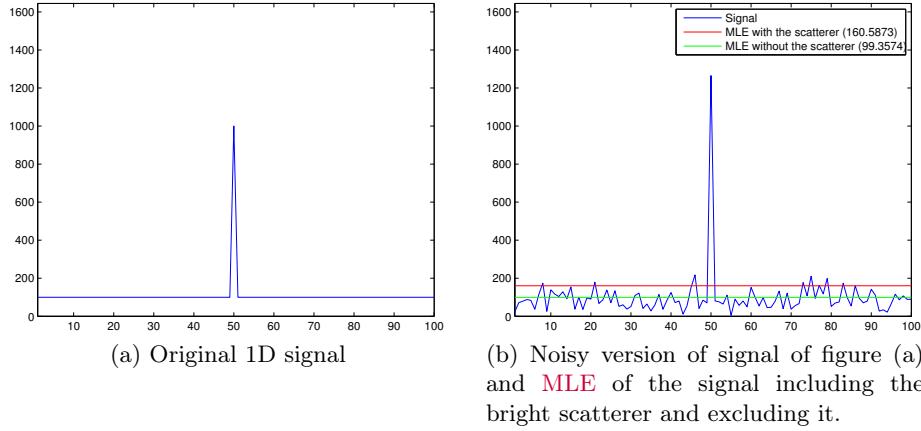


Figure 6.2 – Illustration of the effect of the presence of a strong scatterer for regularization. When the strong scatterer is included in the computation of the **MLE**, it over-estimates the true radiometry. By excluding the strong scatterer, the estimation is close to the true radiometry.

process these points separately to prevent from spreading these large values. Rather than building a statistical model of spatial configurations (edges, textures) in SAR images, many works directly model heterogeneity due both to the presence of strong scatterers and to geometrical features under a stationary assumption (Touzi, 2002; Vasile *et al.*, 2010). It is then possible to derive estimators robust to fluctuations of radiometry within the neighborhood (Pascal *et al.*, 2008).

On the one hand, statistical models that describe geometrical structures such as edges in SAR images generally fail to account for point-like strong scatterers; on the other hand, models of heterogeneous clutter based on a stationary assumption are lacking geometrical information. The work that is presented in this manuscript attempts to account for both the geometrical content and the strong scatterers in an explicit way, i.e., to recover the geometrical part and detect the strong scatterers.

Strong scatterers can be detected using likelihood ratio tests (Lopes *et al.*, 1992; Lopes *et al.*, 1993). These detectors compare the values in the center of a window with the rest of the window, considered as purely background. In dense urban areas, the presence of other point-like scatterers in the vicinity strongly deteriorates the performance and makes such local processing ill-suited.

We also aim at doing change detection between SAR images. During the last few years, the launch of constellations of sensors has allowed to obtain images with a short revisiting time. For instance, using Sentinel-1A and Sentinel-1B from ESA, one can have a new acquisition of a specific site every 6 days. This can be helpful for applications such as disaster monitoring, provided that adapted processing methods are developed. Such methods generally include change detection in a time series.

This problem was first tackled using difference between pairs of images (Singh, 1989) which has the drawback of not being well adapted to the multiplicative nature of the

speckle (see [chapter 2](#)). Comparison between SAR images is usually made using ratios such as in ([Moser and Serpico, 2006](#)). It can also be modeled by statistic tests such as hypothesis tests ([Lombardo and Oliver, 2001; Conradsen *et al.*, 2016](#)). However, these hypothesis tests require the estimation of the radiometry of the scene. This can be done using patch-based estimators such as in ([Su *et al.*, 2014a](#)).

In this work, we aim at performing jointly the detection of strong scatterers with the estimation of the background radiometry. We aim at providing models that explicitly take into account the strong scatterers and that can be applied to multi-temporal stack of images. When doing this joint detection and estimation, it is also possible to account for changes in the strong scatterers.

The idea of decomposing the SAR signal into several components is not new and led to very successful approaches in different SAR modalities. The reconstruction of a sparse distribution of scatterers has first been introduced in ([Cetin and Karl, 2001](#)), where a smoothing and a sparsity inducing penalty were applied to the same image. Sparse regularization has been further investigated under the concept of compressed sensing ([Ender, 2010; Potter *et al.*, 2010](#)) and successfully applied to tomographic SAR focusing ([Zhu and Bamler, 2010](#)). Beyond the recovery of point-like sources, decomposition under a dictionary of SAR signatures has been considered in ([Varshney *et al.*, 2008](#)). Detection of permanent scatterers and possibly of distributed scatterers is at the heart of permanent scatterers processing in multi-pass interferometry ([Ferretti *et al.*, 2011; Fornaro *et al.*, 2009](#)). Decompositions into different scattering behaviors is also central to the analysis of polarimetric images, see for example ([Yamaguchi *et al.*, 2011](#)). Other decomposition schemes have been considered, in particular to analyze the anisotropic behavior of scatterers from sub-aperture decomposition.

The approach followed in this chapter is in the spirit of image decomposition methods studied in the field of image processing ([Aujol *et al.*, 2006; Yin *et al.*, 2005; Elad *et al.*, 2005; Gilles, 2009](#)). These works consider the decomposition of a given image into several components, one typically containing geometrical information (i.e., edges), another capturing the textural information (oscillating components), and the remaining component with the noise. These decompositions are obtained by enforcing the sum of all components to match the initial image while each component gets penalized by a different regularization term promoting a different property (piece-wise constant image, oscillating image).

While the data we will use and the applications are different than those of [Part II](#), we will see that the developed method also relies on [MRF](#). Therefore, we will use an optimization technique based on graphcuts which is an extension of the one used in [chapter 4](#), and the techniques that are developed in this chapter could also be of benefit to the methods developed for SWOT: we did not tackle the case of multi-temporal

SWOT images, but it could be done using the formulation presented in this chapter. Also the non-optimal optimization method which will be presented in [subsection 6.3.2](#) could also be used to process large SWOT images.

This work can be seen as an extension of the decomposition model presented in ([Denis et al., 2010](#)). In the following, we present the general framework for our decomposition models in [subsection 6.2.1](#). The problems are first formulated as detection problems in [subsection 6.2.2](#) then as estimation problems in [subsection 6.2.3](#). We show how the proposed models can be exactly optimized in [subsection 6.3.1](#) and how the computational cost of the minimization can be reduced in [subsection 6.3.2](#). Finally, we include a discussion about the setting of the different parameters in [section 6.4](#). Presentation of the results for the targeted applications is done in [chapter 7](#).

6.2 Decomposition models

6.2.1 A general framework

Modeling urban SAR images as a sum of components Our goal is to explain a multi-temporal stack of T images \mathbf{v} where each image contains N pixels. We consider the following model:

$$\forall t \in \{1, \dots, T\}, \forall i \in \{1, \dots, N\}, v_{t,i} = u_{t,i} \cdot \xi_{t,i} \quad (6.1)$$

$$= (b_{t,i} + s_{t,i}) \cdot \xi_{t,i}, \quad (6.2)$$

where \mathbf{u} is the underlying radiometry of the scene (i.e. the speckle-free version of \mathbf{v}) that we seek to retrieve), \mathbf{b} is a background component and \mathbf{s} is a scatterer component. The general idea of the proposed methods is to find these two components jointly instead of following the more classical approach, where we would seek to retrieve \mathbf{u} directly. In the [MAP](#) framework, estimating \mathbf{u} directly would be expressed in terms of an energy to minimize:

$$\mathcal{E}(\mathbf{u}) = \text{DT}(\mathbf{u}, \mathbf{v}) + R(\mathbf{u}), \quad (6.3)$$

where $\text{DT}(\mathbf{u}, \mathbf{v})$ is a term penalizing a model \mathbf{u} that is not likely considering the observation \mathbf{v} (we will take the negative log-likelihood of the Rayleigh distribution in the case of amplitude SAR data) whereas $R(\mathbf{u})$ encodes prior knowledge on properties that we expect in the model. We have seen in [section 6.1](#), that classical priors (such as [TV](#)) fail because of the presence of strong scatterers. Using the decomposition presented in [Equation 6.2](#) and by making the assumption that both components are independent gives the following type of energy to minimize:

$$\mathcal{E}(\mathbf{u}) = \text{DT}(\mathbf{b} + \mathbf{s}, \mathbf{v}) + \beta_{\text{BG}} R(\mathbf{b}) + \beta_{\text{S}} R(\mathbf{s}), \quad (6.4)$$

where β_{BG} and β_{S} are weights to balance the terms. By doing so, we are able to put different priors on the background \mathbf{b} and on the scatterers \mathbf{s} components. In the following, we present three different models using different priors. In the following, the

prior on the background $R(\mathbf{b})$ is noted $\psi(\mathbf{b})$ and the prior on the scatterer $R(\mathbf{s})$ is noted $\lambda(\mathbf{s})$.

Different priors for different applications We have seen in [section 6.1](#) that a widely-used prior for image regularization is that the resulting image should have a low total variation. This model makes the assumption that the image can be described by constant regions with sharp edges. Whereas this model is not well suited when bright scatterers are present in the image, it can be efficiently applied to our background component as the bright scatterers are not present in it. Since the proposed models aim at processing multi-temporal stacks of images, we present a simple extension of the spatial **TV** penalization to the temporal domain. This extension makes the assumption that changes in time are also rare (which holds in the case of SAR images, thanks to the radiometric stability property):

$$\text{TV}_{3D}^\alpha(\mathbf{b}) = \sum_{t=1}^T \sum_{i \sim j} |b_{t,i} - b_{t,j}| + \alpha \sum_{t=1}^{T-1} \sum_{i=1}^N |b_{t+1,i} - b_{t,i}|. \quad (6.5)$$

This expression makes the hypothesis that the data are perfectly calibrated and registered. Using this formulation makes it possible to set the spatial and the temporal regularization independently via the parameter α .

A strong scatterer can be defined as a point which radiometry is an order of magnitude higher than its surrounding. From this definition, it follows that the bright scatterer component should be sparse. The straightforward way to model this assumption is to consider the pseudo-norm L0:

$$\|\mathbf{s}\|_0 = \sum_{t=1}^T \sum_{i=1}^N \bar{\delta}(s_{t,i}), \quad (6.6)$$

where $\bar{\delta}$ is the operator defined in [Equation 5.17](#). However when using the L0 pseudo-norm as a prior, the minimization of the energy defined in [Equation 6.3](#) is combinatorial (discontinuous and non-convex) which can lead to a computationally heavy optimization process. Typically, the L0 pseudo-norm is relaxed to a L1 norm (this is known as a convex relaxation):

$$\|\mathbf{s}\|_1 = \sum_{t=1}^T \sum_{i=1}^N |s_{t,i}|. \quad (6.7)$$

However, we will see in [section 6.3](#) that when using **TV** as the prior on the background, a model which uses the pseudo-norm L0 for the regularization of the scatterer component can be optimized efficiently. In [section 7.1](#) we present a comparison of L0 and L1 in the framework of scatterer detection. We show that even though L0 is often relaxed to L1, using L0 gives better results.

Using different functions for ψ and λ , we derive three models:

- **TVL0 Regularization (TVR)**: in this model, we use $\psi = \text{TV}_{3D}^\alpha$ with α a finite

number (in our experiment, it will be set to 1) and $\lambda = \|\cdot\|_0$. It is well adapted for applications in regularization or strong scatterers detection.

- **TVL0 One Background (TV1BG):** this model uses $\psi = \text{TV}_{3D}^\infty$ and $\lambda = \|\cdot\|_0$. By using an infinite weight on the temporal regularization, no variations are allowed for one pixel in the time direction, i.e. $\forall i, b_{1,i} = b_{2,i} = \dots = b_{T,i}$. It is based on the assumption that most changes are reflected by changes in strong scatterers and that one background can be sufficient to represent the time series. We will see that doing so greatly reduces the memory need for the model optimization. This model is well suited to scatterers detection.
- **TVL0 One Change (TV1C):** In this model, we also seek only one background (so it uses $\psi = \text{TV}_{3D}^\infty$). It differs from TV1BG as it will not only try to obtain a sparse scatterers component but will also try to obtain a sparse number of temporal changes in it. For the problem to stay tractable, we put constraints on the values one site in the scatterers component can take over time:
 - it can be always equal to 0 (no strong scatterer);
 - it can have the same positive value at every date (no change);
 - it can be 0 until a date t_{app} then be a constant positive value (appearing scatterer);
 - it can be a constant positive value until t_{disp} then be equal to 0 (disappearing scatterer).

We can then define the change map $\mathbf{c}(\mathbf{s})$ as:

$$c_i(\mathbf{s}) = \begin{cases} 0 & \text{if } \exists r > 0, \forall t \in [1, T], s_{t,i} = 0 \text{ or } s_{t,i} = r; \\ 1 & \text{if } \exists r > 0, \forall t \in [1, t_{\text{app}}[, s_{t,i} = 0 \\ & \text{and } \forall t \in [t_{\text{app}}, T], s_{t,i} = r; \\ 2 & \text{if } \exists r > 0, \forall t \in [1, t_{\text{disp}}[, s_{t,i} = r \\ & \text{and } \forall t \in [t_{\text{disp}}, T], s_{t,i} = 0, \end{cases} \quad (6.8)$$

with $r > 0$. To enforce that both the scatterer component and the number of changes in it should be sparse, we use $\lambda(\mathbf{s}) = \|\mathbf{s}\|_0 + \frac{\beta_C}{\beta_S} \|\mathbf{c}(\mathbf{s})\|_0$, where β_C is a weight to balance both terms. By definition, this model is particularly well suited for change detection.

6.2.2 Interpretation of the decomposition as a detection problem

We first tackle the problem of the detection of a strong scatterer at a given pixel (and a change at a given site in the case of the TV1C model) for a given value of the background. We will see that under the L0 prior, this problem is formalized as a [Likelihood Ratio Test \(LRT\)](#). We first start with the simpler case of detecting a strong scatterer for the

models TVR and TV1BG before formalizing the joint detection of strong scatterers and their changes for TV1C.

Detection of a strong scatterer Given a background value $b_{t,i}$, the strong scatterer detection can be expressed as a choice between two hypotheses: the absence (\mathcal{H}_0) or the presence (\mathcal{H}_1) of a strong scatterer in addition to the background scatterers:

$$\begin{cases} \mathcal{H}_0 & : s_{t,i} = 0 ; \\ \mathcal{H}_1 & : s_{t,i} > 0 . \end{cases} \quad (6.9)$$

Under the assumptions of uncorrelated speckle and point-like scatterers, this hypothesis test can be performed as a LRT:

$$\log \frac{p(v_{t,i}|u_{t,i} = b_{t,i} + s_{t,i})}{p(v_{t,i}|u_{t,i} = b_{t,i})} \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \beta_S . \quad (6.10)$$

In Equation 6.10, we know the value of the observation, and the background component's value is an input to our problem. However, the value of the strong scatterer component $s_{t,i}$ is unknown. Replacing it by the value given by the MLE leads to the Generalized Likelihood Ratio Test (GLRT):

$$\max_{s_{t,i} > 0} \log p(v_{t,i}|u_{t,i} = b_{t,i} + s_{t,i}) \stackrel{\mathcal{H}_1}{\underset{\mathcal{H}_0}{\gtrless}} \beta_S + \log p(v_{t,i}|u_{t,i} = b_{t,i}) \quad (6.11)$$

Detection of the strong scatterers and their changes We now examine the case of the detection of a strong scatterer and its potential change at a pixel i in the time series given the constant background value b_i . We recall that we make the (strong) hypothesis that at most one change is present in the time series. This problem can be expressed as a decision between four hypotheses:

- \mathcal{H}_0 : no strong scatterer is present at pixel i .
- \mathcal{H}_{1a} : a strong scatterer is present at pixel i at all dates.
- $\mathcal{H}_{1b}^{\text{app}}$: a strong scatterer appears at a certain date.
- $\mathcal{H}_{1b}^{\text{disp}}$: a strong scatterer disappears at a certain date.

We can also formalize intermediate hypotheses:

- \mathcal{H}_1 : a strong scatterer is present at pixel i (at all or only a fraction of the dates).
- \mathcal{H}_{1b} : a strong scatterer is present but not at all dates.

These hypotheses allow us to express the problem as a hierarchical hypothesis test as presented in Figure 6.3. These hypotheses tests are also performed using a LRT. First

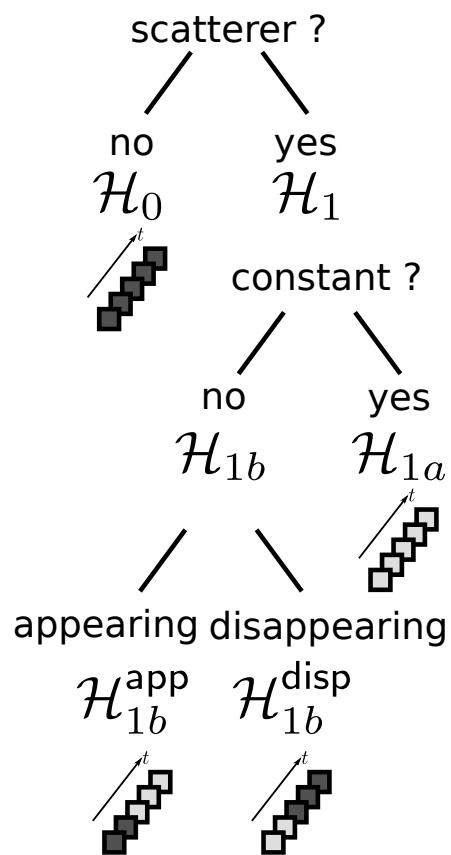


Figure 6.3 – Hierarchical hypothesis test for the detection of strong scatterer and its potential change

the negative log-likelihoods corresponding to the different hypotheses are defined:

$$\mathcal{L}_0(b_i) = - \sum_t \log p(v_{t,i} | u_{t,i} = b_i) \quad (6.12)$$

$$\mathcal{L}_{1a}(b_i, r) = - \sum_t \log p(v_{t,i} | u_{t,i} = b_i + r) \quad (6.13)$$

$$\mathcal{L}_{1b}^{\text{app}}(b_i, r, t_{\text{app}}) = - \sum_{t=1}^{t_{\text{app}}-1} \log p(v_{t,i} | u_{t,i} = b_i) - \sum_{t=t_{\text{app}}}^T \log p(v_{t,i} | u_{t,i} = b_i + r) \quad (6.14)$$

$$\mathcal{L}_{1b}^{\text{disp}}(b_i, r, t_{\text{disp}}) = - \sum_{t=1}^{t_{\text{disp}}-1} \log p(v_{t,i} | u_{t,i} = b_i + r) - \sum_{t=t_{\text{disp}}}^T \log p(v_{t,i} | u_{t,i} = b_i), \quad (6.15)$$

where r is a positive value (to be estimated and considered known at this point) corresponding to the constant radiometry of the strong scatterer when it is present. We first solve the case under \mathcal{H}_1 : we know that we have a strong scatterer and we must decide whether it is constant (\mathcal{H}_{1a}) or changing (\mathcal{H}_{1b}) over time. This is expressed as a LRT:

$$\log \frac{p(\{v_{t,i}\} | \mathcal{H}_{1b})}{p(\{v_{t,i}\} | \mathcal{H}_{1a})} \stackrel{\mathcal{H}_{1b}}{\gtrless} \beta_C \quad (6.16)$$

$$\mathcal{L}_{1a}(b_i, r) - \mathcal{L}_{1b}(b_i, r) \stackrel{\mathcal{H}_{1b}}{\gtrless} \beta_C \quad (6.17)$$

The value of b_i and r are known but solving [Equation 6.17](#) requires knowledge of $\mathcal{L}_{1b}(b_i, r)$. It is therefore replaced by its MLE $\widehat{\mathcal{L}}_{1b}(b_i, r)$:

$$\widehat{\mathcal{L}}_{1b}(b_i, r) = \max \left(\max_{t_{\text{app}}} \mathcal{L}_{1b}^{\text{app}}(b_i, r, t_{\text{app}}), \max_{t_{\text{disp}}} \mathcal{L}_{1b}^{\text{disp}}(b_i, r, t_{\text{disp}}) \right). \quad (6.18)$$

The computation [Equation 6.18](#) is done using an exhaustive search; both $\mathcal{L}_{1b}^{\text{app}}(b_i, r, t_{\text{app}})$ and $\mathcal{L}_{1b}^{\text{disp}}(b_i, r, t_{\text{disp}})$ are tested for each possible date. Using the MLE for the estimation of \mathcal{L}_{1b} gives the following GLRT:

$$\mathcal{L}_{1a}(b_i, r) \stackrel{\mathcal{H}_{1b}}{\gtrless}_{\mathcal{H}_{1a}} \widehat{\mathcal{L}}_{1b}(b_i, r) + \beta_C. \quad (6.19)$$

We also need to solve the upper level of the hierarchical test, which is the strong scatterer detection (choice between \mathcal{H}_0 and \mathcal{H}_1). This is almost equivalent to [Equation 6.10](#), except that this time the test is performed once for each pixel (and not at each date):

$$\log \frac{p(\{v_{t,i}\} | \mathcal{H}_1)}{p(\{v_{t,i}\} | \mathcal{H}_0)} \stackrel{\mathcal{H}_1}{\gtrless}_{\mathcal{H}_0} \beta_S \quad (6.20)$$

$$\mathcal{L}_0(b_i) - \mathcal{L}_1(b_i, r) \stackrel{\mathcal{H}_1}{\gtrless}_{\mathcal{H}_0} \beta_S \quad (6.21)$$

In order to solve [Equation 6.21](#), we need to estimate the value r . Once again, the MLE is used, giving the following GLRT:

$$\mathcal{L}_0(b_i) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \min_{r>0} \mathcal{L}_1(b_i, r) + \beta_S, \quad (6.22)$$

where $\mathcal{L}_1(b_i, r)$ is given by:

$$\mathcal{L}_1(b_i, r) = \min \left(\mathcal{L}_{1a}(b_i, r), \widehat{\mathcal{L}}_{1b}(b_i, r) + \beta_C \right). \quad (6.23)$$

6.2.3 Reformulation as an estimation problem

The detection problems from the previous part can be formalized as estimation problems. In this framework, [Equation 6.11](#) is equivalent to:

$$\widehat{s}_{t,i}(b_{t,i}) = \arg \min_{s_{t,i} \geq 0} (-\log(p(v_{t,i}|u_{t,i} = b_{t,i} + s_{t,i})) + \beta_S \bar{\delta}(s_{t,i})). \quad (6.24)$$

This equation can be simply expressed for the entire stack of images:

$$\widehat{\mathbf{s}}(\mathbf{b}) = \arg \min_{\mathbf{s} \geq 0} \left(- \sum_i \sum_t \log(p(v_{t,i}|u_{t,i} = b_{t,i} + s_{t,i})) + \beta_S \|\mathbf{s}\|_0 \right). \quad (6.25)$$

In a similar way, [Equation 6.23](#) can be written as:

$$\begin{aligned} \widehat{\mathbf{s}}(\mathbf{b}) = \arg & \min_{\substack{\mathbf{r} \geq 0 \\ \mathbf{t}_c \in [1, T]^N}} \left(- \sum_i \sum_t \log(p(v_{t,i}|b_i + s_{t,i})) + \beta_S \|\mathbf{d}\|_0 + \beta_C \|\mathbf{c}(\mathbf{s})\|_0 \right) \\ & \text{s.t. } \forall i, \forall t, \quad s_{t,i} \geq 0 \\ & \forall i, \forall t, \quad \delta(c_i(\mathbf{s}))(s_{t,i} - r_i) = 0 \\ & \forall i, \forall t < t_{ci}, \quad \delta(c_i(\mathbf{s}) - 1)s_{t,i} = 0 \\ & \forall i, \forall t \geq t_{ci}, \quad \delta(c_i(\mathbf{s}) - 1)(s_{t,i} - r_i) = 0 \\ & \forall i, \forall t < t_{ci}, \quad \delta(c_i(\mathbf{s}) - 2)(s_{t,i}) - r_i = 0 \\ & \forall i, \forall t \geq t_{ci}, \quad \delta(c_i(\mathbf{s}) - 2)s_{t,i} = 0 \end{aligned} \quad (6.26)$$

where \mathbf{d} is an image such that:

$$d_i = \begin{cases} 1 & \text{iff } \exists t \text{ s.t. } s_{t,i} > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (6.27)$$

The constraint on [Equation 6.26](#) are required to ensure that in the "no change" case the radiometry of the strong scatterer is constant, in the case of an appearing strong scatterer it is 0 before the appearance and a constant value after, and that in the case of a disappearing one, it is constant before the disappearance date and a constant value after.

We can not directly solve [Equation 6.25](#) and [Equation 6.26](#) as they are function of the value of the background \mathbf{b} which is unknown. We have seen in [section 6.1](#) that in order to be able to properly estimate the background, strong scatterers must be excluded. Thus we propose to jointly estimate both components. By adding the desired prior on the background, it gives the equation that was sought in the first place:

- In the case of the TVR model:

$$\begin{aligned} (\hat{\mathbf{b}}, \hat{\mathbf{s}}) = \arg \min_{\substack{\mathbf{b} \geq 0 \\ \mathbf{s} \geq 0}} & - \sum_i \sum_t \log(p(v_{t,i} | b_{t,i} + s_{t,i})) + \beta_{BG} \text{TV}_{3D}^{\alpha}(\mathbf{b}) + \beta_S \|\mathbf{s}\|_0 \end{aligned} \quad (6.28)$$

- TV1BG model:

$$\begin{aligned} (\hat{\mathbf{b}}, \hat{\mathbf{s}}) = \arg \min_{\substack{\mathbf{b} \geq 0 \\ \mathbf{s} \geq 0}} & - \sum_i \sum_t \log(p(v_{t,i} | b_{t,i} + s_{t,i})) + \beta_{BG} \text{TV}_{3D}^{\infty}(\mathbf{b}) + \beta_S \|\mathbf{s}\|_0 \end{aligned} \quad (6.29)$$

- TV1C model:

$$\begin{aligned} (\hat{\mathbf{b}}, \hat{\mathbf{s}}) = \arg \min_{\substack{\mathbf{b} \geq 0 \\ \mathbf{r} \geq 0 \\ \mathbf{t}_c \in [1, T]^N}} & - \sum_i \sum_t \log(p(v_{t,i} | b_{t,i} + s_{t,i})) + \beta_{BG} \text{TV}_{3D}^{\infty}(\mathbf{b}) \\ & + \beta_S \|\mathbf{d}\|_0 + \beta_C \|\mathbf{c}(\mathbf{s})\|_0 \end{aligned} \quad (6.30)$$

s.t. $\forall i, \forall t, \quad s_{t,i} \geq 0$
 $\forall i, \forall t, \quad \delta(c_i(\mathbf{s}))(s_{t,i} - r_i) = 0$
 $\forall i, \forall t < t_{ci}, \quad \delta(c_i(\mathbf{s}) - 1)s_{t,i} = 0$
 $\forall i, \forall t \geq t_{ci}, \quad \delta(c_i(\mathbf{s}) - 1)(s_{t,i} - r_i) = 0$
 $\forall i, \forall t < t_{ci}, \quad \delta(c_i(\mathbf{s}) - 2)(s_{t,i}) - r_i = 0$
 $\forall i, \forall t \geq t_{ci}, \quad \delta(c_i(\mathbf{s}) - 2)s_{t,i} = 0$

Note that even though in [Equation 6.29](#) and in [Equation 6.30](#) the background component is a function of time, it is not allowed to vary along time (from the TV_{3D}^{∞} prior), and the stack of background components can be replaced by a single image.

Summary: Decomposition models

We presented three models adapted for the processing of multi-temporal series of urban SAR images. The main idea is to separate the image as a sum of two components: the background and the strong scatterers.

We can then apply meaningful priors on each of the components. By selecting different priors, we propose three models:

- TVR, for the regularization of time series.
- TV1BG, for an efficient strong scatterers detection and a mean representation of the background.
- TV1C, for change detection.

We have first presented the models as detection problems. We have then presented a reformulation as estimation problems which will allow us to apply an exact optimization method in the next section.

6.3 Optimization

6.3.1 Exact optimization using graph-cuts

Graph construction Because of the L0 term, previously defined energies ([Equation 6.28](#), [Equation 6.30](#) and [Equation 6.30](#)) are non-convex. As seen in [section 6.1](#), this typically prevents the model from being efficiently optimized and is generally solved by replacing the L0 pseudo-norm by an L1 norm. However, when an optimization problem involves (possibly non-convex) separable terms and convex pairwise terms, it can be restated to a minimum cut / maximum flow search on an S-T graph ([Ishikawa, 2003](#)). In ([Pock et al., 2008](#)) a continuous variational counterpart was proposed to solve the same type of problems. With the particular choice of the total variation as a prior for the background component, the defined energies can be solved using this method.

To use this optimization method, we need the energy to be formulated as:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \sum_i \sum_t f_0(x_{t,i}) + \sum_{(i,t) \sim (j,t')} f_1(x_{i,t}, x_{j,t'}) , \quad (6.31)$$

where f_0 is a function for the separable terms, and f_1 for the pairwise terms (and f_1 should be convex). This function only optimizes a single variable. However, we have expressed in [Equation 6.25](#) and in [Equation 6.26](#) the optimal value for a strong scatterer given a background. Hence, we can easily define the single-variable problem as in [Equation 6.31](#). For models TVR and TV1BG and data in single-look amplitude

distribution (hence following a Rayleigh law), it yields:

$$\mathbf{x} = \mathbf{b} \quad (6.32)$$

$$f_0(x_{t,i}) = 2 \log(x_{t,i} + s_{t,i}(x_{t,i})) + \frac{v_{t,i}^2}{(x_{t,i} + s_{t,i}(x_{t,i}))^2} + \beta_S \bar{\delta}(s_{t,i}) \quad (6.33)$$

$$f_1(x_{t,i}, x_{t,j}) = \beta_{BG} |x_{t,i} - x_{t,j}| \quad (6.34)$$

$$f_1(x_{t,i}, x_{t',i}) = \alpha \beta_{BG} |x_{t,i} - x_{t',i}| \quad (6.35)$$

Our problem is now stated following [Equation 6.31](#) so it can be solved using Ishikawa's method ([Ishikawa, 2003](#)). As in [chapter 4](#), the main idea behind it is to build a graph such that an S-T cut in this graph corresponds to a solution of our problem, and that the energy of solution equals the value of the cut. It is then possible to use one of the algorithms developed in the field of Computer Science to compute a minimum-cut (e.g. ([Boykov and Kolmogorov, 2004](#)) which is particularly adapted to grid-structured graphs). Details about the graph construction can be found in [subsection 3.2.2](#). In the particular case of this model, vertical links going from one quantization level to another for the same site are not only set to the data-term value, but to $f_0(x)$, which includes the energy for the scatterer component.

In the case of TV1R, the graph has one node for each pixel at each date and each quantization level. For large series of images, this can raise problems in memory load. One advantage of models TV1BG and TV1C is that, as they do not allow variations in the background, [Equation 6.35](#) can be replaced with $f_1(x_{t,i}, x_{t',i}) = \infty$ (as $\alpha = \infty$). Therefore, the graph can be constructed with only one node for each pixel at each possible quantization level. In this latter case, we need to sum all the contributions for the possibly changing strong scatterers; in the case of TV1BG:

$$f_0(x_i) = \sum_t 2 \log(x_i + s_{t,i}(x_i)) + \frac{v_{t,i}^2}{(x_i + s_{t,i}(x_i))^2} + \beta_S \bar{\delta}(s_{t,i}) \quad (6.36)$$

In the case of TV1C, we also need to take into account the prior on the changes:

$$f_0(x_i) = \sum_t 2 \log(x_i + s_{t,i}(x_i)) + \frac{v_{t,i}^2}{(x_i + s_{t,i}(x_i))^2} + \beta_S \bar{\delta}(s_{t,i}) + \beta_C \bar{\delta}(c_{t,i}(s(x_i))) \quad (6.37)$$

Scatterer component computation In order to actually build the graph, we need to compute $s_{t,i}(b_{t,i})$. We start by the expression of the GLRT of [Equation 6.11](#). At a given pixel, under Rayleigh distribution, we have the following expression of the [MLE](#) for the model $u_{t,i} = b_{t,i} + s_{t,i}$ given an observation $v_{t,i}$:

$$\widehat{u}_{t,i} = \arg \max_{u_{t,i}} \log(p(v_{t,i} | u_{t,i})) \quad (6.38)$$

$$= \arg \max_{u_{t,i}} \log \left(\frac{2v_{t,i}}{u_{t,i}^2} \right) - \frac{v_{t,i}^2}{u_{t,i}^2} \quad (6.39)$$

Then:

$$\frac{\partial}{\partial u_{t,i}} \log(p(v_{t,i}|u_{t,i})) = 0 \iff \frac{\partial}{\partial u_{t,i}} \left(\log \left(\frac{2v_{t,i}}{u_{t,i}^2} \right) - \frac{v_{t,i}^2}{u_{t,i}^2} \right) = 0 \quad (6.40)$$

$$\iff -\frac{2}{u_{t,i}} + \frac{2v_{t,i}^2}{u_{t,i}^3} = 0 \quad (6.41)$$

$$\implies u_{t,i} = v_{t,i} \quad (6.42)$$

Hence the **MLE** of $u_{t,i}$ is $\widehat{u_{t,i}} = v_{t,i}$, then $\widehat{s_{t,i}} = v_{t,i} - b_{t,i}$, giving:

$$\max_{s_{t,i} > 0} \log p(v_{t,i}|u_{t,i} = b_{t,i} + s_{t,i}) = \log \left(\frac{2}{v_{t,i}} \right) - 1 \quad (6.43)$$

and [Equation 6.11](#) can be written as:

$$\max_{s_{t,i} > 0} \log p(v_{t,i}|u_{t,i} = b_{t,i} + s_{t,i}) \stackrel{\mathcal{H}_1}{\gtrless}_{\mathcal{H}_0} \beta_S + \log p(v_{t,i}|u_{t,i} = b_{t,i}) \quad (6.44)$$

$$\iff \frac{v_{t,i}^2}{b_{t,i}^2} - \log \left(\frac{v_{t,i}^2}{b_{t,i}^2} \right) \stackrel{\mathcal{H}_1}{\gtrless}_{\mathcal{H}_0} \beta_S + 1 \quad (6.45)$$

This computation holds for the choice of the Rayleigh distribution. In [section 2.3](#), we saw that Rice distribution is more adapted when a scatterer is dominant in the resolution cell (which is the case when we consider a non-null strong scatterer component). However, the **MLE** of $u_{t,i}$ can not be derived in closed-form (as in the case of [Equation 6.42](#) for the Rayleigh distribution) and must be numerically evaluated. Another advantage of using the Rayleigh distribution is that the **GLRT** only depends on the ratio between the observation and the background $\frac{v_{t,i}}{b_{t,i}}$ which allows us to set β_S using a desired constant false alarm rate (see [section 6.4](#)). For the scatterer detection, in the cases of TVR and TV1BG, using [Equation 6.45](#) and the equivalence between [Equation 6.11](#) and [Equation 6.24](#) gives:

$$\widehat{s_{t,i}}(b_{t,i}) = \begin{cases} v_{t,i} - b_{t,i} & \text{if } v_{t,i} > b_{t,i} \\ & \text{and } \frac{v_{t,i}^2}{b_{t,i}^2} - \log \left(\frac{v_{t,i}^2}{b_{t,i}^2} \right) > \beta_S + 1 \\ 0 & \text{otherwise.} \end{cases} \quad (6.46)$$

For the model TV1C, the same formula applies over the hierarchical hypothesis test shown in [Figure 6.3](#). The only difference in the resolution of both models is the computation of the optimal scatterer value. In the case of TV1C, the strong scatterer at one site only takes at most one single strictly positive value. In the case of a constant strong scatterer which is present at all date (\mathcal{H}_{1a}), the **MLE** of $\widehat{u_i}$ is:

$$\widehat{u_i} = \arg \max_{u_i} \frac{\sum_{t=1}^T \log(p(v_{t,i}|u_i))}{T}. \quad (6.47)$$

Then:

$$\frac{\partial}{\partial u_i} \frac{\sum_{t=1}^T \log(p(v_{t,i}|u_i))}{T} = 0 \iff \frac{\sum_{t=1}^T \left(-\frac{2}{u_i} + \frac{2v_{t,i}^2}{u_i^3} \right)}{T} = 0 \quad (6.48)$$

$$\iff -\frac{2}{u_i} + \frac{1}{T \cdot u_i^3} \cdot 2 \sum_{t=1}^T v_{t,i}^2 = 0 \quad (6.49)$$

$$\iff u_i^2 = \frac{\sum_{t=1}^T v_{t,i}^2}{T} \quad (6.50)$$

$$\implies u_i = \sqrt{\frac{\sum_{t=1}^T v_{t,i}^2}{T}}. \quad (6.51)$$

In words, the **MLE** of the model given amplitude observations at different dates is the square root of the average of the squares of the observations. It is equivalent to converting the data to intensity before doing the average. Note that when $T = 1$, the **MLE** is equivalent to [Equation 6.42](#). We can then express the optimal value for r in each hypothesis:

- for \mathcal{H}_{1a} :

$$\hat{r}(b_i) = \sqrt{\frac{\sum_{t=1}^T v_{t,i}^2}{T}} - b_i, \quad (6.52)$$

- for an appearing scatterer at date t_{app} ($\mathcal{H}_{1b}^{\text{app}}$):

$$\hat{r}(b_i) = \sqrt{\frac{\sum_{t=t_{\text{app}}}^T v_{t,i}^2}{T - t_{\text{app}} + 1}} - b_i, \quad (6.53)$$

- for a disappearing scatterer at date t_{disp} ($\mathcal{H}_{1b}^{\text{disp}}$):

$$\hat{r}(b_i) = \sqrt{\frac{\sum_{t=1}^{t_{\text{disp}}-1} v_{t,i}^2}{t_{\text{disp}} - 1}} - b_i, \quad (6.54)$$

It is then possible to apply the hypothesis tests defined in [Equation 6.18](#), [Equation 6.19](#) and in [Equation 6.23](#).

Computational and memory complexity The graph construction requires a large amount of memory that limits its application to small regions of interest and/or short time series. The number of vertices is proportional to the number of images in the time series (in the case of TV1R), the number of pixels of each image, and the number of quantization levels. The number of edges is proportional to the number of

nodes (about 8 times the number of nodes). We used the graph-cuts implementation described in (Boykov and Kolmogorov, 2004). In this graph implementation, each vertex requires 48 bytes of storage, and each edge requires 32 bytes. Hence, the memory footprint of the graph representation limits the size of the series of images that can be processed. For example, a series of 20 images of size 300×400 with a quantization into 50 levels requires 33.7GiB (or 36.2GB) of RAM for the graph construction when using TVR. However, when $\psi = \text{TV}_{3D}^{\infty}$ (which is the case in TV1BG and TV1C), the problem only involves one background. Solving the same problem with this regularization only requires 1.35GiB, simplifying the exact optimization of the proposed model.

Regarding the computational cost, the worst-case complexity of the minimum cut is $O(EV^2|C|)$ where $|C|$ is the value of the minimum cut, E is the number of edges and V the number of vertices. In practice, the experimental complexity scales almost linearly with the number of nodes (Boykov and Kolmogorov, 2004). On a computer with an Intel® Xeon(R) CPU E5-1620 with 16Gb of RAM, the algorithm takes 52,04s to compute a decomposition on 2 images of 300×400 pixels with 50 levels of quantization. Note that we do not fully benefit from the power of the processor as the used algorithm is single-core. Algorithms computing the grid-cut in parallel have also been proposed (see (Liu and Sun, 2010)) and more compact memory representations that exploit the regularity of the graph have been introduced in (Jamriska *et al.*, 2012).

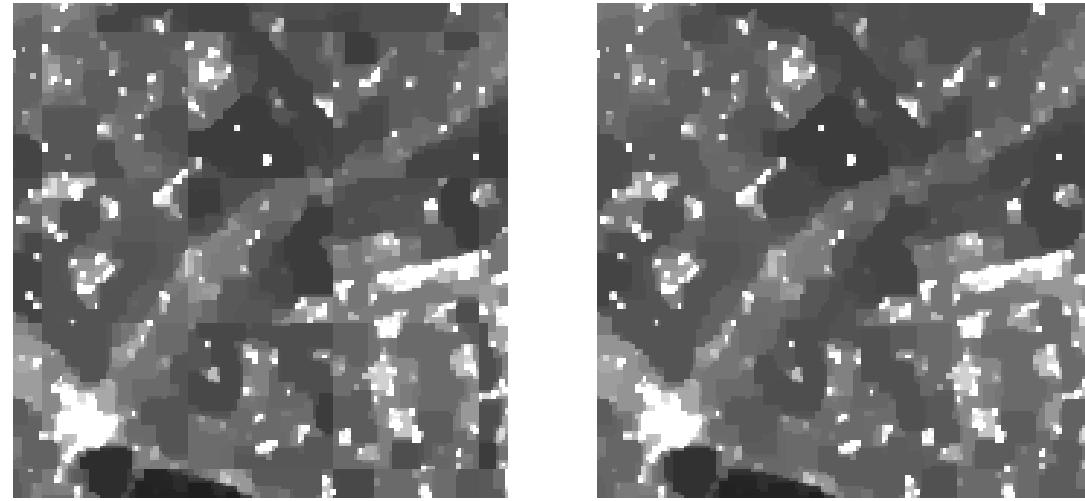
To reduce the computational and memory complexity, a subset of all quantization levels can be considered at a time to get an approximate solution, see (Denis *et al.*, 2009; Shabou *et al.*, 2011). We describe in subsection 6.3.2 another approach based on block processing to further reduce memory requirements.

6.3.2 Efficient optimization

To apply the method to large images and/or long time series, it is necessary to develop a method for limited-memory graph-cuts optimization.

Memory usage of the graph-cut method is proportional to the number of pixels in the series. The required memory can thus be reduced by computing the optimization locally, on spatio-temporal blocks extracted from the time series. Even though the models involve only second order cliques, limiting the direct interactions to the immediate neighborhood, the maximum a posteriori estimate involves long-range correlations, i.e., during the optimization, regularization effects are propagated over long distances. Simple division of the image into smaller blocks thus results into visible *block artifacts*, as can be observed in figure 6.4a. These *block artifacts* are due to the lack of context: a constant area in the global optimum that gets divided into two regions during the block-processing is represented by two different (constant) values with an artificial discontinuity between the blocks.

To introduce context in the window F of interest, it is necessary to perform the optimization on a bigger window C containing all the objects partially presents in F .



(a) Solution obtained by local optimization (computation window of size 50^2 leading to a block effect)

(b) Exact solution obtained using the presented algorithm for a computation window of size 150^2

Figure 6.4 – Crops of the background component of the first image of the Saint-Gervais set obtained with the presented method (filling window of size 50^2).

The proposed method works as follows:

1. To process a given window F (named filling window in the following), extract a larger computation window C such that $F \subset C$.
2. Perform the optimization using the graph-cuts method described in [subsection 6.3.1](#) on the spatio-temporal window C .
3. Keep the results of the decomposition only in the (smaller) filling window F .
4. Slide the filling window F and repeat from step 1 until the whole image has been covered by the filling window.

This process is illustrated in [Figure 6.5](#). We show in [Figure 6.6](#) the root mean squared error obtained by the block-processing approach for different sizes of the computation window (the size of the block F is kept constant and equal to 50×50 pixels for this single-date image). The required memory grows quadratically with the spatial window size. Note that as we increase the size of the computation window, the algorithm is slower. This is explained by the fact that the number of cuts stays the same for any size of the computation window (as we do not change the size of F) while the size of the graph increases (which also increases the complexity of finding the minimal cut, see [subsection 6.3.1](#)). As a comparison, running the optimal algorithm on the same dataset takes 505s. It can be observed that on this image, the exact solution is obtained for computation windows C larger or equal to 150×150 pixels (which requires only 8.6% of the amount of needed memory to process the whole image at once). When the computation window F is strictly smaller than the image size, the solution computed on

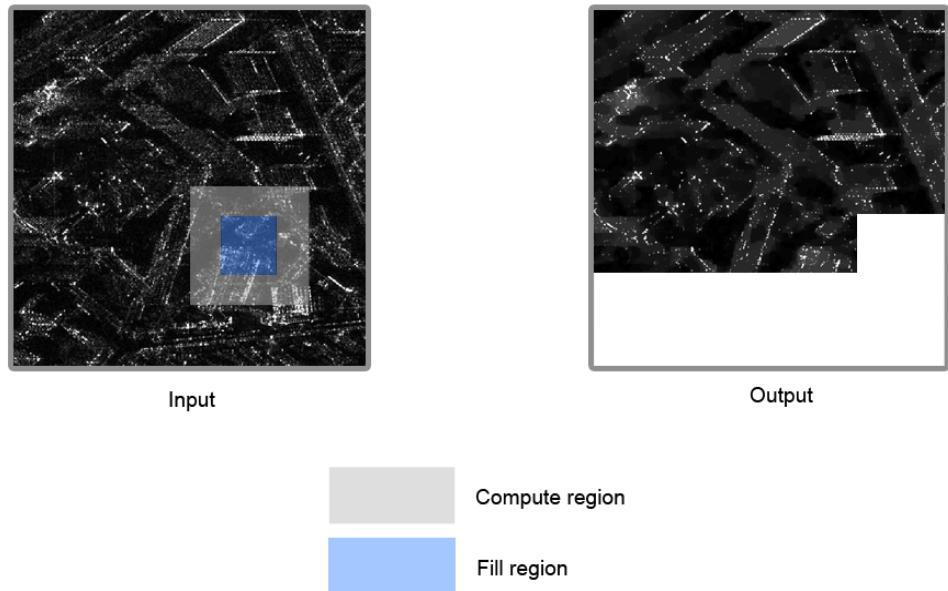


Figure 6.5 – Illustration of the proposed method for a memory efficient optimization of the decomposition models. A sliding window is used to compute the output on part of the image. The computations are done in a larger window (the compute region) in order to include context and avoid the *block effect*. Animation available at www.sylvainlobry.com/phd.

the block F is not guaranteed to match the solution obtained from the whole image, i.e., it is only approximate. Therefore, this algorithm makes it possible to select a trade-off between memory/time complexity and quality of the result.

Summary: Optimization of decomposition models

Thanks to the choice of the TV prior for the background, we can use an exact optimization method based on graph-cuts for our model provided that a quantization of the possible values for the background component is performed. To construct the graph, the priors on the strong scatterers are integrated in the function representing the clique of order one. In practice, at each pixel and for each possible value for the background, the optimal value for the strong scatterer components are computed.

6.4 Parameters computation

One of the drawbacks of the proposed methods is that they require the user to tune a number of parameters. Each method requires to provide:

- The number of looks (number) L : we have shown the equations for the simple case of the Rayleigh distribution (where the number of looks is set to 1), however

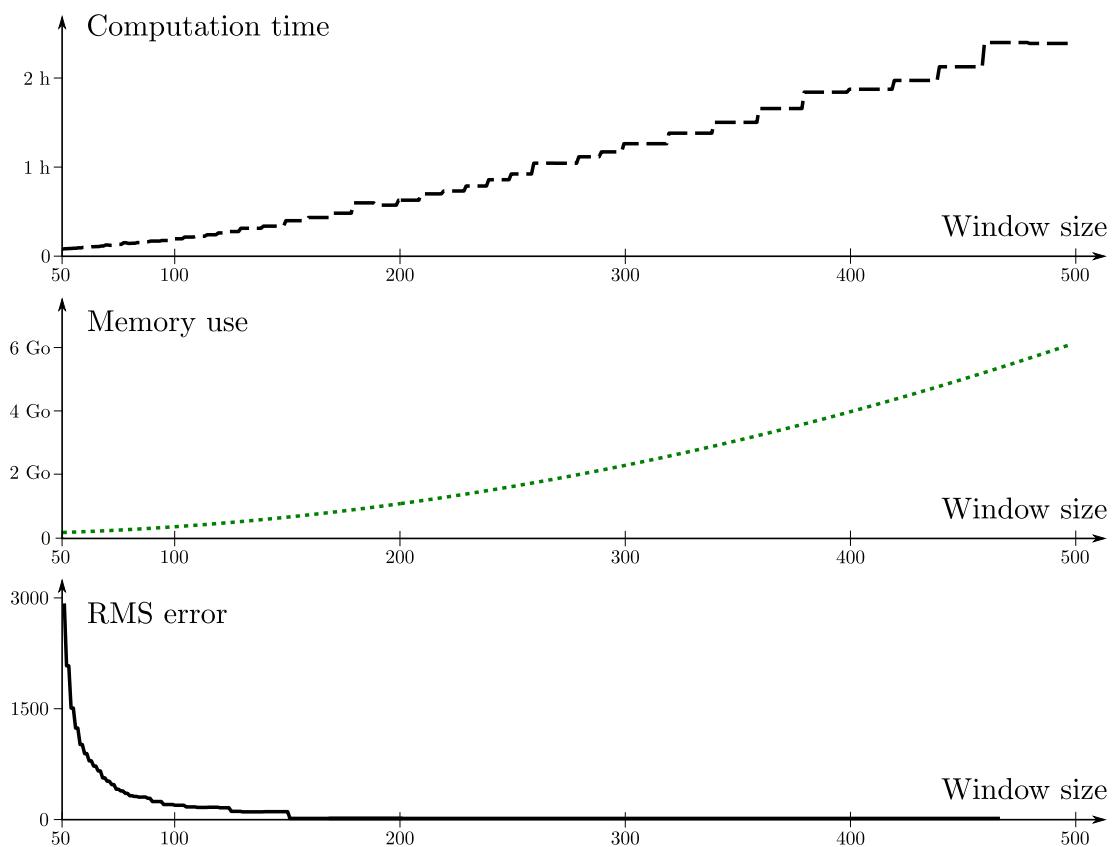


Figure 6.6 – Trade-off between computation time, memory use (top two graphs) and accuracy (bottom graph): evolution as a function of the size of the spatial computation window C , for the first image of the Saint-Gervais data-set.

it is easily extended to the Rayleigh-Nakagami distribution. In this case, it is another parameter in the likelihood. This number of look is usually known for each image. When it is not known, the **Equivalent Number Of Looks (ENL)** can be easily estimated by computing the coefficient of variation in an homogeneous region ([Jean-Marie Nicolas, 2012](#)).

- The quantization levels (array of numbers): this parameter is required in order to apply the exact optimization scheme presented in [section 6.3](#). However it is hard to tune it manually. We present a method to tune it based on two meaningful parameters.
- The weight of the prior on the background component (number) β_{BG} : this parameter is heavily application dependent. Depending on the targeted application, it will change. However, we provide in this section a formula to obtain a first approximation of its value.
- The weight of the prior on the strong scatterers component (number) β_S : this parameter depends on one's definition of a strong scatterer. It can, however, be easily set as a function of the contrast between a strong scatterer and the background.

In addition, when using the change detection model TV1C, we have to set the weight put on the number of detected changes. Once again, this is heavily application dependent.

The quantization levels: while the quantization levels are hard to tune manually, they have a direct influence on the complexity (both computational and in memory, see [section 6.3](#)) and on the quality of the result. Instead of requiring an array of numbers, we propose to tune this parameter based on:

- the number of desired levels (n_{quantif});
- the desired coverage of the histogram of the image for the background component (p).

Concerning the number of desired levels, it will usually be tuned based on the available memory (i.e. the larger number the better). A notable exception is when the targeted application is a classification based on the radiometries, where the quantization levels would be easily tuned in order to match the radiometries of the different classes. The second parameter is the portion p (between 0 and 1) of the histogram that is to be represented by the background (a typical value is $p = 0.95$). This can be manually set by looking at the histogram of the image. The quantization levels are then obtained using the quantiles on the p parts of the total histogram of the image.

The weight of the prior on the strong scatterers component (β_S): from [Equation 6.45](#) it is easy to derive a desired contrast $c = \frac{v_{t,i}}{b_{t,i}}$ to the weight that have to be

put on the prior of the strong scatterers component:

$$\beta_S = c^2 - 2 \log(c) - 1 \quad (6.55)$$

The weight of the prior on the background (β_{BG}): this prior highly depends on the application. In some desired cases, a noisy background (obtained with a low β_{BG}) is sufficient, whereas for images interpretation purposes, it can be useful to have a highly regularized background. However it also depends on other variables such as the number of looks and the quantization. Following the idea presented for the tuning of β in [subsection 4.4.4](#), an average level of regularization can be obtained using the following formula for model TVR:

$$\beta_{BG} = \frac{L}{\delta_{\text{quantif}}}, \quad (6.56)$$

where δ_{quantif} is the average distance between two levels of quantization. In the case of models TV1BG and TV1C, the number of images represented by the background must be taken into account:

$$\beta_{BG} = \frac{L \cdot T}{\delta_{\text{quantif}}}. \quad (6.57)$$

The weight for the prior on the number of changes (β_C): This weight can hardly be set automatically. TV1C suffers from the fact that the distribution of \mathcal{L}_{1b} is not constant through the dates: for instance, it is more likely to detect an appearing strong scatterer at the beginning of the series than it is at the end. Along the fact that it is a problem for the detection of changes, it is also a problem for a meaningful setting of β_C . The variation of the empirical distributions of $\mathcal{L}_{1b}^{\text{app}}$ (defined in [Equation 6.14](#)) is shown in [Figure 6.7](#). Compensating the variations of the distribution under the change cases is a problem that needs to be addressed in the future.

6.5 Conclusion

We have presented three different models using decompositions into background and scatterer components in order to process multi-temporal series of SAR images. The models use different priors on the components, allowing to target different applications such as regularization, strong scatterer detection and change detection that will be covered in chapter [chapter 7](#). The proposed formulation has several advantages:

- it makes it possible to put meaningful priors on each components;
- it can be optimized exactly up to a chosen discretization;
- there is no other requirement on the prior for the strong scatterers component other than it has to be separable. Then, many different applications can be considered.

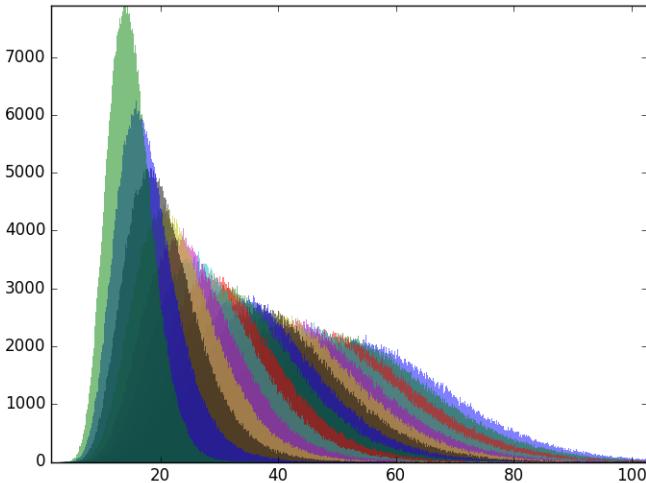


Figure 6.7 – Empirical distribution of $\mathcal{L}_{1b}^{\text{app}}$ under \mathcal{H}_0 (no strong scatterer). The radiometry of the background is set to $\mu = 20$, the number of looks is $L = 1$ and there are $T = 15$ dates in the series.

Given the fact that new models could be developed based on different priors for the strong scatterers, future work includes enhancements based on the strong scatterers priors for the proposed models. For instance, while we only allow one change for the TV1C model, a more realistic model would allow cyclic changes. It has been discarded for now because of the computational complexity it would raise, but it could be formulated in the proposed framework. Another improvement would be to enforce a **Constant False Alarm Rate (CFAR)** of the change detection in TV1C. As seen in Figure 6.7, the number of detected changes is not constant through the time.

Future work could also be dedicated to the optimization of the models, using the variational method from (Pock *et al.*, 2008) which would allow to have a continuous (as opposed to the discretization currently needed) optimization.

From the methods presented in this chapter, we can list contributions that could be transferred in the context of the water/land classification (see Part II):

- The temporal links defined with TV_{3D}^α could be used with the Ising model that was used for the classification part of chapter 4 when multi-temporal series will be available.
- The efficient optimization scheme presented in subsection 6.3.2 to reduce the memory usage can be used on any MRF defined on the pixel level, and therefore could be used for the classification and estimation MRF defined in chapter 4.

Chapter 7

Applications

This chapter is dedicated to applications of the three models presented in [chapter 6](#). The targeted applications are the following:

- Strong scatterer detection ([section 7.1](#)): the detection of strong scatterers, either as features for the detection/classification of man-made structure or simply to take them into account in further processing steps. It is generally done when working on SAR images of urban area. Classically used methods such as ([Lopes et al., 1993](#)) estimate the background radiometry on a window and then compute a [GLRT](#). The results show the benefits of the joint detection and estimation compared to such models. Also, this section compares the results using the L0 pseudo-norm against using an L1 norm, as it is usually done for convex relaxation.
- Image regularization ([section 7.2](#)): regularizing SAR images can be helpful in interpretation tasks. We compare the results from the TVR and TV1BG models to two widely used models used for regularization: a simple [TV](#) regularization and a temporal multi-looking. This comparison outlines the difference in the results between these models, and shows how the proposed models can be useful for regularization.
- Change detection ([section 7.3](#)): as we have seen in [section 6.1](#), change detection is one of the primary application of [SAR](#) images. In [section 7.3](#), we compare our method to widely-used models in this particular task. We show that our method gives interesting results, especially when we do not want spatial regularization on the changes.

7.1 Strong scatterers detection

In this section, we compare the scatterers detection of the TVR/TV1BG model to the method proposed by ([Lopes et al., 1993](#)). This method computes the ratio of the radiometry of a given pixel and a cross centered on the pixel. This ratio is then thresholded to decide if the pixel is a strong scatterer. Also, we address the benefits of using the L0 pseudo-norm versus a convex relaxation with the L1 norm when doing

strong scatterers detection. The comparison is presented based on results computed on a synthetic image. In subsection 6.2.1, we defined a strong scatterer as a point which radiometry is an order of magnitude higher than its surrounding. However, it still not clear what can be considered "an order of magnitude higher" given the complexity of the background that surrounds a target. In this respect, a ground truth would be necessarily biased toward a method.

The combinatorial optimization method described in section 6.3 is applicable to the L0 pseudo-norm and to the L1 norm, and provides in both cases the global optimum. It is thus possible to compare the performance of the two formulations: the L0 penalty derived from the GLRT and the L1 norm widely used to obtain a convex relaxation of the minimization problem, see for example (Tropp and others, 2006).

We either consider the model TVR/TV1BG (in our experiments, only one background is considered, making both models exactly the same), or a modified version of this model where the L0 pseudo-norm is replaced by the L1 norm. When considering the modified minimization problem involving the L1 norm, the optimal $s_{t,i}$ is no longer given by a simple expression as in Equation 6.46. Optimal values $s_{t,i}$ are either zero, or if strictly positive, they must cancel the first partial derivative of the sum of the log-likelihood and the L1 norm, given below:

$$\begin{aligned} \frac{\partial}{\partial s_{t,i}} \left[2 \log(s_{t,i} + b_{t,i}) + \frac{v_{t,i}^2}{(s_{t,i} + b_{t,i})^2} + \beta_S s_{t,i} \right] = \\ - \frac{2 v_{t,i}^2}{(s_{t,i} + b_{t,i})^3} + \frac{2}{s_{t,i} + b_{t,i}} + \beta_S. \end{aligned} \quad (7.1)$$

The only positive and real-valued root, as obtained by finding the roots of a third degree polynomial using a computer algebra system, is:

$$\widehat{s}_{t,i}(b_{t,i}) = t^{1/3} + \frac{4}{9\beta_S^2 t^{1/3}} - \frac{3\beta_S b_{t,i} + 2}{3\beta_S}, \quad (7.2)$$

with:

$$t = \frac{v_{t,i} \sqrt{27\beta_S^2 v_{t,i}^2 - 16}}{3^{3/2} \beta_S^2} + \frac{27\beta_S^2 v_{t,i}^2 - 8}{27\beta_S^3}. \quad (7.3)$$

We consider a numerical experiment to compare the detection performance of L0 and L1 formulations, for various contrasts between point-like scatterers and a piecewise constant background, see Figure 7.1c. The contrast between each scatterer and the background is given in Figure 7.1d. 100 noisy versions are then generated using a Rayleigh multiplicative model. The L0 and L1 models are applied to each of these images with a fixed value of β_{BG} (set to 0.05) and various β_S values to obtain the receiver operating characteristic (ROC) curves drawn on Figure 7.2. They are compared with the ROC curve obtained using the method based on a local estimation of the background proposed in (Lopes *et al.*, 1993). The L0 model outperforms both the L1

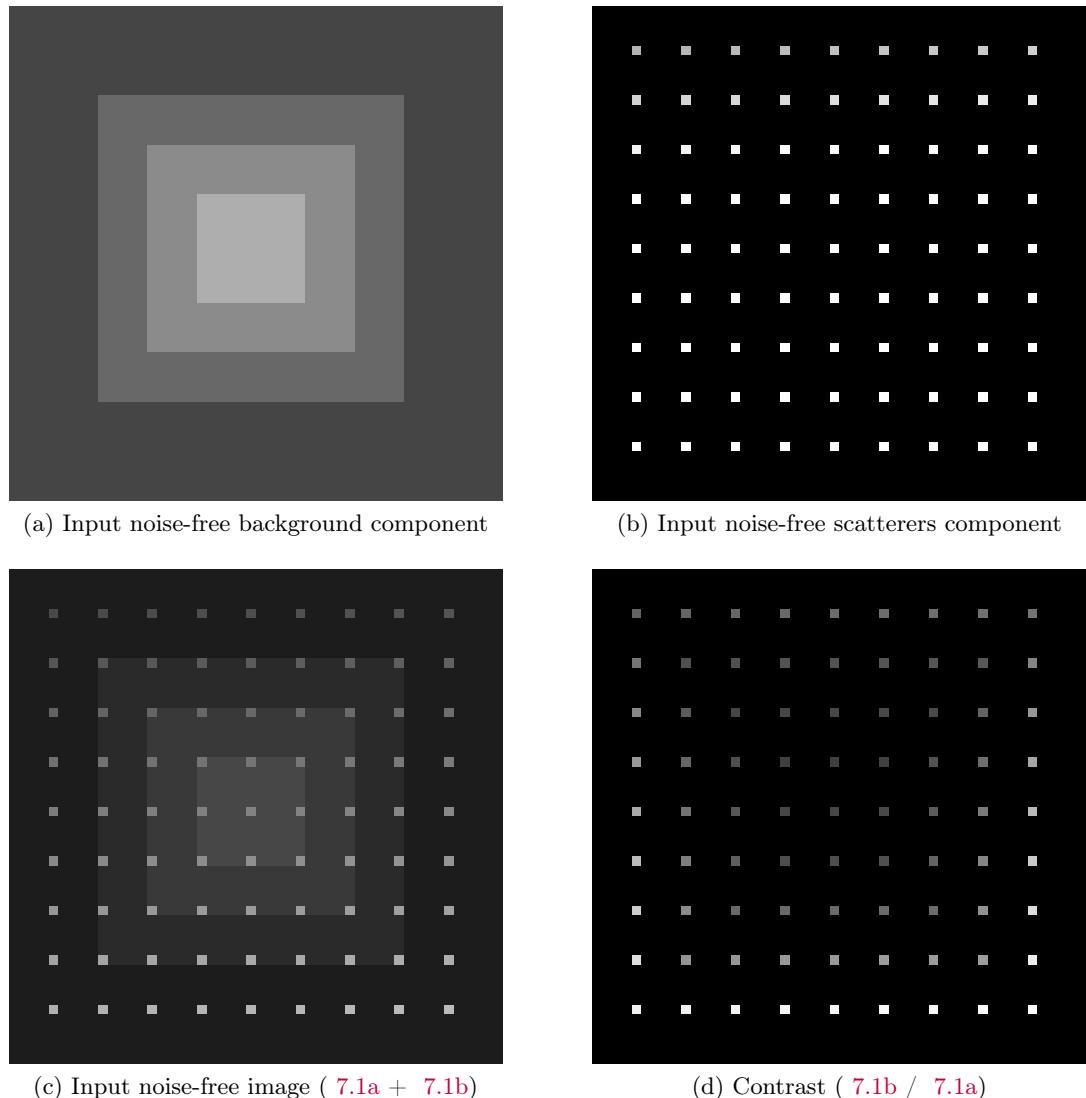


Figure 7.1 – Input image used to compare L0 and L1 models in the scatterer detection application. 100 noise realizations of 7.1c are generated and scores obtained in scatterers detection using both the L0 and L1 models are presented in Figure 7.4.

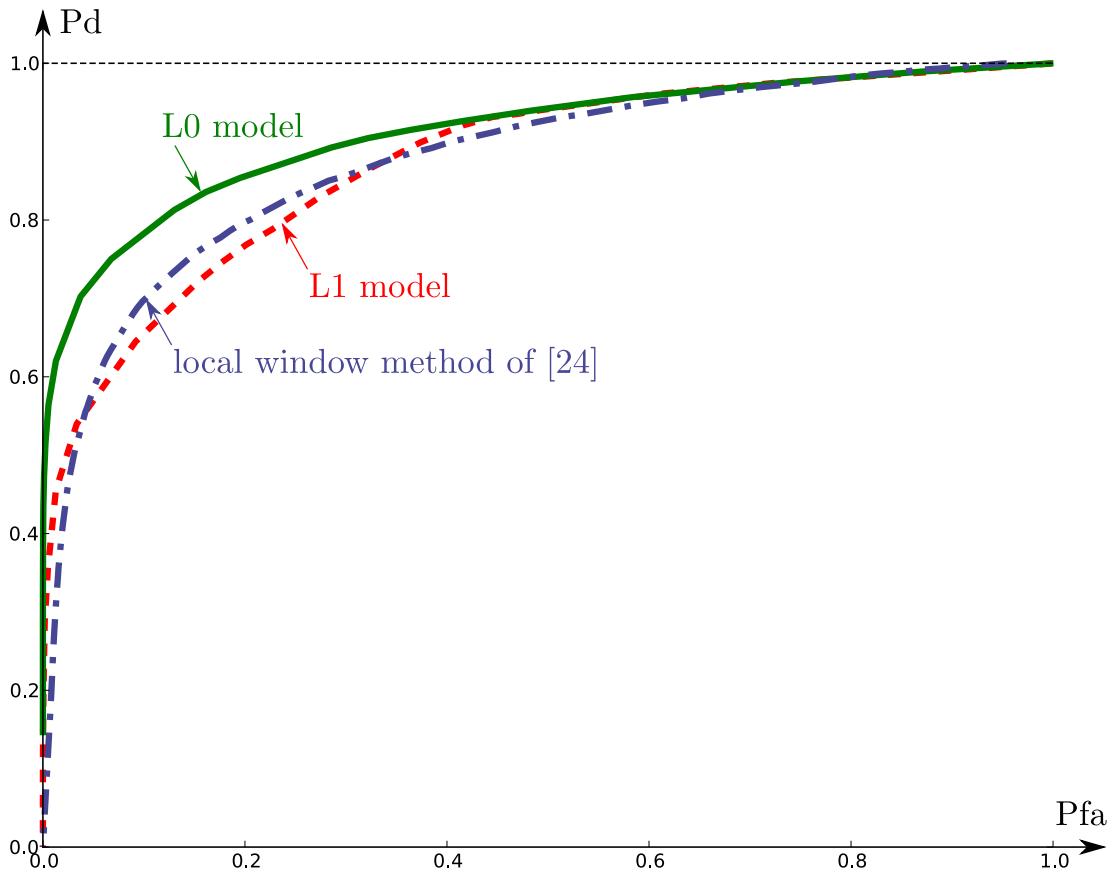


Figure 7.2 – Receiver operating curve (ROC) between L0 and L1 version of our model and a method based on local window analysis (Lopes *et al.*, 1993).

model and the detection based on local background estimation. We show in Figure 7.3 the scatterers detected by each model, for an identical proportion of correct detections. The L0 model gives a uniform performance (correct detections and false detections are well distributed in the image), while the L1 model fails to detect scatterers in the regions with higher background radiometry and gives more false detections in low-radiometry areas. This phenomenon is confirmed by representing the evolution of the Probability of False Alarm (PFA) and the Probability of Detection (PD) as a function of the background radiometry, Figure 7.2. The L0 term produces constant PFA and PD (note that, in the numerical simulation, when the background radiometry changes, the scatterer radiometry is adapted so as to keep a unitary ratio: $s_{t,i}/b_{t,i} = 1$). In contrast, the PFA and PD of the L1 model decrease when the background radiometry increases, which is consistent with the non-uniformity of detected scatterers observed in 7.3b.

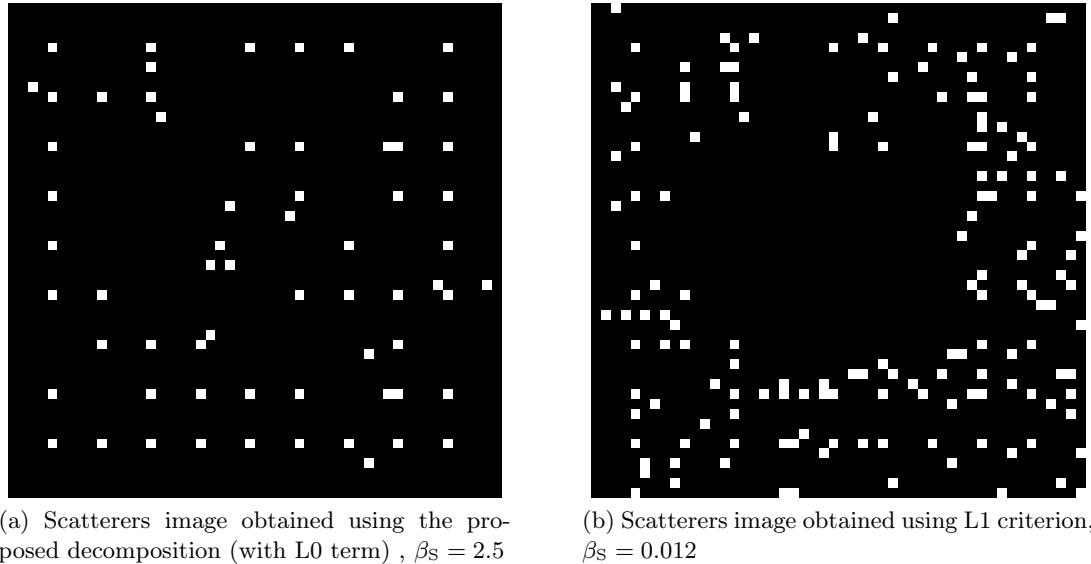


Figure 7.3 – 2 images of scatterers achieving the same probability of detection: left with an L0 penalty, right with an L1 penalty.

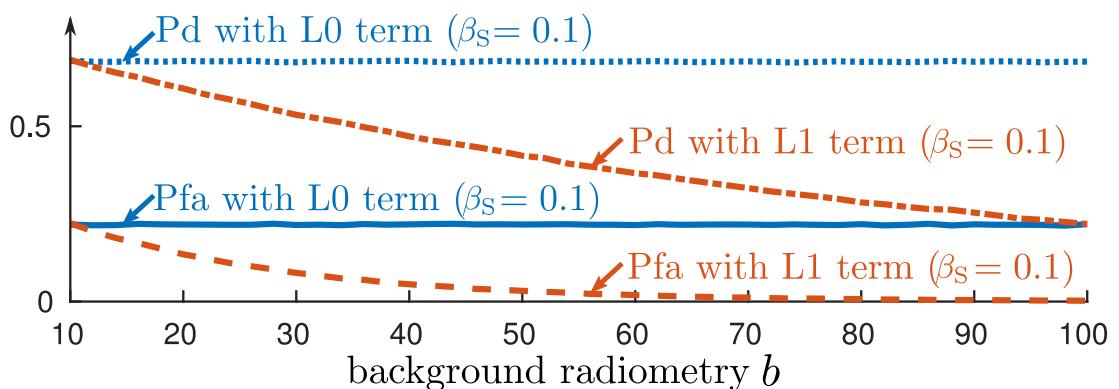


Figure 7.4 – With an L1 penalty, the probability of false alarm **PFA** and the probability of detection **PD** vary with the background radiometry.

Summary: Strong scatterers detection

The three proposed models allow for a strong scatterer detection based on the pseudo-norm L0. We compared the performances of one of our model (TV1BG) on a synthetic set of images to (Lopes *et al.*, 1993) and a L1 version of our model. Compared to (Lopes *et al.*, 1993), our model benefits from a better estimation of the background. The results presented also indicate that a convex relaxation to the L1 norm has strong incidences on the performances.

7.2 Image regularization

For image regularization, both models TVR and TV1BG can be used efficiently. We show the results on a dataset acquired by TerraSAR-X in spotlight mode (resolution of $1m \times 1m$). The multi-temporal stack is made of 56 images with dimensions of 700×700 pixels. Acquisition dates ranges from 11/22/2007 to 03/13/2012. The first and the last image of the multi-temporal stack are shown in Figure 7.5. In these images, we can expect a lot of variations at different scales:

- A lot of traffic can be present at the border of the river "La Seine".
- In the river, we can have variations due to the presence of boats. Also there are two train lines in the image (from the top of the image to the right, and from the right to the bottom).
- Two areas are in construction. On the upper-left part of the image, the "Maison de la Radio" was being renovated (from early 2009 until today), while a mall was first demolished then built on the bottom-right part of the image (from 2007 to 10/23/2013).

As using the exact optimization method described in subsection 6.3.1 for the TVR model (124 GiB/133 GB of memory are required) is hardly possible, we use for all computations of this section the suboptimal method presented in subsection 6.3.2. The parameters are as follow and are set using the methods described in section 6.4:

- $n_{\text{quantif}} = 20$
- $p = 0.95$
- $c = 3$ (which leads to $\beta_S = 5.8$).
- F (side of the filling window for the efficient optimization) = 100
- C (side of the computing window for the efficient optimization) = 150

The weight on the prior on the background β_{BG} is set automatically (using Equation 6.56 and Equation 6.57).

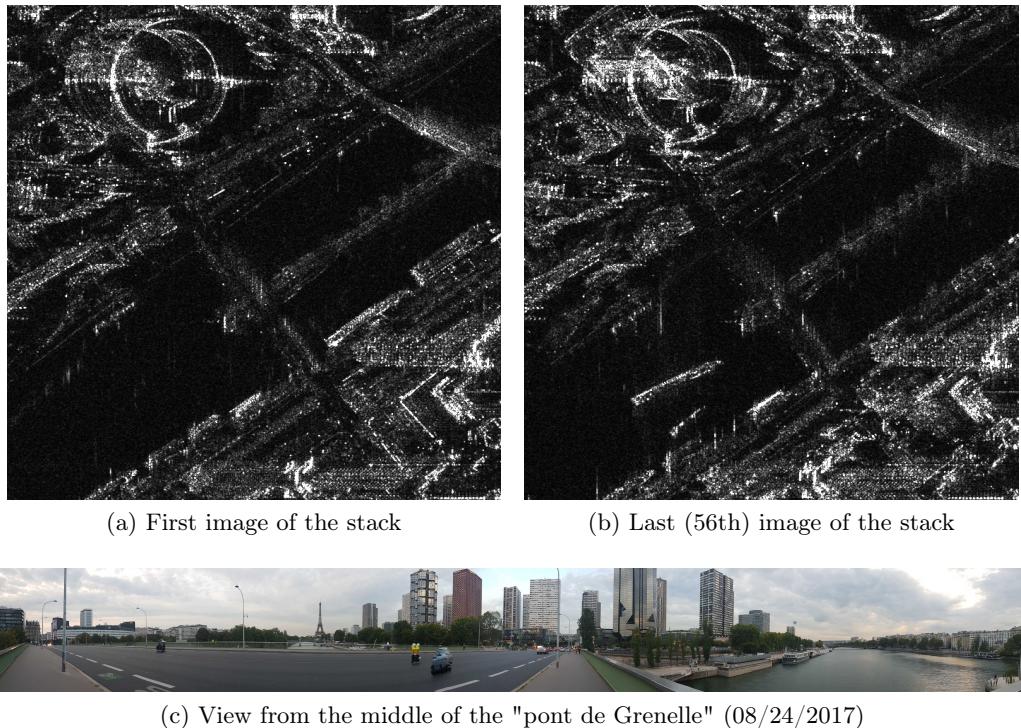


Figure 7.5 – First and last image of the Paris data. Acquired by TerraSAR-X in spotlight mode. Animation showing the data at each date is available at www.sylvainlobry.com. The 360° photography taken from the middle of the SAR images shows the "Maison de la Radio" on the left (which is on the top-left corner of the SAR images) and the mall in the middle (bottom-right of the SAR images).

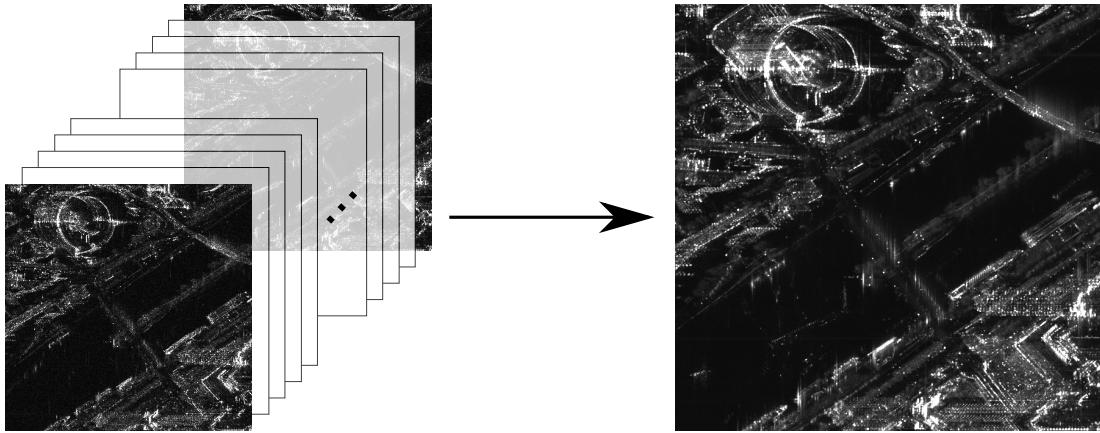


Figure 7.6 – Result of the temporal multi-looking applied on the full stack.

In this section, we compare the results obtained by TVR and TV1B to classical regularization methods: classical **TV** or temporal multi-looking. We recall that temporal multi-looking apply the **MLE** at each site taking all the pixels for the different dates. In amplitude, it gives:

$$u_i = \sqrt{\frac{1}{T} \sum_{t=1}^T v_i^2} \quad (7.4)$$

From [section 6.1](#), we know that these methods are not fully adapted to an application on multi-temporal series of SAR images: the assumption of a low **TV** is broken by the presence of numerous strong scatterers, while the assumption that there is no change between the different acquisitions that is made by the temporal multi-looking does not hold in highly populated urban area.

We show the results from the temporal multi-looking in [Figure 7.6](#) and the results obtained using the classical **TV** model in [Figure 7.7](#). The results obtained by our method are presented in [Figure 7.8](#) for model TVR and in [Figure 7.9](#) for model TV1BG. Finally, all the results are presented in [Figure 7.10](#) for an easier comparison.

We can observe in the result of the temporal multi-looking ([Figure 7.6](#)) that the image has almost no speckle remaining. However, in changing areas, we have many strong scatterers. Basically, when a strong scatterer is present at a site for at least one date, it will be present in the temporal multi-looking. This results in many strong scatterers for areas under construction, and a "shadow" effect on the river, where boats have been present at one date.

On the opposite, when looking at the results from the **TV** method, we can realize that no strong scatterer is present. This is easily explained: as they break the assumption of large homogeneous area, the model is not able to represent them. However, their presence affects the value of the homogeneous area. It is worth noticing that when lowering the regularization parameter, it is possible to recover the strong scatterers at the price of a poor localization. If we want to recover the right localization, the regularization parameter should be very close to 0, leading to very little regularization.

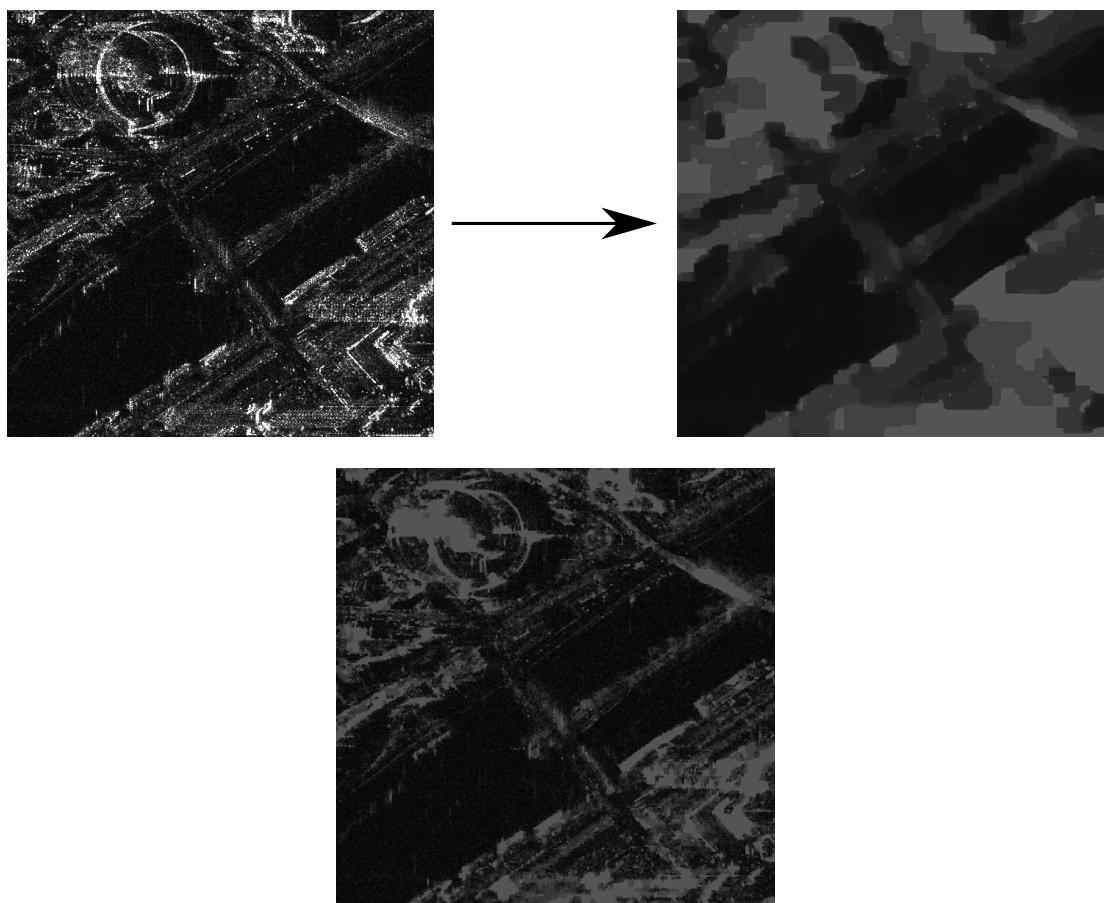


Figure 7.7 – Result of the classical **TV** regularization on the first image of the stack, with the automatically obtained regularization parameter (on top), or a lower (factor 10) regularization parameter (bottom)

The results obtained by the first of the proposed model (TVR) are presented in [Figure 7.8](#). We can see the effect of the two regularization terms: the background contains a few but large homogeneous area while the strong scatterer component is sparse. It is worth noticing that there is only few changes between the two pictured backgrounds, which is coherent with the assumption that the majority of changes are from the strong scatterer component. This assumption is exploited by the model TV1BG (see [Figure 7.9](#)). While the results are visually nearly as good, close examination reveals some incoherency, for instance in the bottom-right corner of the image for the first image of the series.

Summary: Image regularization

We presented results obtained using the proposed TVR and TV1BG models and compared them to two widely used regularization methods: [TV](#) and a simple temporal multi-looking. As expected the proposed models perform better than these methods when strong scatterers are present (which breaks the hypothesis made by [TV](#)) and potentially moving (which breaks the assumption made by the temporal multi-looking).

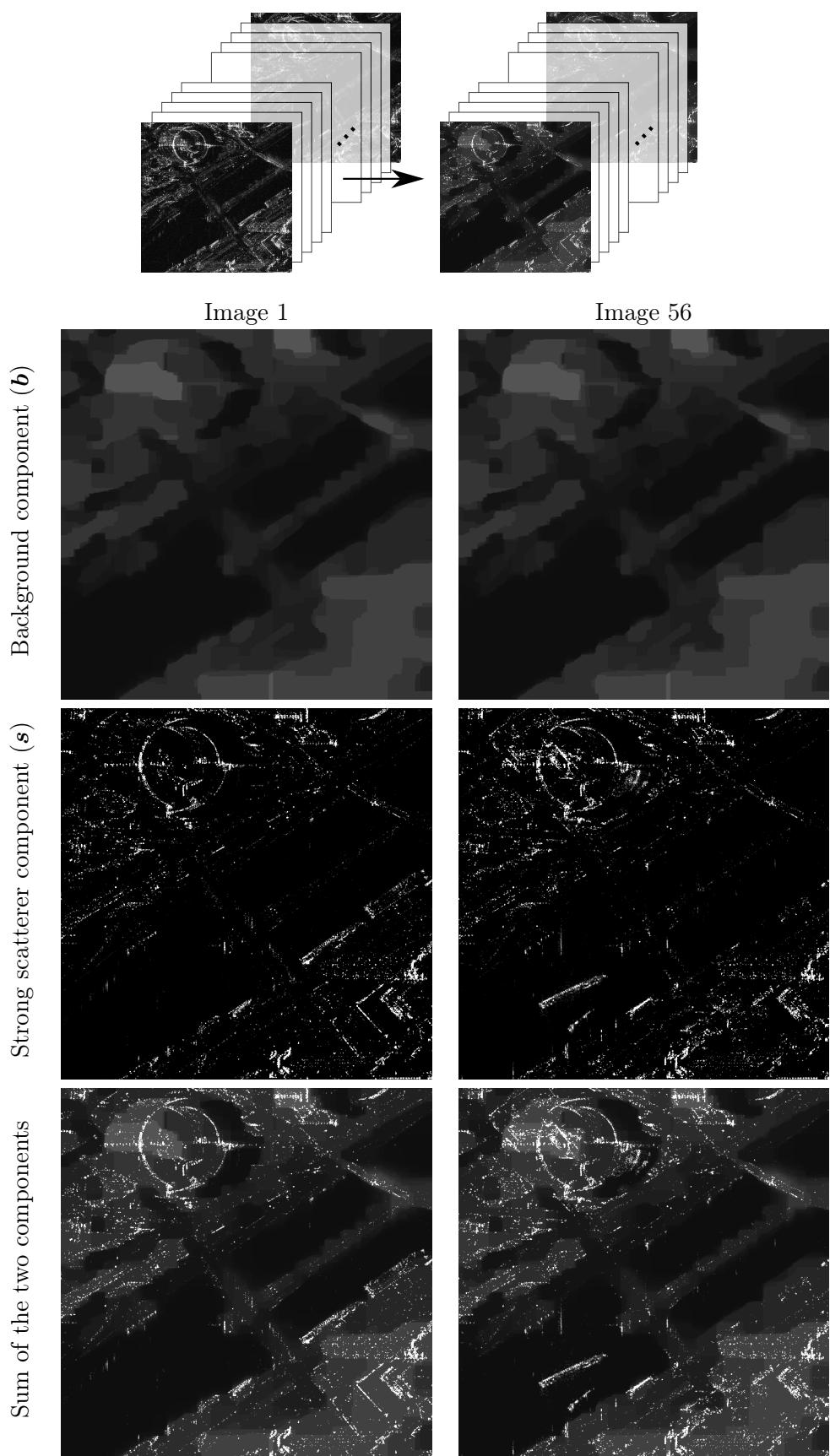


Figure 7.8 – Result of the TVR model on the full stack. An animation showing all the results is available at www.sylvainlobry.com/phd.

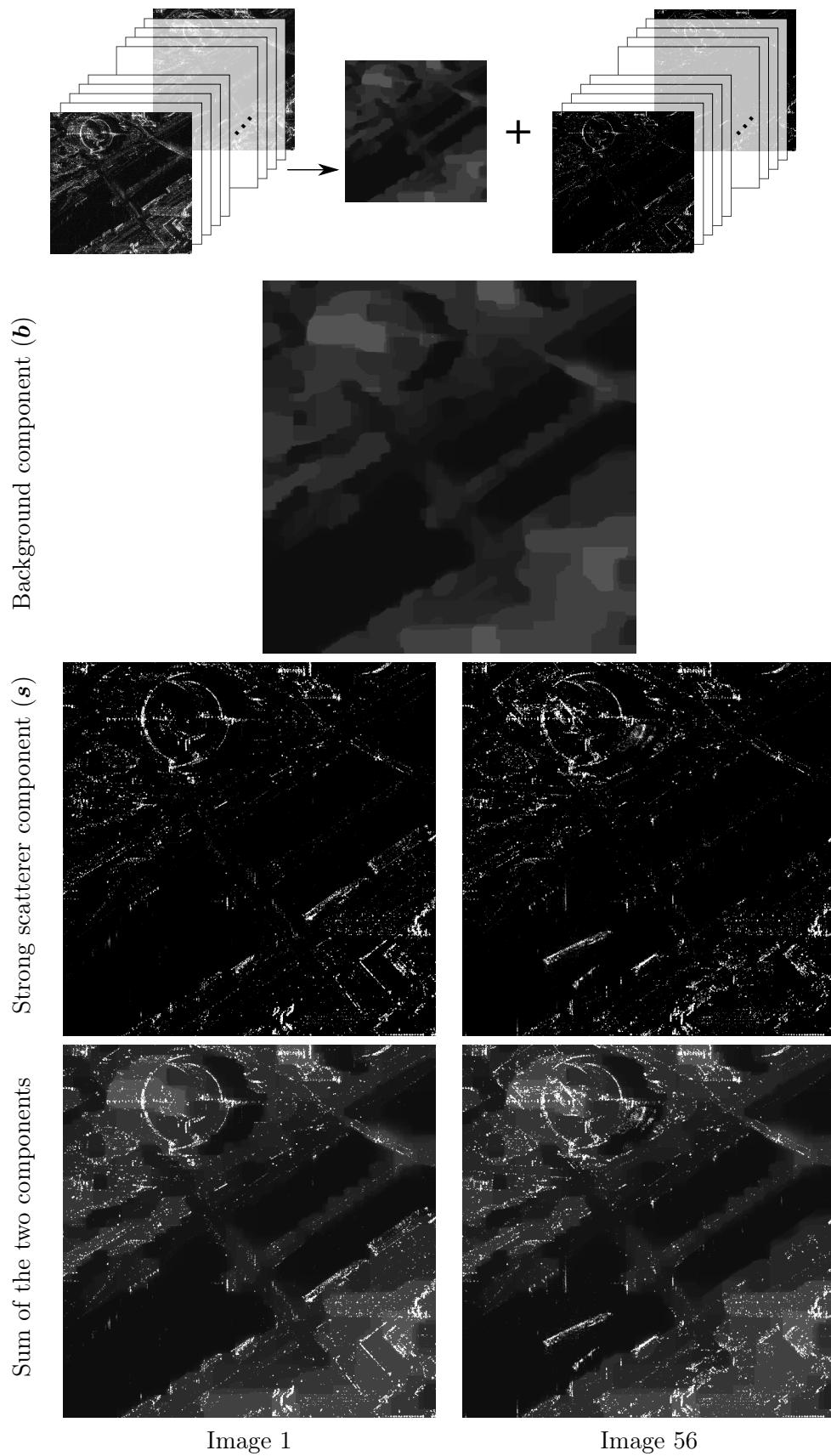


Figure 7.9 – Result of the TV1BG model on the full stack. An animation showing all the results is available at www.sylvainlobry.com/phd.

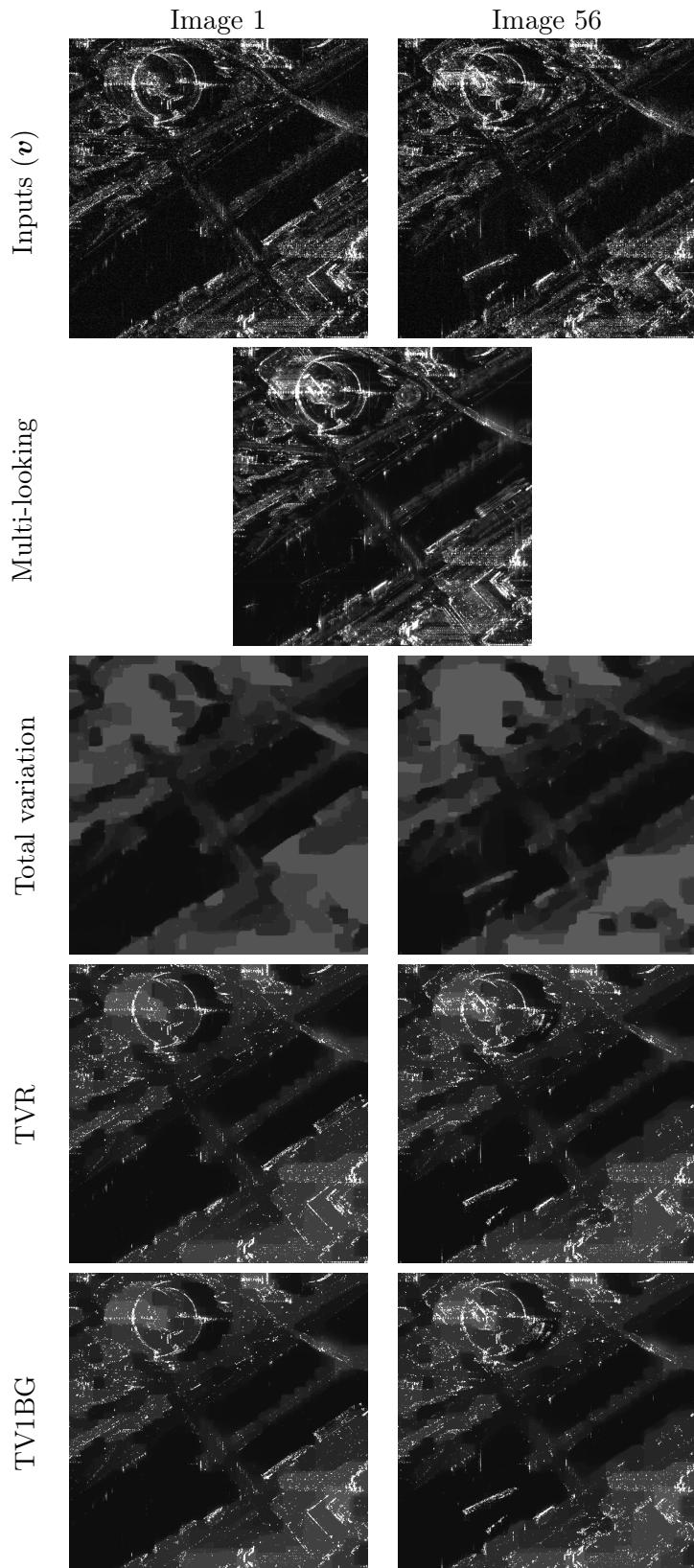


Figure 7.10 – Comparison of the different results.

7.3 Change detection

While all three models can be used for change detection, TV1C is particularly well suited for this purpose. We have seen in subsection 6.2.1 that it accounts for the changes directly in the model via a L0 regularization enforced on the number of changes.

Qualitative evaluation: First we present results on the dataset shown in Figure 7.5. They have been computed with the following parameters for the tuning rules described in section 6.4:

- $n_{\text{quantif}} = 10$
- $p = 0.95$
- $c = 3$.
- $\beta_C = 3$.
- F (side of the filling window for the efficient optimization) = 350
- C (side of the computing window for the efficient optimization) = 400

The results are presented in Figure 7.11. We can note that even though the model does not enforce any spatial regularity in the changes, changes are quite grouped. This indicates that it is not necessary to add such a regularization.

In Figure 7.12, Figure 7.13 and Figure 7.14, we look at specific area:

- "Maison de la Radio" (Figure 7.12): as explained in the dataset description (see section 7.2), it was being renovated during the acquisitions. We can see that we detect the construction of the central tower.
- "Beaugrenelle" (Figure 7.13): this area contains a mall that was first destroyed, then re-built. A few changes are detected for the destruction part but not for the construction. This is interesting, as it outlines the main weakness of our proposed method: it considers at most one change per pixel during the time series. In the case of destruction followed by a construction, it is not well adapted.
- "Boat" (Figure 7.14): In this crop, a boat is present during the last acquisition of the time series. In this case, it should probably not be considered as a change and be represented as "no strong scatterers" (\mathcal{H}_0) pixels. However, as we have seen in section 6.4, the proposed model is not CFAR, meaning that a strong scatterer that is only present at the last date is more likely to be detected as a change than a strong scatter only present at the first date.

We also present the results of the change detection on the following datasets:

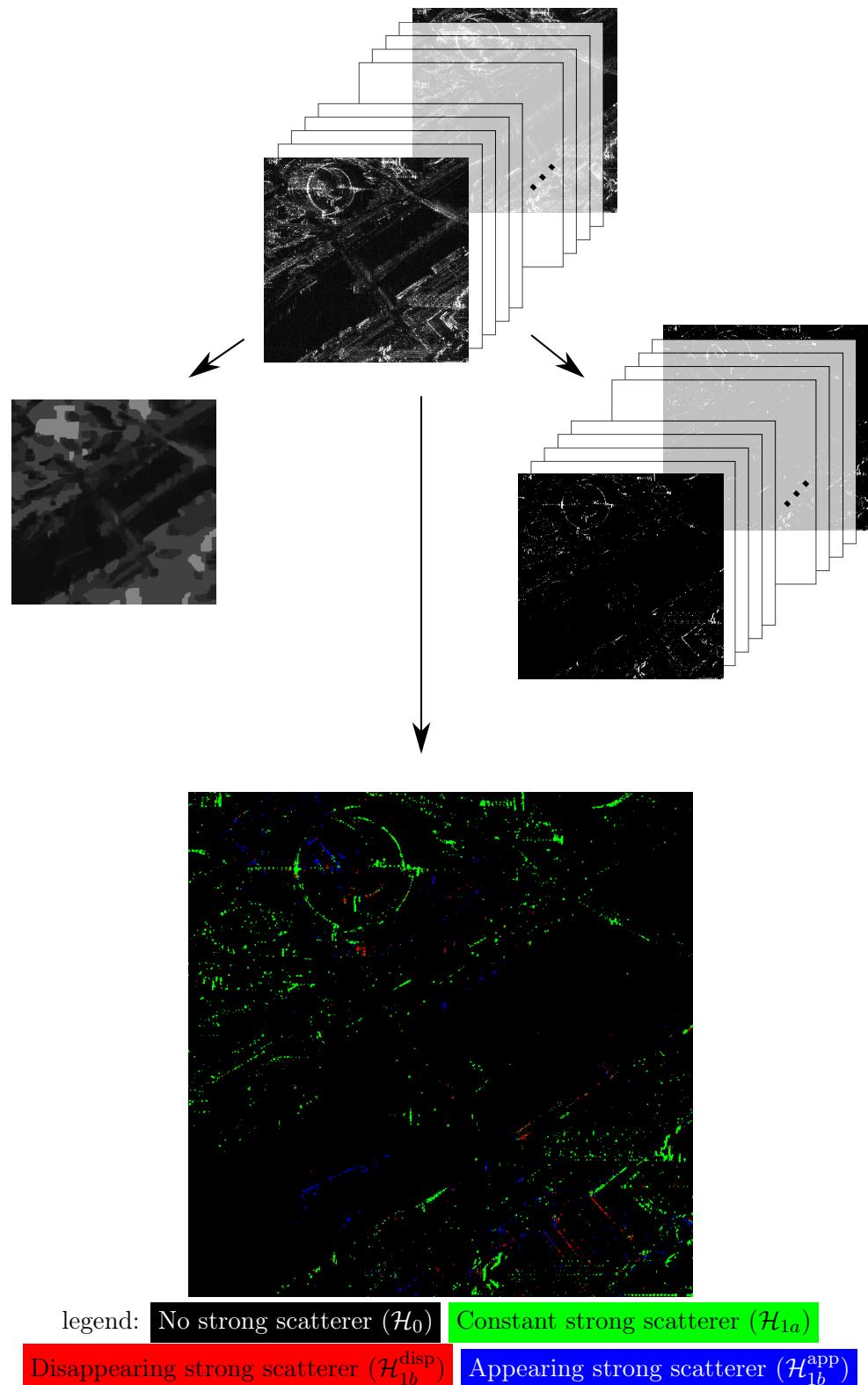


Figure 7.11 – Results of the TV1C model on the Paris dataset.

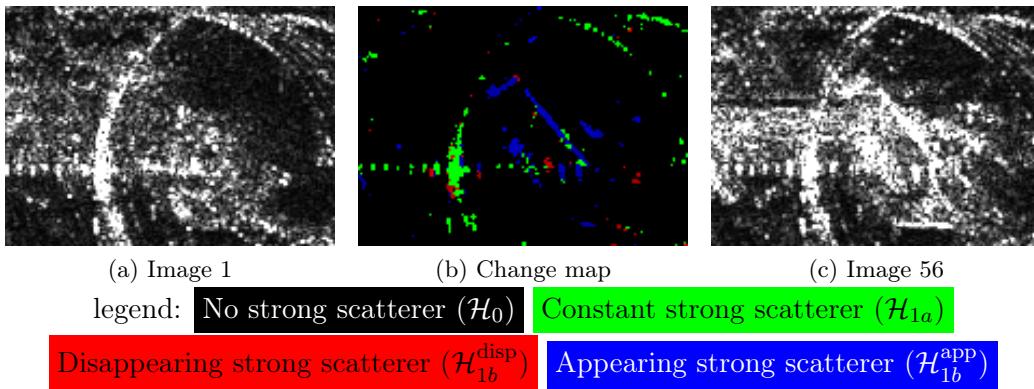


Figure 7.12 – Changes in the "Maison de la Radio". We can see that in this case, the newly constructed tower at the middle of the building was well detected.

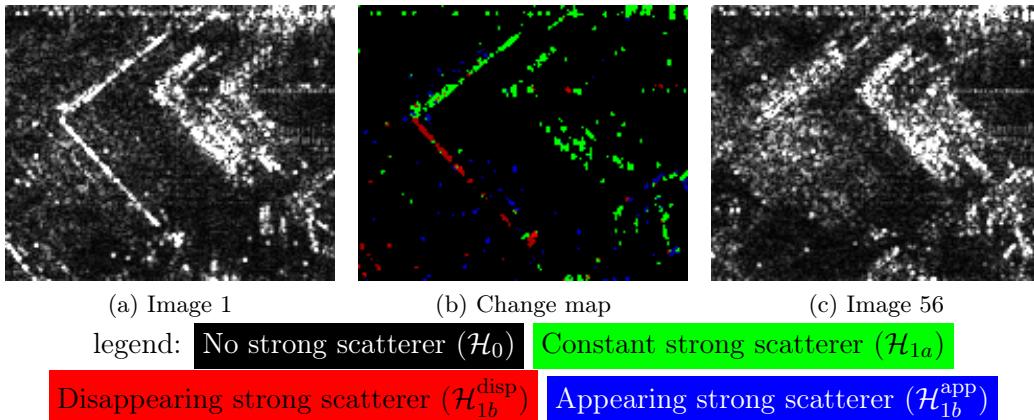


Figure 7.13 – Changes in the Beaugrenelle area (bottom-right part of the images). This area was in construction during the acquisitions. Only a few changes are detected, and no appearing strong scatterers. This area is typically not well supported by our model as it features two changes (in this case a disappearance and an appearance).

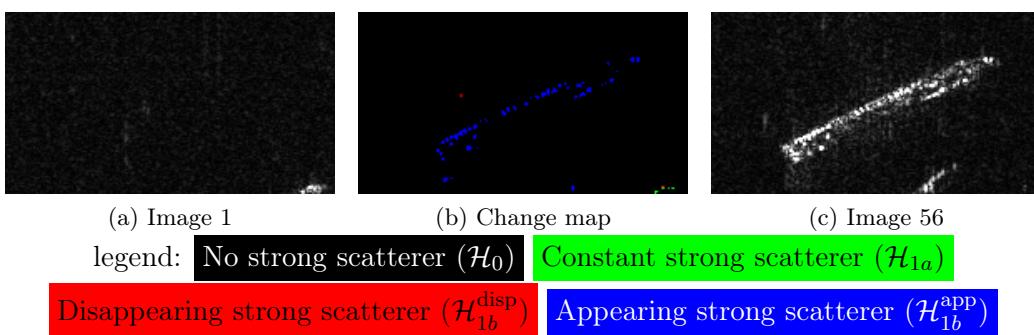


Figure 7.14 – Crop on "La Seine". Here, boats are sometimes sailing and this is the case in the last image of the time series. While this should probably not be detected as an appearance, it is the case using our model. This is due to the fact that our model is not CFAR as explained in section 6.4.

- A series of 17 images of size 2048×2048 from Paris acquired by TerraSAR-X (in spotlight HR mode, resolution: $1m \times 1m$) between 01/24/2009 and 04/09/2010 in [Figure 7.15](#);
- A series of 13 images of size 512×512 from Saint-Gervais, France acquired by TerraSAR-X (in stripmap mode, $2m \times 2m$) between 05/31/2009 and 09/25/2011 in [Figure 7.16](#);
- A series of 25 images of size 512×512 from Saclay, France acquired by Sentinel-1A (in Interferometric Wide Swath, $5m \times 20m$) between 03/06/2015 and 02/29/2016 in [Figure 7.17](#).

All the results are computed using a fixed set of parameters:

- $\beta_{BG} = 0.1$
- $\beta_S = 3 \times T$ (where T is the number of images).
- $\beta_C = 5.$

We can see that using a fixed set of parameters on datasets with different resolutions gives similar results. It shows that it will be feasible to apply this model on different datasets with minimal user intervention.

Comparison with state of the art: We now compare our algorithm to other change detection algorithms for amplitude SAR images.

In ([Lobry et al., 2016](#)), we also presented a change detection application using models TVR and TV1BG. The same hypothesis was made as a starting point: changes in urban areas are mostly reflected by changes in the strong scatterers. To perform the change detection between two dates, the two strong scatterers components were:

- first binarized, to achieve robustness with respect to radiometric fluctuations;
- compared on a spatial neighboring, to achieve robustness with respect to small changes in the vicinity of a detected strong scatterer:

$$c_i = \left| \sum_{\delta_i} s_{\delta_i, t_1}^{\text{bin}} - \sum_{\delta_i} s_{\delta_i, t_2}^{\text{bin}} \right|, \quad (7.5)$$

were s^{bin} is the binarization of s , and δ_i a set of pixels in the neighborhood of i . c can then be thresholded in order to retrieve the changes.

We compare the results on the Saint-Gervais dataset of the TV1C method and the method using TV1BG with the followings:

- the GLRT from ([Lombardo and Oliver, 2001](#));
- an approach based on Wilcoxon tests from ([Krylov et al., 2012](#));

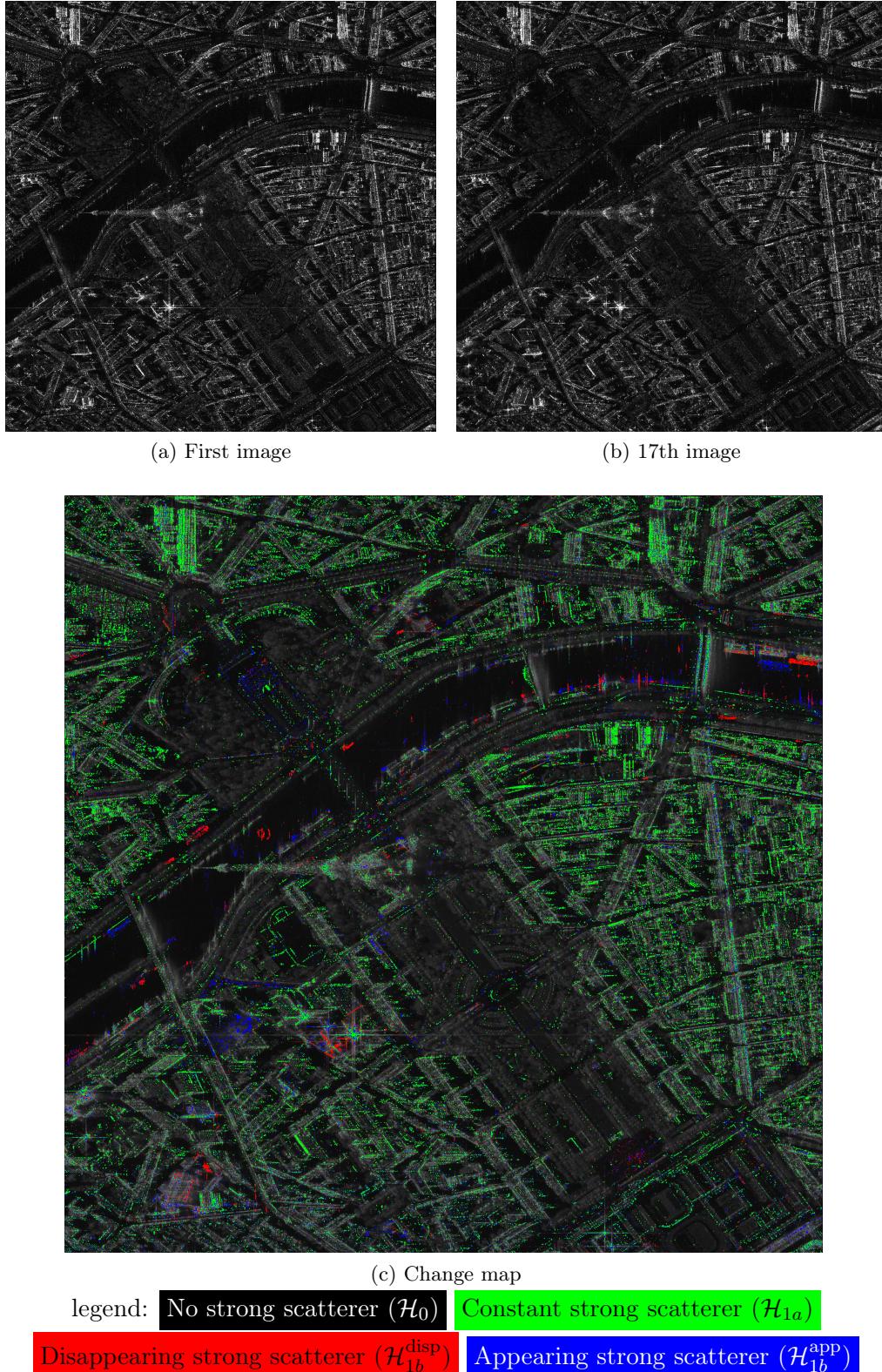
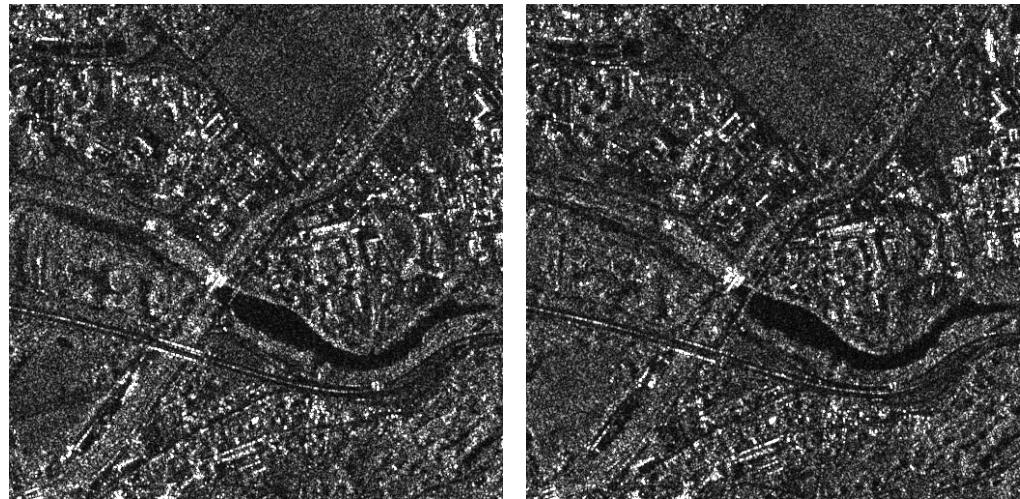
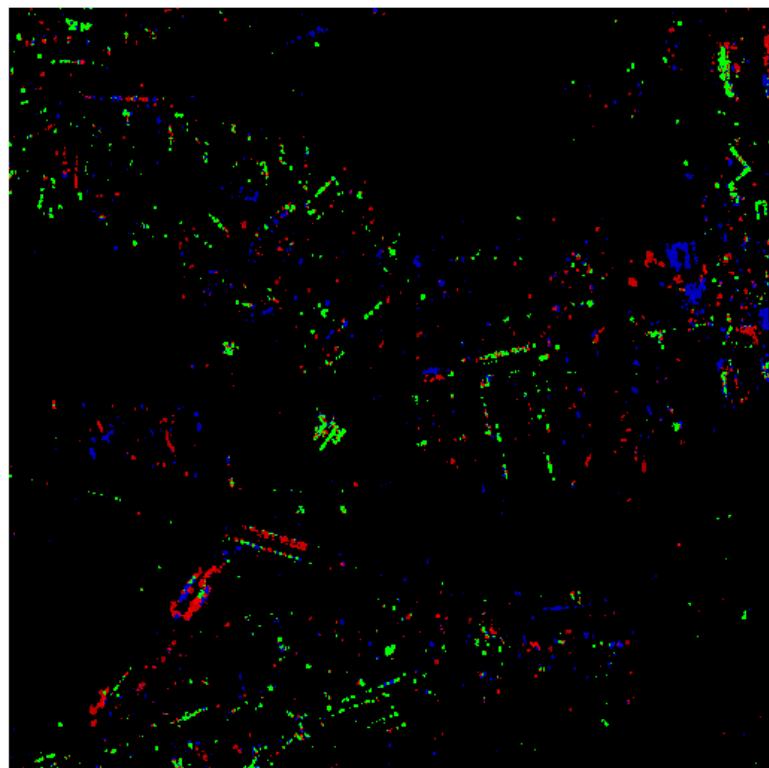


Figure 7.15 – Results on the second Paris dataset, acquired by TerraSAR-X in spotlight mode.



(a) First image

(b) 13th image



(c) Change map

legend: No strong scatterer (\mathcal{H}_0) Constant strong scatterer (\mathcal{H}_{1a})

Disappearing strong scatterer ($\mathcal{H}_{lb}^{\text{disp}}$)

Figure 7.16 – Results on the Saint-Gervais dataset, acquired by TerraSAR-X in stripmap mode.

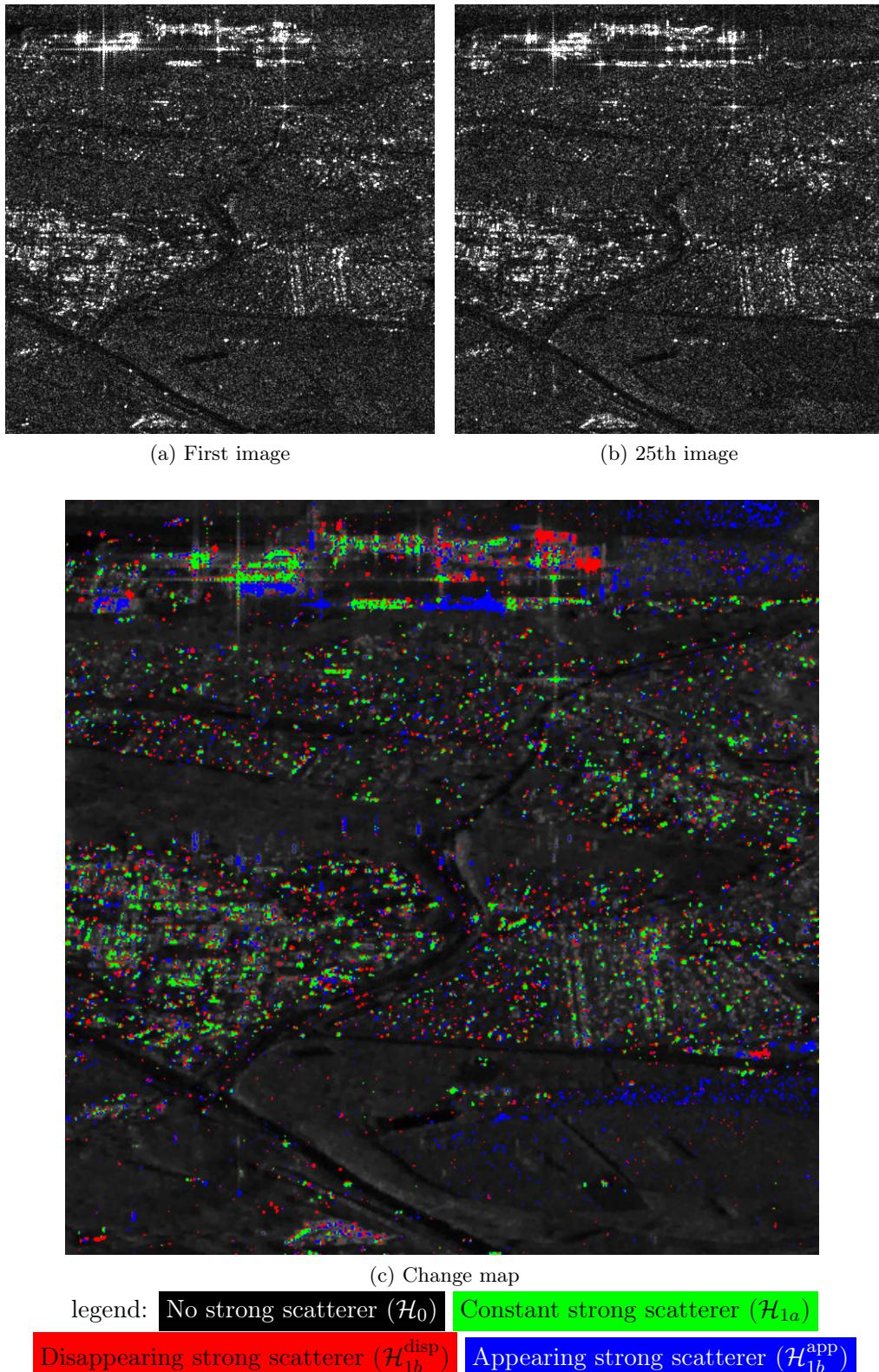


Figure 7.17 – Results on the Saclay dataset, acquired by Sentinel-1A in Interferometric Wide Swath mode.

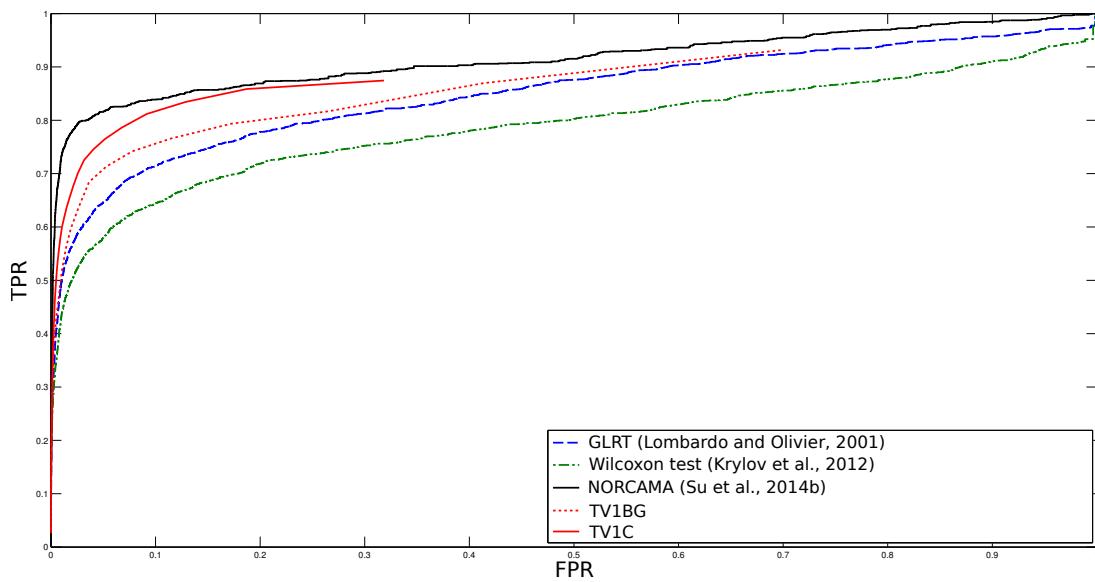


Figure 7.18 – ROC curve comparing the results of Change detection on the Saint-Gervais dataset.

- NORCAMA from ([Su et al., 2014a](#)) which applies a likelihood ratio test on images denoised with a non-local algorithm ([Su et al., 2014b](#)).

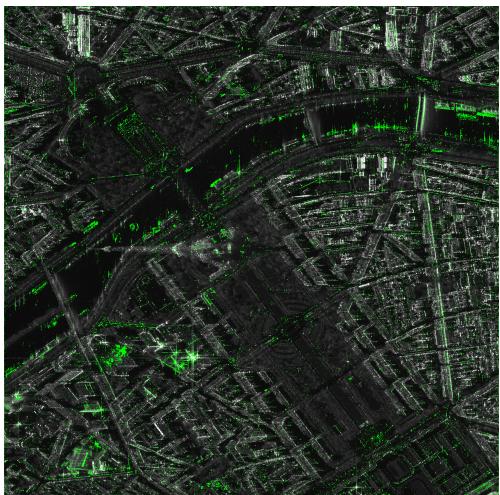
The results, computed using a ground truth made manually by ([Su et al., 2014a](#)), are presented as a **Receiver Operating Characteristic (ROC)** curve (comparing **TPR** to **FPR**) in [Figure 7.18](#). The reason the proposed models can not achieve 100% **TPR** is because it makes the assumption that all changes occurs in the strong scatterers, which is not always verified in practice. We can see that our model obtains poorer performances compared with ([Su et al., 2014a](#)). On the opposite, it also provides the background and the strong scatterers components that can be useful for the interpretation of the changes. We propose to visually compare the results of TV1C and ([Su et al., 2014a](#)) on the second Paris dataset. Since ([Su et al., 2014a](#)) uses a first step of denoising, results show more spatial regularity. On the other hand, using our method, only changes in strong scatterers are detected. This is more visible when we zoom in as in [Figure 7.20](#) where the spatial regularity results in false detections. Therefore the choice of one model versus another has to be made with respect to the desired application.



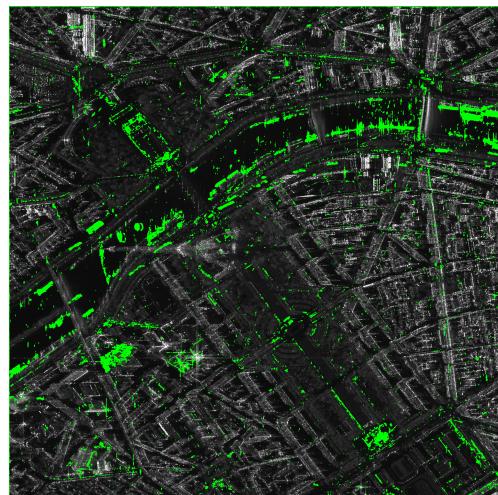
(a) First image



(b) 17th image



(c) Results using TV1C



(d) Results using NORCAMA

Figure 7.19 – Comparison of the results using our method and (Su *et al.*, 2014a). Changes are in green.

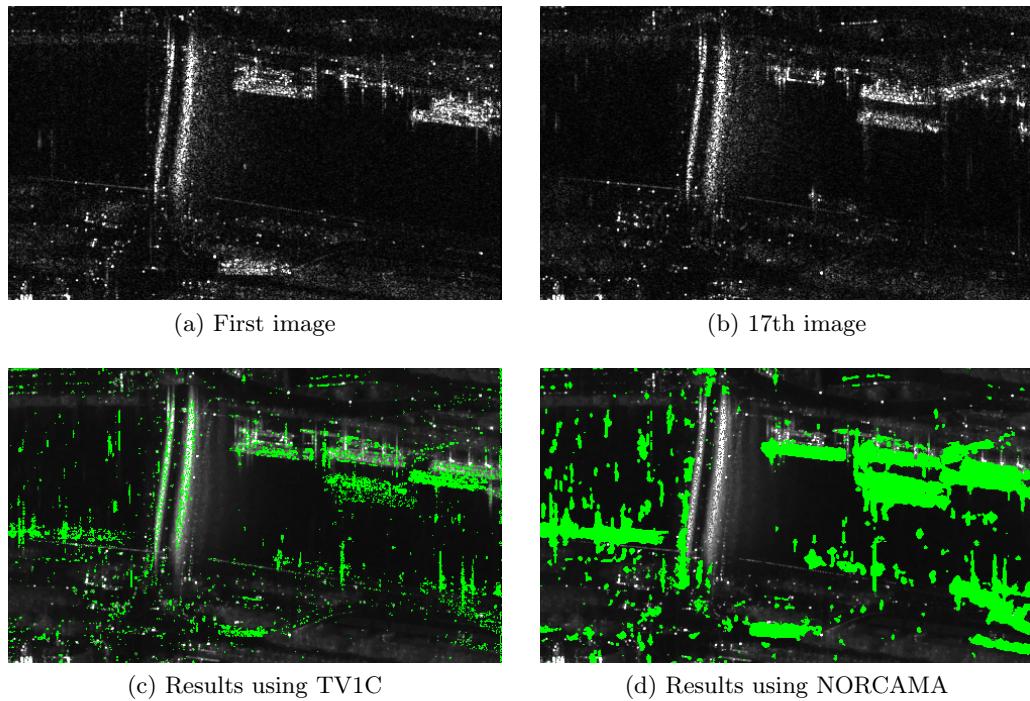


Figure 7.20 – Comparison of the results using our method and (Su *et al.*, 2014a). Changes are in green.

Summary: Change detection

We have presented results using the TV1C model and a change detection technique as a post-processing step to TV1BG. These results are compared to state-of-the art methods. TV1C model can be efficiently applied to find well located changes in urban areas. However, it has several drawbacks:

- it make the (strong) hypothesis that changes occur only in the strong scatterer component;
- it can only detect at most one change per pixel in the whole time series;
- the probability of detecting a change is not the same depending on the date.

While these drawbacks must be kept in mind when using the model, we saw that it performs well compared to other classically used methods. Note that we could allow for more changes at each pixel at the price of an higher computational cost, and the third point could be solved by adding a time-depending factor to the β_C parameter (which is work in progress). It offers an alternative to widely used methods in terms of results (as it will also provide a background regularization and a strong scatterer detection).

7.4 Conclusion

We have presented results of application in strong scatterers detection, regularization and change detection of the models developed in [chapter 6](#). From the results, we can see that the proposed method performs generally well on these applications. We showed that it is more interesting to use the L0 pseudo-norm than the L1 for the strong scatterers detection, and that a convex relaxation comes at a cost in performances. We showed that our models can be a good alternative for amplitude images regularization. Finally, the change detection results showed that our method has good results on images from different sensors and of different resolutions. However, the hypothesis made that there is at most one change per pixel and the fact that the probability of detecting a change in not the same depending on the date leads to some bad detections in the results.

It would be theoretically possible to allow for more than one change by modifying the hierarchical hypothesis tests to include hypothesis for more than one change but this would lead to an exponential increase in the complexity.

Conclusion and perspectives

Conclusion

The main objective of this thesis is to provide reliable methods for the land/water classification of **SWOT** images. This task is a critical step towards the reliable estimation of water elevation. We have divided the main objects of interest for hydrologists in two categories defined on their general shape:

- Large and compact objects: widely-used binary classification methods fail to take into account the intra-class variation present in **SWOT** (described in [section 2.4](#)). The contributions presented in this thesis are a straightforward application of a general framework (**MRF** for an Ising model) to variable class parameters and their estimation. We presented two novel methods for the estimation of the parameters that can be embedded in an iterative process (alternating classification and estimation). The first method constructs a partition of the image with regions where parameters can be considered constant. The second method relies on another **MRF** for the estimation of the parameters. We have shown that these methods are not only relevant for **SWOT** images with strong variations, but can also be applied to very different sensors (we have used the example of an image acquired by SETHI, a P-band airborne sensor) which presents similar features.
- Narrow curvilinear networks: mainly representing rivers, they do not respect the hypothesis of the isotropic spatial compactness made by most classification algorithm. To this effect, we adapted previous works developed for road detection and vessels detection in fundus of eyes images. This method is a combination of two steps: the first step detects segments at the pixel level and the second step connects the detected segments and selects the meaningful connections with respect to their global contribution to geometrical properties of river networks. This method only improves by a fraction the water detection performances but is of prime importance for the study of these objects using **SWOT** data.

Part of the PhD was also dedicated to the operational implementation of the developed methods in the **CNES** processing chain. To this date, we implemented the **MRF** taking into account the variational class parameters and the Markovian estimation for the large and compact objects detection. The method dedicated to the detection of narrow curvilinear networks is in the process of being implemented.

Considering the revisiting time of satellites (approximately 21 days in the case of **SWOT**), multitemporal processing of the data can represent a strong improvement in terms of classification results. As this thesis is an upstream work (the satellite will be launched in April 2021), meaningful multitemporal data was not available for **SWOT**. This prevents efficient evaluation of the methods in this context. Therefore, we worked on the adaptation of the algorithms proposed for the large and compact objects detection to various problems in a different context: multitemporal **SAR** images of urban areas. We presented models taking into account the strong scatterers that appear in such images and that usually break the assumptions made by regularization algorithms. The general idea is to describe the image as a sum of two components: a component representing the backgrounds and another one representing the strong scatterers. Spatial regularity can then be enforced on the backgrounds, while the strong scatterers component is controlled by its sparsity.

- TVR: this model is adapted to the regularization of time series. It outputs one background and one strong scatterers component for each input image.
- TV1BG: this model only outputs one background representing the whole multitemporal stack, along with one strong scatterers component for each input image. It is well adapted for strong scatterers detection.
- TV1C: as for TV1BG, it outputs one background and as many strong scatterers components as the number of input images. In addition, it features a sparsity constraint on the number of changes that can occur in the strong scatterers component stack, making it adapted for change detection.

The proposed models can be optimized exactly and could be easily adapted to new applications. One drawbacks of the proposed optimization method, based on graphcuts, is the memory needed which can prevent its application on large time series. To this effect, we proposed a non optimal adaptation of **MRF** models which makes possible to set a trade-off between needed memory, the computation time and the quality of the result. This contribution will be useful for the processing of large time series of **SWOT** data.

Perspectives

When realistic data will be available, it will be of prime importance to study the adaptation of the proposed methods for the multi-temporal processing of urban **SAR** data to the case of water detection in **SWOT** images. The adaptation of the TV_{3D} defined in the third part to an Ising model is straightforward. Adapting the methods for the estimation of the parameters has to be approached differently for the two proposed methods:

- considering the region-based approach, the partition found on a first image could be used when new images of the same area are available. However, this will

not be robust to changes, which is of interest for **SWOT**. A solution could be to restart the partitioning process on a new image using a coarse partition (i.e. a few iterations before convergence) from a first image as an initialization.

- Considering the Markovian approach, the prior could be modified to add a temporal link (i.e. adding an L2 distance between pixels at the same site and at consecutive dates). However, this would not be adapted to abrupt changes (e.g. floods). Another solution would be to use a previous parameter map as an initialization (instead of the Xfactor for instance).

Another task which will have to be tested is the impact of a difference between the real σ_0 of each class and those assumed for the computation of the initialization using the Xfactor. The difference could be problematic for the simple methods that do not re-estimate the parameters.

The work achieved on the processing of multi-temporal could be extended in many ways. The model adapted to change detection (TV1C) as presented has two main drawbacks: it is not able to take into account more than one change at each pixel and the probability of detecting a change is not constant through time. The fact that only one change can be considered at each pixel can be problematic when processing large time series. An extension to a fixed number of changes is straightforward, but the complexity grows exponentially with the maximum number considered. Therefore, other strategies should be explored. The fact that the probability of detection is not constant comes from the variability induced by speckle. A strategy of normalization of the probability should be applied.

Publications

International journal

- Sylvain Lobry, Loïc Denis, Florence Tupin
Multi-temporal SAR image decomposition into strong scatterers, background, and speckle
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS), 2016

National journal

- Sylvain Lobry, Loïc Denis, Florence Tupin, Weiying Zhao
Décomposition de séries temporelles d'images SAR pour la détection de changement
Traitement du Signal (GRETSI, Lavoisier) (accepted)

International conferences

- Sylvain Lobry, Loïc Denis, Florence Tupin, Roger Fjørtoft
Double MRF for water classification in SAR images by joint detection and reflectivity estimation
International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, USA, July 2017.
- Sylvain Lobry, Florence Tupin, Roger Fjørtoft
Unsupervised detection of thin water surfaces in SWOT images based on segment detection and connection
International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, USA, July 2017.
- Weiying Zhao, Sylvain Lobry, Henri Maître, Jean-Marie Nicolas, Florence Tupin
Urban area change detection based on generalized likelihood ratio test
International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp), Bruges, Belgium, June 2017
- Sylvain Lobry, Florence Tupin, Loïc Denis
A decomposition model for scatterers change detection in multi-temporal series of SAR images
International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, July 2016
- Sylvain Lobry, Florence Tupin, Roger Fjørtoft
Non-Uniform Markov Random Fields for Classification of SAR Images
European Conference on Synthetic Aperture Radar (EUSAR), Hamburg, Germany, June 2016

- Sylvain Lobry, Florence Tupin, Loïc Denis

Sparse+Smooth decomposition model for multi-temporal SAR images

International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp), Annecy, France, July 2015

National conferences

- Sylvain Lobry, Loïc Denis, Florence Tupin, Roger Fjørtoft

Détection de l'eau dans les images radar du futur satellite SWOT

GRETSI, September 2017

- Sylvain Lobry, Florence Tupin, Loïc Denis

Un modèle de décomposition pour la détection de changement dans les séries temporelles d'images RSO

RFIA, Clermont-Ferrand, France, July 2016

Acronyms

CFAR Constant False Alarm Rate. 111, 126, 128, 142

CNES Centre National D'Études Spatiales. 2, 59, 60, 137, 142

DEM Digital Elevation Model. 3, 19, 48–50, 54, 59, 62, 142

ENL Equivalent Number Of Looks. 109, 142

ER Error Rate. 62, 142

FPR False Positive Rate. 62, 63, 133, 142

GLRT Generalized Likelihood Ratio Test. 96, 98, 99, 103, 113, 142

HR High Rate. 17, 142

ICM Iterated Conditional Modes. 24–26, 29, 80, 142

JPL Jet Propulsion Laboratory. 2, 59, 60, 142

KaRIn Ka-band Radar Interferometer. 9, 17, 142

LOS Line of Sight. 10, 142

LR Low Rate. 17, 142

LRT Likelihood Ratio Test. 95, 96, 98, 142

MAP Maximum a posteriori. 21, 23, 38, 59, 60, 62–68, 93, 142

MCC Matthews Correlation Coefficient. 62, 64, 142

MLE Maximum Likelihood Estimate. 22, 38, 39, 45, 47, 49, 52, 53, 56, 59, 60, 63, 65, 67, 73, 91, 96, 98, 99, 102–104, 120, 142

MRF Markov Random Field. 3, 22–24, 27, 29, 35, 38, 42, 44–47, 49, 51, 56–60, 63–68, 71, 72, 77, 81, 84, 85, 92, 111, 137, 138, 142

ONERA Office National d'Études et de Recherches Aérospatiales. 59, 60, 142

PD Probability of Detection. 116, 117, 142

PFA Probability of False Alarm. 116, 117, 142

RADAR RAdio Detection and Ranging. 10, 11, 142

ROC Receiver Operating Characteristic. 133, 142

SA Simulated Annealing. 24, 26, 29, 142

SAR Synthetic Aperture Radar. 2–5, 9, 11, 12, 17, 20, 33, 36, 37, 40, 41, 113, 138, 142

SLC Single-Look Complex. 16, 19, 142

SNR Signal-to-noise ratio. 3, 17, 19, 20, 35, 48, 142

SWOT Surface Water & Ocean Topography. 1–5, 9, 17–20, 33, 35, 40, 41, 45, 46, 48, 59, 60, 137–139, 142

TPR True Positive Rate. 62–64, 133, 142

TV Total Variation. 4, 27, 29, 93, 94, 107, 113, 120–122, 142

Bibliography

- A. Achim, P. Tsakalides, and A. Bezerianos. SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling. *IEEE Transactions on Geoscience and Remote Sensing*, 41(8), 2003. [90](#)
- R. Amhaz, S. Chambon, J. Idier, and V. Baltazart. 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, Oct 2016. [XVIII, 71](#)
- F. Argenti, T. Bianchi, and L. Alparone. Multiresolution MAP Despeckling of SAR Images Based on Locally Adaptive Generalized Gaussian pdf Modeling. *IEEE Transactions on Image Processing*, 15(11):3385–3399, 2006. [90](#)
- G. Aubert and J.-F. Aujol. A variational approach to removing multiplicative noise. *SIAM Journal on Applied Mathematics*, 68(4):925–946, 2008. [90](#)
- Jean-François Aujol, Gilles Aubert, Laure Blanc-Féraud, and Antonin Chambolle. Image decomposition application to SAR images. In *Scale Space Methods in Computer Vision*, pages 297–312. Springer, 2003. [90](#)
- J.F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition: modeling, algorithms, and parameter selection. *IJCV*, 67(1), 2006. [92](#)
- R Azencott. Markov field approach: parameter estimation by qualitative boxes. *Course*, 1992. [43](#)
- M. Bachmann et al. TerraSAR-X antenna calibration and monitoring based on a precise antenna model. *IEEE Transactions on Geoscience and Remote Sensing*, 48(2):690–701, 2010. [18, 48](#)
- I. Ben Ayed, A. Mitiche, and Z. Belhadj. Multiregion level-set partitioning of synthetic aperture radar images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(5):793–800, 2005. [41](#)
- I. Ben Ayed, A. Mitiche, and Z. Belhadj. Polarimetric image segmentation via maximum-likelihood approximation and efficient multiphase level-sets. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(9):1493–1500, 2006. [41](#)

- J. Bioucas-Dias and M. Figueiredo. Multiplicative noise removal using variable splitting and constrained optimization. *Image Processing, IEEE Transactions on*, 19(7):1720–1730, 2010. [90](#)
- I. Bloch. Information combination operators for data fusion: A comparative review with classification. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 26(1):52–67, 1996. [XX](#), [74](#)
- C. Bouman, M. Shapiro, et al. A multiscale random field model for bayesian image segmentation. *Image Processing, IEEE Transactions on*, 3(2):162–177, 1994. [52](#)
- Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(9):1124–1137, 2004. [28](#), [29](#), [102](#), [105](#)
- Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11):1222–1239, Nov 2001. [27](#)
- A. Buades, B. Coll, and J.-M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005. [XXII](#), [90](#)
- F. Cao, F. Tupin, J.-M. Nicolas, R. Fjørtoft, and N. Pourthié. Extraction of water surfaces in simulated Ka-band SAR images of KaRIn on SWOT. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 3562–3565, 2011. [XI](#), [XIV](#), [XVIII](#), [3](#), [40](#), [71](#), [75](#)
- C. Cazals, S. Rapinel, P.-L. Frison, A. Bonis, G. Mercier, C. Mallet, S. Corgne, and J.-P. Rudant. Mapping and characterization of hydrological dynamics in a coastal marsh using high temporal resolution sentinel-1a images. *Remote Sensing*, 8(7):570, 2016. [XIV](#), [40](#)
- M. Cetin and W. Karl. Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization. *Image Processing, IEEE Transactions on*, 10(4):623–631, 2001. [92](#)
- T. Chan and L. Vese. Active contours without edges. *Image processing, IEEE transactions on*, 10(2):266–277, 2001. [40](#), [41](#)
- J. Chen, Y. Chen, W. An, Y. Cui, and J. Yang. Nonlocal filtering for polarimetric SAR data: A pretest approach. *IEEE Transactions on Geoscience and Remote Sensing*, 49(5), 2011. [90](#)
- C. Chesnaud, P. Refregier, and V. Boulet. Statistical region snake-based segmentation adapted to different physical noise models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11):1145–1157, Nov 1999. [41](#)

- K. Conradsen, A. A. Nielsen, and H. Skriver. Determining the Points of Change in Time Series of Polarimetric SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(5):3007–3024, May 2016. [92](#)
- D. Cozzolino, S. Parrilli, G. Scarpa, G. Poggi, and L. Verdoliva. Fast Adaptive Nonlocal SAR Despeckling. *Geoscience and Remote Sensing Letters, IEEE*, 11(2):524–528, Feb 2014. [90](#)
- C.-A. Deledalle, L. Denis, and F. Tupin. Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. *IEEE Transactions on Image Processing*, 18(12):2661–2672, 2009. [90](#)
- C.-A. Deledalle, L. Denis, and F. Tupin. NL-InSAR: Nonlocal interferogram estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 49(4):1441–1452, 2011. [XXII, 90](#)
- C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäger. NL-SAR: A Unified Nonlocal Framework for Resolution-Preserving (Pol)(In)SAR Denoising. *Geoscience and Remote Sensing, IEEE Transactions on*, 53(4):2021–2038, April 2015. [XXI, 40, 81, 90](#)
- C.-A. Deledalle, L. Denis, G. Ferraioli, V. Pascazio, G. Schirinzi, and F. Tupin. Very high resolution and interferometric SAR: Markovian and patch-based non-local mathematical models. In Gabriele Moser and Josiane Zerubia, editors, *Mathematical Models for Remote Sensing Image Processing*. Springer International Publishing, 2018. [11](#)
- L. Denis, F. Tupin, J. Darbon, and M. Sigelle. SAR image regularization with fast approximate discrete minimization. *Image Processing, IEEE Transactions on*, 18(7):1588–1600, July 2009. [90, 105](#)
- L. Denis, F. Tupin, and X. Rondeau. Exact discrete minimization for TV+L0 image decomposition models. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 2525–2528. IEEE, 2010. [XI, 3, 93](#)
- E.-W. Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959. [XX, 76](#)
- J. Edmonds and R. Karp. Theoretical improvements in algorithmic efficiency for network flow problems. *Journal of the ACM (JACM)*, 19(2):248–264, 1972. [29](#)
- M. Elad, J.L. Starck, P. Querre, and DL Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *App. Comp. Harm. Anal.*, 19(3):340–358, 2005. [92](#)
- J. Ender. On compressive sensing applied to radar. *Signal Processing*, 90(5):1402–1414, 2010. [92](#)

- A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci. A new algorithm for processing interferometric data-stacks: SqueeSAR. *Geoscience and Remote Sensing, IEEE Transactions on*, 49(9):3460–3470, 2011. [92](#)
- R. Fjørtoft et al. KaRIn on SWOT: Characteristics of Near-Nadir Ka-Band Interferometric SAR Imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 52(4):2172–2185, April 2014. [17](#), [18](#)
- R. Fjørtoft, A. Lopes, P. Marthon, and E. Cubero-Castan. An optimal multiedge detector for SAR image segmentation. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(3):793–802, 1998. [XIV](#), [40](#)
- R. Fjørtoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin. Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(3):675–686, 2003. [47](#)
- R. Fjørtoft, J. M. Gaudin, N. Pourthie, C. Lion, A. Mallet, J. C. Souyris, C. Ruiz, F. Koudogbo, J. Duro, P. Ordoqui, and A. Arnaud. 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, July 2010. [17](#)
- L. Ford and D. Fulkerson. Maximal flow through a network. *Canadian journal of Mathematics*, 8(3):399–404, 1956. [29](#)
- G. Fornaro, D. Reale, and F. Serafino. Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers. *Geoscience and Remote Sensing, IEEE Transactions on*, 47(1):224–237, 2009. [92](#)
- M. Fredman and R. Tarjan. Fibonacci heaps and their uses in improved network optimization algorithms. *Journal of the ACM (JACM)*, 34(3):596–615, 1987. [76](#)
- F. Galland, N. Bertaux, and P. Réfrégier. Minimum description length synthetic aperture radar image segmentation. *IEEE Transactions on Image Processing*, 12(9):995–1006, 2003. [41](#)
- S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-6(6):721–741, Nov 1984. [XIV](#), [22](#), [25](#)
- J. Gilles. Image Decomposition: Theory, Numerical Schemes, and Performance Evaluation. *Adv. in Imag. Electron Phys.*, 158:89–137, 2009. [92](#)
- P. Gleick. Water in crisis: a guide to the worlds fresh water resources. 1993. [IX](#), [1](#)
- A. Goldberg and R. Tarjan. A new approach to the maximum-flow problem. *Journal of the ACM (JACM)*, 35(4):921–940, 1988. [29](#)

- J. Goodman. Some fundamental properties of speckle. *JOSA*, 66(11):1145–1150, 1976. [9](#), [12](#)
- D. Greig et al. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 271–279, 1989. [27](#), [43](#), [44](#), [45](#), [51](#), [55](#), [58](#)
- S. Grosdidier et al. River network detection on simulated SWOT images based on curvilinear denoising and morphological detection. In *2012 IEEE International Geoscience and Remote Sensing Symposium*, pages 5454–5457, July 2012. [71](#)
- H. Ishikawa. Exact optimization for Markov random fields with convex priors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(10):1333–1336, 2003. [XXVI](#), [27](#), [28](#), [101](#), [102](#)
- O. Jamriska, D. Sykora, and A. Hornung. Cache-efficient graph cuts on structured grids. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 3673–3680, June 2012. [105](#)
- Emmanuel Trouvé Jean-Marie Nicolas, Florence Tupin. Les bases de l'imagerie satellitaire. Technical report, Télécom ParisTech, 2012. [14](#), [109](#)
- M. Jung, G. Peyré, and L. Cohen. Nonlocal active contours. *SIAM Journal on Imaging Sciences*, 5(3):1022–1054, 2012. [41](#)
- V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, 2004. [27](#)
- V. Krylov, G. Moser, A. Voisin, S. Serpico, and J. Zerubia. Change detection with synthetic aperture radar images by wilcoxon statistic likelihood ratio test. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 2093–2096. IEEE, 2012. [129](#)
- J. Lambin, R. Morrow, L.-L. Fu, J. Willis, H. Bonekamp, J. Lillibridge, J. Perbos, G. Zaouche, P. Vaze, W. Bannoura, et al. The ostm/jason-2 mission. *Marine Geodesy*, 33(S1):4–25, 2010. [1](#)
- J. Lee, S. Cloude, K. Papathanassiou, M. Grunes, and I. Woodhouse. Speckle filtering and coherence estimation of polarimetric SAR interferometry data for forest applications. *IEEE Transactions on Geoscience and Remote Sensing*, 41(10), 2003. [90](#)
- J.-S. Lee. Speckle analysis and smoothing of synthetic aperture radar images. *Computer graphics and image processing*, 17(1):24–32, 1981. [XXII](#), [40](#)
- S. Li. *Markov random field modeling in image analysis*. Springer Science & Business Media, 2009. [22](#)

- H. Liu and K. Jezek. A complete high-resolution coastline of antarctica extracted from orthorectified radarsat sar imagery. *Photogrammetric Engineering & Remote Sensing*, 70(5):605–616, 2004. [40](#)
- J. Liu and J. Sun. Parallel graph-cuts by adaptive bottom-up merging. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2181–2188, June 2010. [29](#), [105](#)
- S Lobry, F Tupin, and L Denis. A decomposition model for scatterers change detection in multi-temporal series of sar images. In *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*, pages 3362–3365. IEEE, 2016. [129](#)
- P. Lombardo and C.J. Oliver. Maximum likelihood approach to the detection of changes between multitemporal SAR images. *IEEE Proceedings-Radar, Sonar and Navigation*, 148(4):200–210, 2001. [92](#), [129](#)
- A. Lopes, E. Nezry, Goze, R. Touzi, and A. Solaas. Adaptive processing of multilook complex SAR images. In *Geoscience and Remote Sensing Symposium, 1992. IGARSS '92. International*, volume 2, pages 890–892, May 1992. [91](#)
- A. Lopes, E. Nezry, R. Touzi, and H. Laur. Structure detection and statistical adaptive speckle filtering in SAR images. *International Journal of Remote Sensing*, 14(9):1735–1758, 1993. [XXVI](#), [91](#), [113](#), [114](#), [116](#), [118](#)
- G. Moser and S. B. Serpico. Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10):2972–2982, 2006. [92](#)
- F. Palsson, J. R. Sveinsson, M. Ulfarsson, and J. Benediktsson. SAR image denoising using total variation based regularization with SURE-based optimization of the regularization parameter. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, pages 2160–2163, July 2012. [90](#)
- S. Parrilli, M. Poderico, C.V. Angelino, and L. Verdoliva. A Nonlocal SAR Image Denoising Algorithm Based on LLMSE Wavelet Shrinkage. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2), 2012. [90](#)
- F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, and P. Larzabal. Covariance structure maximum-likelihood estimates in compound gaussian noise: Existence and algorithm analysis. *Signal Processing, IEEE Transactions on*, 56(1):34–48, 2008. [91](#)
- T. Perciano et al. 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, 2016. [75](#)
- P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *pattern analysis and machine intelligence, IEEE Transactions on*, 12(7):629–639, 1990. [40](#)

- T. Pock, T. Schoenemann, G. Graber, H. Bischof, and D. Cremers. A convex formulation of continuous multi-label problems. *Computer Vision–ECCV 2008*, pages 792–805, 2008. [101](#), [111](#)
- L. Potter, E. Ertin, J. Parker, and M. Cetin. Sparsity and compressed sensing in radar imaging. *Proceedings of the IEEE*, 98(6):1006–1020, 2010. [92](#)
- G. Ramillien, J. Famiglietti, and J. Wahr. Detection of continental hydrology and glaciology signals from grace: a review. *Surveys in Geophysics*, 29(4-5):361–374, 2008. [IX](#), [1](#)
- J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978. [41](#)
- F. Rossant et al. A morphological approach for vessel segmentation in eye fundus images, with quantitative evaluation. *Journal of Medical Imaging and Health Informatics*, 1(1):42–49, 2011. [XVIII](#), [71](#)
- L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992. [XXII](#), [27](#), [90](#)
- H. Samet. The quadtree and related hierarchical data structures. *ACM Comput. Surv.*, 16(2):187–260, jun 1984. [52](#)
- A. Shabou, J. Darbon, and F. Tupin. A Markovian Approach for InSAR Phase Reconstruction With Mixed Discrete and Continuous Optimization. *Geoscience and Remote Sensing Letters, IEEE*, 8(3):527–531, May 2011. [105](#)
- M. Silveira et al. Separation Between Water and Land in SAR Images Using Region-Based Level Sets. *IEEE Geoscience and Remote Sensing Letters*, 6(3):471–475, July 2009. [40](#), [41](#)
- A. Singh. Digital change detection techniques using remotely-sensed data. *International journal of remote sensing*, 10(6):989–1003, 1989. [91](#)
- G. Steidl and T. Teuber. Removing multiplicative noise by Douglas-Rachford splitting methods. *Journal of Mathematical Imaging and Vision*, 36(2):168–184, 2010. [90](#)
- X. Su, C.-A. Deledalle, F. Tupin, and H. Sun. NORCAMA: Change Analysis in SAR Time Series by Likelihood Ratio Change Matrix Clustering. *ISPRS Journal of Photogrammetry and Remote Sensing*, pages 247–261, May 2014. [XXVIII](#), [92](#), [133](#), [134](#), [135](#)
- X. Su, C.-A. Deledalle, F. Tupin, and H. Sun. Two-step multitemporal nonlocal means for synthetic aperture radar images. *Geoscience and Remote Sensing, IEEE Transactions on*, 52(10):6181–6196, Oct 2014. [133](#)
- G. Sullivan and R. Baker. Efficient quadtree coding of images and video. *Image Processing, IEEE Transactions on*, 3(3):327–331, 1994. [52](#)

- B. Tapley and C. Reigber. GRACE: a satellite-to-satellite tracking geopotential mapping mission. *Bollettino di Geofisica Teorica ed Applicata*, 40(3-4), 1999. [1](#)
- C. Tison et al. 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, 2004. [16](#)
- R. Touzi, A. Lopes, and P. Bousquet. A statistical and geometrical edge detector for SAR images. *geoscience and remote sensing, IEEE Transactions on*, 26(6):764–773, 1988. [XIV](#), [XIX](#), [40](#), [73](#)
- R. Touzi. A review of speckle filtering in the context of estimation theory. *Geoscience and Remote Sensing, IEEE Transactions on*, 40(11):2392–2404, 2002. [91](#)
- J. Tropp et al. Just relax: Convex programming methods for identifying sparse signals in noise. *Information Theory, IEEE Transactions on*, 52(3):1030–1051, 2006. [114](#)
- F. Tupin, H. Maitre, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky. Detection of linear features in SAR images: application to road network extraction. *Geoscience and Remote Sensing, IEEE Transactions on*, 36(2):434–453, 1998. [XI](#), [XVIII](#), [XX](#), [3](#), [71](#), [72](#), [73](#), [74](#), [75](#)
- K. Varshney, M. Cetin, J. Fisher III, and A. Willsky. Sparse representation in structured dictionaries with application to synthetic aperture radar. *Signal Processing, IEEE Transactions on*, 56(8):3548–3561, 2008. [92](#)
- G. Vasile, E. Trouvé, and J. Lee. Intensity-Driven Adaptive-Neighborhood Technique for Polarimetric and Interferometric SAR Parameters Estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6), 2006. [XXII](#), [90](#)
- G. Vasile, J.-P. Ovarlez, F. Pascal, and C. Tison. Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(4):1809–1826, 2010. [91](#)
- F. Weissgerber. *Coherent processing of multi-mode multi-resolution SAR images for the characterization of urban environment*. Theses, Télécom ParisTech, November 2016. [16](#)
- G.-S. Xia, G. Liu, W. Yang, and L. Zhang. Meaningful object segmentation from SAR images via a multiscale nonlocal active contour model. *IEEE Transactions on Geoscience and Remote Sensing*, 54(3):1860–1873, 2016. [41](#)
- H. Xie, L. E. Pierce, and F. T. Ulaby. Statistical properties of logarithmically transformed speckle. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3):721–727, Mar 2002. [15](#)
- H. Xie, L.E. Pierce, and F.T. Ulaby. SAR speckle reduction using wavelet denoising and Markov random field modeling. *IEEE Transactions on Geoscience and Remote Sensing*, 40(10):2196–2212, January 2002. [90](#)

Y. Yamaguchi, A. Sato, W.-M. Boerner, R. Sato, and H. Yamada. Four-component scattering power decomposition with rotation of coherency matrix. *Geoscience and Remote Sensing, IEEE Transactions on*, 49(6):2251–2258, 2011. [92](#)

W. Yin, D. Goldfarb, and S. Osher. Image Cartoon-Texture Decomposition and Feature Selection Using the Total Variation Regularized L^1 Functional. *LNCS*, 3752:73, 2005. [92](#)

H. Zhong, Y. Li, and L. Jiao. SAR Image Despeckling Using Bayesian Nonlocal Means Filter With Sigma Preselection. *IEEE Geoscience and Remote Sensing Letters*, 8(4), 2011. [90](#)

X. X. Zhu and R. Bamler. Tomographic SAR inversion by-norm regularization—The compressive sensing approach. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(10):3839–3846, 2010. [92](#)

Modèles Markoviens pour les images SAR: Application à la détection de l'eau dans les images satellitaires SWOT et analyse multi-temporelle de zones urbaines

Sylvain Lobry

RÉSUMÉ : Afin d'obtenir une meilleure couverture, à la fois spatiale et temporelle de leurs mesures les hydrologues utilisent des données spatiales en plus de celles acquises sur place. Fruit d'une collaboration entre les agences spatiales française (le CNES) et américaine (JPL, NASA), la future mission SWOT a notamment pour but de fournir des mesures de hauteur des surfaces d'eau continentales en utilisant l'interférométrie radar à synthèse d'ouverture (SAR). Dans cette thèse, nous nous intéressons au problème de la détection de l'eau dans les images d'amplitude SWOT qui est ici un prérequis au traitement interférométrique.

Dans cette optique, nous proposons d'utiliser une méthode dédiée à la détection des larges cours d'eau ainsi qu'un traitement spécifique pour la détection de rivières fines. La première méthode est basée sur un champ de Markov (MRF) pour la classification, conjointement à une estimation des paramètres de classes qui ne peuvent être supposés constants dans le cas de SWOT. L'estimation des paramètres peut également être modélisée par des champs de Markov. La seconde méthode s'appuie sur une détection de segments au niveau pixellique complétée par une connexion de ces segments.

Afin d'étudier l'extension aux séries multi-temporelles, nous proposons des méthodes de traitement adaptées aux données SAR de zones urbaines. Ces zones présentent de forts rétro-diffuseurs, ayant une radiométrie largement supérieure à celle des autres points dans l'image. Les modèles présentés prennent explicitement en compte la présence de ces forts rétro-diffuseurs en considérant les images comme une somme de deux composantes (le fond et les cibles fortes). Différents termes de régularisation peuvent alors être utilisés pour chacune de ces deux composantes. Modélisés comme des champs de Markov, ils peuvent alors être optimisés exactement par recherche de coupure minimale dans un graphe. Nous présentons des applications en détection de cibles fortes, régularisation et détection de changement dans ces séries.

MOTS-CLEFS: Traitement d'images, Télédétection, Imagerie satellitaire, Radar à synthèse d'ouverture (RSO), SWOT, Classification binaire, Détection d'eau, Régularisation, Détection de changement, Multi-temporel.

ABSTRACT: To obtain a better coverage both spatially and temporally, hydrologists use spaceborne data in addition to data acquired *in situ*. Resulting from a collaboration between NASA's Jet Propulsion Laboratory (JPL) and the French Space Agency (CNES), the upcoming SWOT mission will provide global continental water elevation measures using Synthetic Aperture Radar (SAR) interferometry. In this dissertation, we address the problem of water detection in SWOT amplitude images, which is to be performed before the interferometric processing.

To this end, we propose to use a method dedicated to the detection of large water bodies and a specific algorithm for the detection of narrow rivers. The first method is based on Markov Random Fields (MRF). The classification is regularized and the class parameters, which cannot be assumed constant in the case of SWOT, are jointly estimated. The second method is based on segment detection at the pixel level, completed by a connection step.

In order to study the extension to multi-temporal data, we propose methods adapted to the processing of series of SAR images of urban areas. These areas feature strong scatterers, having a radiometry orders of magnitude higher than the other points in the image. The proposed models explicitly account for the presence of these strong scatterers by considering the images as a sum of two components (the background and the strong scatterers). Different regularization terms can then be applied to each of these components. Modeled as MRF, they can then be optimized exactly using graph cuts. We present applications for strong scatterers detection, regularization and change detection.

KEY-WORDS: Image processing, Remote sensing, satellite imagery, Synthetic Aperture Radar (SAR), SWOT, Binary classification, Water detection, Regularization, Change detection, Multitemporal processing.

