Analyse des données lidar aéroportées à Retour d'Onde Complète pour la cartographie des milieux urbains

Clément Mallet

École Doctorale
d'Informatique,
Télécommunications
et Électronique de Paris

Thèse

présentée pour obtenir le grade de docteur

de Télécom ParisTech

Spécialité : Signal et Images

Clément MALLET

Analyse de données lidar à Retour d'Onde Complète pour la classification en milieu urbain

Soutenue le 26 novembre 2010 devant le jury composé de

Frédéric Jurie	Président
Xavier Briottet	Rapporteurs
Norbert Pfeifer	
Frédéric Bretar	Examinateurs
Uwe Soergel	
Florence Tupin	
Christian Heipke	Directeurs de thèse
Michel Roux	

Acknowledgements / Remerciements

This PhD thesis work has been carried in the Laboratoire MATIS of the Institut Géographique National, in conjunction with the TSI département of Télécom ParisTech and the Institute of Photogrammetry and GeoInformation (IPI) of the Leibniz University of Hanover, Germany.
First, I would like to acknowledge Christian Heipke. He accepted to supervise this work from the early beginning, and allowed me to benefit from visiting IPI each time I asked him. I also would like thank to Michel Roux for his efforts and his availibility, especially during the last months. Frédéric Bretar initiated me into airborne laser scanning research and offered me this great opportunity to make this work. I fell tremendously lucky to have been supervised by Uwe Soergel. I greatly appreciated his knowledge, his constant advices, and the numerous discussions we had everywhere we met. He managed to make him self-available whenever I needed. I did learn a lot, and I hope our cooperation will continue.

I thank Frédéric Jurie for accepting to preside the committee, and for its interest to my work. I would like to address my aknowledgements to Norbert Pfeifer and Xavier Briottet for their thorough analysis of this work. They deeply reviewed my thesis, and I appreciated their helpful comments. Thanks also to Florence Tupin for her relevant remarks.

Merci ensuite à Didier Boldo et Nicolas Paparoditis, chefs successifs du laboratoire MATIS pour les excellentes conditions de recherche au laboratoire et pour m'avoir laissé une totale liberté dans mes recherches, ainsi qu'à Marc Pierrot-Deseilligny. Un grand merci également à François Boyero et Marie-Claude Foubert, sans qui le laboratoire ne serait pas le même.

Je tiens ensuite à témoigner ma sincère amitié et mes remerciements aux membres du laboratoire avec qui j'ai pris (et prends toujours) beaucoup de plaisir à travailler depuis mon arrivée au MATIS: Nicolas David, pour tout ce qui est imaginable, Adrien Chauve, pour avoir souffert avec moi, Nesrine Chehata et enfin Florent Lafarge.
Merci également à Mathieu et Olivier, pour m'avoir conseillé ces trois dernières années et pour avoir relu ce manuscript, Bertrand, Benjamin et Sébastien, pour leur aide, Antonio et Thanasis qui m'ont supporté dans un environnement clos commun, tout comme Jérôme et Thomas, ainsi que tous les autres membres: Bahman, Jean-Pierre, Alex, Isabelle, Nico, M. Arnaud, Bruno, Adrien, Fabien, Lâmân, Karim, Jean-Pascal, Emilie, Mélanie D. & S., Daniela, Erwann, Corina, Lionel, Mahzad, Matthieu, David et Grégoire.

I also would like to deeply thank the IPI members for their warm welcome, especially the B222 fellows, Alex, Jens, Jan and Thorsten, as well as Petra and Joachim.

J'ai également une pensée émue pour Francis Schmitt et pour sa famille.

Enfin, je tiens à remercier mes famille et belle-famille pour leur intérêt au coloriage des

cartes. Par dessus tout, je ne serais pas le même sans Camille pour son soutien inestimable, et sans Cléo et son petit frère, qui me rappelent le vrai sens des choses.

Contents

Résumé étendu en français

1 Introduction

1.1 Télédétection des milieux urbains

Les dernières années ont permis de constater qu'avec la très grande avancée des techniques de télédétection, les informations disponibles sont beaucoup plus riches et l'analyse des milieux urbains peut désormais se faire à une échelle beaucoup plus fine. Des images satellites à une résolution proche de 10 cm sont disponibles. Les derniers capteurs hyperspectraux peuvent acquérir plusieurs centaines de bandes spectrales en même temps à des résolutions là aussi de plus en fine. Les Radars à Synthèse d'Ouverture sont eux très prisés en situations d'urgences ou quand une forte répétitivité est souhaitée. Le traitement de ces données a permis l'émergence d'un très grande nombre d'applications pour la génération de bases de données et en conséquence la cartographie automatique des paysages. En fonction des données utilisées, les bases de données ne sont pas à la même échelle et n'ont pas la même finalité : détection de changements, gestion des risques naturels, tourisme, navigation etc.

Dans un tel contexte, les systèmes laser aéroporté topographiques (ou lidar) présentent de nombreux avantages. Il s'agit d'une technologie active fournissant directement l'information de distance entre le capteurs et les cibles atteintes. Cette information est ensuite géoréférencée et toutes les mesures successives fournissent au final un nuage de points en trois dimensions. Ces nuages sont reconnus pour leur très grande précision altimétrique ($<$0,1 m), même s'ils échantillonnent la surface terrestre de manière plus irrégulière que les capteurs image. De plus, en fonction de la géométrie des objets illuminés par le faisceau laser, plusieurs rétrodiffusions peuvent être enregistrés pour une seule impulsion laser (par exemple, la canopée des arbres et le terrain sous-jacent). On parle de systèmes lidar multi-échos. Cette capacité a permis leur rapide adoption, dès les années 1990, pour l'étude des milieux forestiers, puis pour des cas plus spécifique comme la détection de ponts ou de lignes à haute tension. Les données lidar sont également très prisées dans toutes les applications nécessitant une représentation 3D de grande qualité altimétrique des surfaces acquises : reconstruction 3D de villes, Modèles Numériques de Terrain (MNT) ou de Surface (MNS) pour, par exemple, la gestion des zones à enjeux, la métrologie, ou les applications patrimoniales.

1.2 L'avènement du lidar aéroporté à retour d'onde complète

Les systèmes lidar traditionnels ne fournissent qu'une simple représentation géométrique des scènes sous forme de coordonnées tri-dimensionnelles. Les systèmes nommés **à retour d'onde complète** ont émergé il y a vingt ans environ et sont devenus très populaires les cinq dernières années avec l'apparition des systèmes topographiques commerciaux. Ces capteurs permettent d'enregistrer des signaux rétrodiffusés par chaque impulsion laser émise. Ces signaux sont appelés "ondes" (*waveform*) et sont numérisés à fréquence constante. Une telle séquence représente donc l'avancée de l'impulsion laser interagissant avec les surfaces réfléchissantes. On nomme alors "onde complète" (OC) chacun de ces ensembles quantifiés de données d'amplitude rétrodiffusée. La forme de ces ondes varie en fonction des objets atteints : cette variation est due à la fois à leurs propriétés géométriques et radiométriques. En conséquence, les données à retour d'onde complète (ROC) apportent des informations supplémentaires par rapport aux données multi-échos classiques. En lieu et place de nuages de points 3D faiblement structurés, les systèmes ROC fournissent une succession de profils 1D des scènes 3D, apportant des connaissances physiques additionnelles sur les objets atteints par le faisceau laser.

L'enregistrement de ces ondes permet leur analyse *a posteriori* de la phase d'acquisition, et donc le développement de traitements plus avancés que ceux implémentés dans les logiciels embarqués avec le capteur. Dans un premier temps, la détection des maxima d'amplitude dans les ondes, que l'on appelle échos, qui deviendront des points 3D après géoréférencement, peut être améliorée (nombre et position), en particulier en cas de structures complexes (végétation), mais également qualifiée. Dans un second temps, la forme des ondes, non disponible auparavant, peut être analysée pour retrouver des informations sur les propriétés physiques sur les objets atteints.

Les premiers systèmes ont été conçus pour des applications bathymétriques, puis, sous l'impulsion de la NASA, des systèmes topographiques expérimentaux, souvent satellitaires, ont vu le jour. Ils ont démontré la pertinence de l'enregistrement des ondes complètes pour caractériser les zones de végétation tropicales. Puis, en 2004 est apparu le système commercial petite empreinte (tâche au sol du faisceau laser inférieure à 1 m) LiteMapper 5600, développé sans application particulière. De nombreux travaux ont alors été initiés sur ce type de données, soit en zones forestières soit en zones urbaines, et se sont accrus les trois dernières années avec l'augmentation du nombreux de systèmes commerciaux ayant cette capacité de numérisation. L'intérêt de ce type de données est désormais démontré. Certains travaux de thèse récents ont développé des méthodes de traitements originales [Jut07; Lin09; Rei10], mais la contribution spécifique des donnés ROC pour la cartographie automatique des milieux urbains a été peu investigée jusqu'à maintenant.

1.3 Objectifs et limitations de la thèse

L'Institut Géographique National (IGN) a pour vocation de développer et d'entretenir le Référentiel à Grande Echelle (RGE), et parmi ses objectifs, se trouve la production, la mise à jour et l'enrichissement de la composition topographique du RGE (BD TOPO®). Dans ce contexte opérationnelle de cartographie à échelle nationale, l'IGN a acquis en 2007 un système lidar aéroporté à retour d'onde complète (Optech 3100). Le premier objectif est d'améliorer la qualité des MNT sur les zones à enjeux (inondations etc.) et de forêts, avec des précisions altimétrique et planimétrique supérieures à respectivement 0,3 m et 1 m.

Pour cela, une chaîne de production fondée sur le simple nuage de points 3D a été mise en place, et il a été démontré que des acquisitions avec de faibles densités de points ($\sim$ 2-3 points/m^2 par bande) sont suffisantes pour atteindre les qualités souhaitées

L'émergence des systèmes ROC a soulevé beaucoup de questions quant à leurs potentiel réel pour la cartographique à la fois urbaine et rurale. Que faire des informations physiques désormais disponibles en supplément des informations géométriques traditionnelles ?
L'objectif de la thèse est donc d'évaluer la contribution des données lidar à retour d'onde complète petite empreinte en zones urbaines. Cette tâche sera menée dans le cadre de la classification des zones d'occupation du sol, étape préliminaire à grands nombres de méthodes pour constituer, mettre à jour, qualifier et enrichir les bases de données topographiques. Une analyse similaire a été mené en parallèle de ce travail pour les zones de forêt.
Le processus mis en place doit prendre en entrée les signaux 1D lidar bruts (avec toutes les métadonnées nécessaires pour ensuite les géoréférencer), les traiter pour générer un nuage de points 3D, et enfin classer ce dernier en se servant des informations supplémentaires extraites durant l'étape de traitement du signal.

Puisque la problématique porte sur l'analyse des milieux urbains avec des données haute résolution, les ondes issus des systèmes petite empreinte seront utilisées. Cela signifie que la tâche au sol du faisceau laser est inférieure à 1 m, c'est-à-dire inférieure à la taille des objets d'intérêt (sauf pour la végétation). En conséquence, chaque écho de chaque onde correspond à une cible spécifique et son analyse morphologique peut donc apporter des informations physiques propres à cette dernière. Par ailleurs, notre problématique de travail porte uniquement sur la contribution des ROC vis-à-vis des données lidar multi-échos. Nous ne l'évaluerons pas par rapport à des données images, par exemple multi- ou hyper-spectrales, comme on pourrait légitimement s'y attendre dans un contexte de caractérisation de l'occupation du sol.
Trois classes d'intérêt ont été choisies : *sol*, *bâtiment* et *végétation*. La classe "*sol*" regroupe tous les types de sol ainsi que les objets du bas-sursol (*i.e.*, légèrement au-dessus du sol, par exemple, les voitures, les poteaux). La classe "*végétation*" comprend elle les arbres ainsi que la basse-végétation, comme les haies. Nous avons décidé de ne pas abordé le problème de discrimination des espèces d'arbres, des matériaux de toits et des types de surfaces de sol. En effet, nous pensons que la première est principalement fondée sur l'analyse de la répartition des points 3D (et donc non spécifique onde complète). A l'inverse, les deux derniers problèmes sont difficilement abordables avec la simple connaissance de la géométrie 3D et donc toute comparaison s'avèrerait fallacieuse.

L'objectif final de la thèse est la classification du nuage de points 3D issu des ondes lidar. Un grand nombre d'autres sujets d'intérêts relatifs aux données ROC ont donc été volontairement laissés de côté. Il s'agit de :

- la classification directe des signaux 1D, sans décomposition sous forme d'échos ;
- la simulation d'ondes lidar, dans un objectif d'analyse/synthèse, pour améliorer la compréhension de l'influence des propriétés physiques des objets ;
- la reconnaissance de forme sur les nuages 3D générés ou sur les données OC brutes (volume spatio-temporel), trop redondant avec la chaîne de reconstruction 3D semi-automatique de bâtiments déjà mise en place à l'IGN.

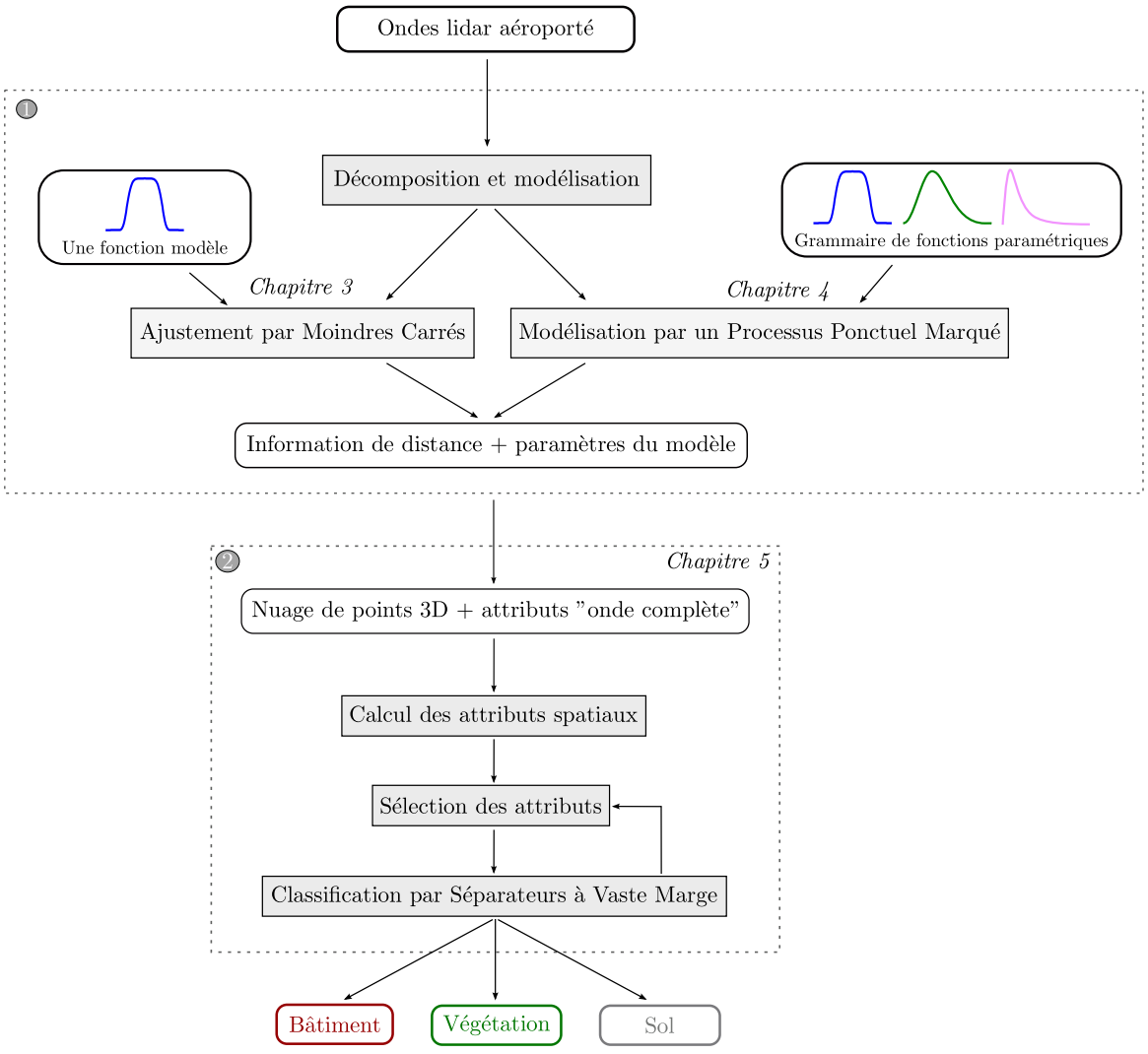


FIG. 1 – Vue globale de l'approche retenue.

1.4 Stratégie adoptée

La stratégie retenue est donc composée de deux étapes principales, se suivant :

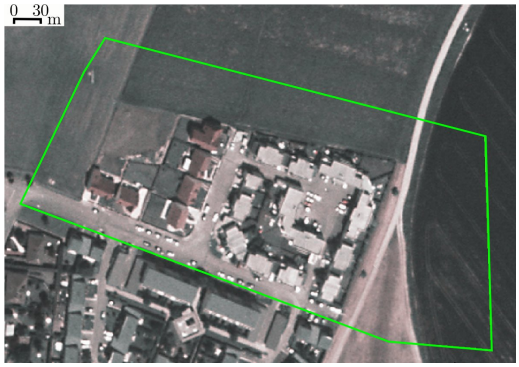
1. **Décomposition et modélisation des ondes.** La **décomposition** consiste à détecter pour chaque onde la position des maxima d'amplitude, chaque maximum (écho) correspondant à la rétrodiffusion d'un objet spécifique. L'étape de **modélisation** vise elle à caractériser la forme de chaque écho via l'estimation des paramètres d'une fonction analytique, ces paramètres étant ensuite introduits lors de l'étape de classification.
   Une première stratégie cherche à accomplir ces deux tâches successivement, une fonction modèle unique étant sélectionnée. Dans un second temps, une approche stochastique, fondé sur un processus ponctuel marqué, ne se limitant plus à un unique modèle, est mise en place : on cherche la meilleure fonction analytique dans une grammaire donnée.
2. **Classification.** Après géoréférencement des échos extraits, nous sommes en possession d'un nuage de points 3D avec des attributs morphologiques. Ces attributs sont fusionnés avec des attributs spatiaux (calculés sur la simple analyse géométrique locale autour de chaque point laser) et sont introduits dans un classifieur supervisé par Séparateurs à Vaste Marge (SVM). Afin de sélectionner les attributs les plus pertinents, mais également pour évaluer la contribution des caractéristiques morphologiques des échos vis à vis de celles spatiales, une étape de sélection d'attributs est ajoutée en amont de celle de classification.

2 Le lidar topographique à retour d'onde complète

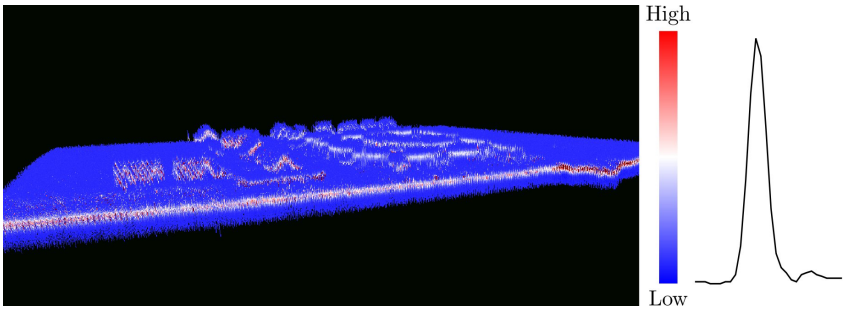
2.1 Motivation, principes et connaissances actuelles

En cas d'unique objet atteint par le faisceau laser, un capteur enregistrent un unique retour est suffisant. Cependant, même pour de petites empreintes au sol (de 0,2 à 2 m), plusieurs cibles peuvent être atteintes. Cela a motivé le développement de capteurs multi-échos, capables désormais d'enregistrer jusqu'à six retours pour une unique impulsion (zones de végétation par exemple). Pour ce type de capteur, la détection des échos (*ie* le passage d'un signal quantifié à un ensemble discret de valeurs de distance) s'effectue lors de l'acquisition par un algorithme temps-réel embarqué avec le système. De tels algorithmes sont inconnus pour les utilisateurs finaux, et en particulier leurs performances qui peuvent être critiques en cas de paysages complexes. Les inconvénients suivant ont été relevés : 1/ mauvaise estimation de la position des échos (et donc dégradation de la qualité altimétrique des points lidar, pouvant aller jusqu'à 0,3 m) ; 2/ omission d'échos dans les ondes (car les systèmes multi-échos sont limités en nombre d'échos par onde) ; 3/ non détection du deuxième écho en cas de superposition ; 4/ non détection des échos de faible amplitude (car les systèmes sont réglés pour détecter au dessus d'un certain seuil prédéfini). Ces raisons ont motivés le développement des capteurs ROC, avec l'ajout d'un numériseur et de capacité de stockage supplémentaire.

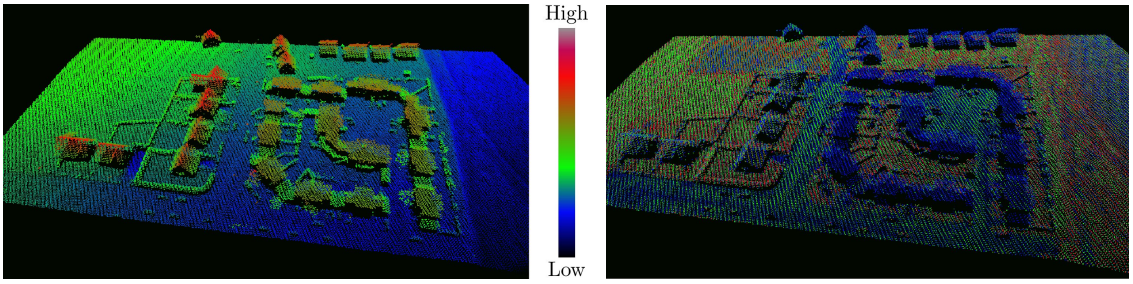
Ils sont apparus très récemment (2004), après des phases opérationnelles en contexte bathymétrique (années 80 et 90) puis expérimentale en contexte topographique pour l'analyse des zones d'occupation du sol et des forêts (dès 1994). Les données résultant de ces



(a) Orthoimage ©IGN.

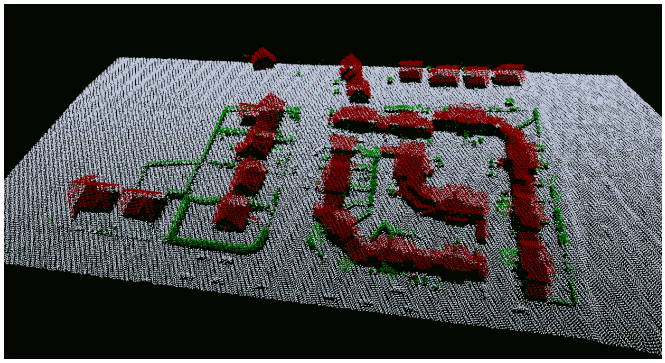


(b) Volume brut de données onde complète.



(c) Nuage de points 3D estimé (altitude).

(d) Nuage de points 3D estimé (amplitude).



(e) Classification : Bâtiment – Végétation – Sol.

FIG. 2 – (a) Illustration de la stratégie mise en place sur une zone d'intérêt. (b) Données d'entrée : signaux 1D, ici géoréférencés et où chaque échantillon de chaque onde est colorié selon son amplitude. (c) Ces signaux sont décomposés en échos, ce qui permet de générer un nuage de points 3D. (d) Ces échos sont ajustés avec une fonction modèle, fournissant des attributs de forme. (e) Ces attributs sont regroupés avec des attributs spatiaux puis sont introduits dans un classifieur supervisée SVM.

capteurs sont donc également très nouvelles et peu maîtrisées. La forme des ondes dépend à la fois des paysages survolés mais également des spécifications d'acquisition : plus le faisceau est divergent, plus on augmente la probabilité d'atteindre un grand nombre d'objets.

L'enregistrement des ondes lidar se fait avec un échantillonnage toutes les nanosecondes. Il débute dès que le signal rétrodiffusé dépasse un seuil prédéfini, au-delà d'une distance données (par exemple 100 m), et est limité en nombre d'échantillons (souvent entre 50 et 80). Si toutefois, à la fin d'une séquence, l'amplitude du signal est encore supérieure à un seuil prédéfini, une nouvelle séquence, toujours de taille fixe, peut être enregistrée. Les numériseurs onde complète n'enregistrent toutefois pas mieux les signaux de faible amplitude et donc, en cas de faible rapport Signal à Bruit et de configurations défavorables (fort angle d'incidence couplé à une forte pente de la cible atteinte), des ondes retour peuvent ne pas être enregistrées. Enfin, les signaux émis, appelés T0, ou du moins une proportion non connue, sont enregistrés pour chaque impulsion. Ils sont nécessaires pour connaître exactement le temps d'aller-retour du faisceau laser et pour mener à bien l'étape de correction des valeurs estimées d'amplitude et de largeur des échos.
Une telle numérisation présente de nombreux avantages, souvent liés à la possibilité de pouvoir, a posteriori de l'acquisition, traiter les signaux lidar. Un certain nombre d'inconvénients existent également : l'enregistrement des signaux entraîne une baisse de la cadence de tir, les volumes de données enregistrés sont beaucoup plus important (entre cinq et dix fois), et enfin, il n'existe pas à l'heure actuelle de solution logicielle permettant à un utilisateur final de gérer, visualiser et traiter les ondes lidar.

2.2 Traitements des ondes lidar

À l'heure actuelle, le traitement des ondes lidar consiste principalement en l'extraction (1) d'échos, qui géoréférencés donneront des points 3D, et (2) de paramètres morphologiques de ces échos qui serviront à l'étape de classification. Une partie de la littérature met donc l'accent sur des méthodes de détections d'échos avancées (pour trouver les d'échos de faible amplitude ou se recouvrant ainsi que pour estimer très finement la distance cible-capteur). Une autre partie se focalise sur la modélisation des échos par des fonctions paramétriques simples. L'hypothèse communément adoptée est qu'une onde est un mélange de fonctions gaussiennes. A l'étape de détection des échos succède donc une étape d'estimation des paramètres de la gaussienne (amplitude et largeur), effectuée soit par Moindres Carrés non linéaires, soit par Expectation-Maximization.
La connaissance des ondes émises permet elle deux types de traitements. En amont de la modélisation, certains auteurs effectuent une étape de déconvolution afin d'éliminer l'influence de la forme des ondes émises sur l'onde retour. Cette étape est d'autant plus nécessaire que la forme d'onde émise est spécifique (c'est-à-dire fortement asymétrique à droite). En aval, en modélisation également l'onde émise par une gaussienne, il est possible de normaliser les amplitudes et largeurs estimées par écho de celle de l'onde émise. L'objectif est là d'obtenir des valeurs homogènes entre zones d'acquisition et au sein même d'une acquisition. Il a été en effet noté qu'au cours du temps les pics émis ont tendance à s'élargir et à diminuer d'amplitude.
Plus de détails sur les traitements des ondes peuvent être trouvés dans [MB07].

2.3 Applications

Les applications fondées sur les données lidar ROC sont de deux types : celles qui utilisent le nuage de points 3D avec attributs de forme (amplitude+largeur) obtenu par modélisation, et celles qui utilisent directement les signaux 1D géoréférencés en trois dimensions. On renvoie là aussi le lecteur vers [MB07] pour de plus amples détails.
Les applications fondées "nuage de points 3D" concernent la classification des zones d'occupation du sol, l'étape de filtrage préalable à la génération de Modèles Numériques de Terrain, l'estimation des paramètres forestiers, la détection et la classification des arbres à la fois en milieux naturels et urbains. Les paramètres forestiers (hauteurs d'arbres etc.) sont estimés de manière plus fiable grâce à un plus grande nombre de points 3D détectés et à une meilleure qualité de détection. Les classifications en zones naturelles et urbaines tirent elles parti des attributs de forme pour segmenter un plus grand nombre de classes (surfaces de sol par exemple) ou pour améliorer les méthodes existantes (en particulier la détection de la végétation). Le filtrage des nuages de points et la classification des arbres bénéficient à la fois du plus grand nombre de points 3D détectés et des attributs de forme estimés. Ainsi, la largeur des échos est un indice fondamental pour discriminer les points 3D appartenant au sol de ceux situés sur des buissons bas dans le cadre de la génération de MNT. Les points 3D supplémentaires se situant eux souvent dans la canopée des arbres, ils permettent une meilleure description de la forme des arbres pour leur classification.
Les applications fondées "signaux 1D" sont de deux types. En milieux urbains, l'analyse d'ondes successives permet un traitement spatio-temporel du volume de données. Des images spatio-temporelles sont ainsi générées à plusieurs intervalles de distance. Les lignes détectées dans chacune des images sont ensuite accumulées en 3D pour retrouver des pans de toits complets. En zones de végétation, dans une approche analyse/synthèse, les paramètres forestiers sont estimés en comparant des ondes lidar réelles (souvent moyenne empreinte 10-25 m) à celles simulées à partir de modèles de forêts. Une étape d'inversion du modèle de transfert radiatif utilisé, plus ou moins complexe, est effectuée pour cela.

3 Données

Deux jeux de données sont disponibles pour notre étude. Comme annoncé en introduction, nous travaillerons avec des données issues de capteurs lidar à petite empreinte, *i.e.*, telles que les faisceaux laser ont une tâche au sol inférieure à 1m. Le premier jeu correspond à une unique bande laser acquise sur la ville de Biberach (Allemagne) avec le capteur Riegl LMS-Q560. La densité de points est proche de 5 points/m^2. Le deuxième jeu de données a été acquis sur la ville d'Amiens (France) avec le système Optech 3100 EA : 11 bandes laser avec une densité de 2 points/m^2 couvrent des zones urbaines plus ou moins denses. Toutefois, les données seront traitées bandes par bandes, sans bénéficier du recouvrement entre elles.
Le tableau 1 récapitule les principales spécifications des deux zones test.

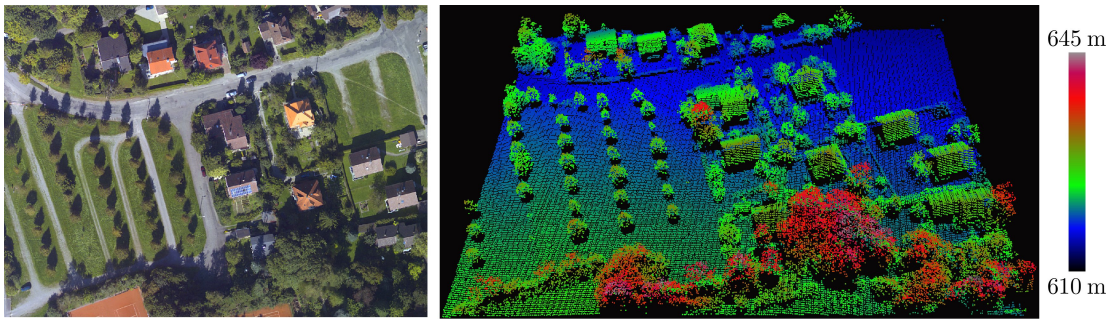
4 Traitement et modélisation des ondes lidar

L'extraction de points 3D à partir d'ondes lidar passe nécessairement par une étape de traitement du signal, désormais laissée à l'utilisateur final des données. La détection

Zone / Métadonnées	**Biberach**	**Amiens**
Date	Septembre 2006	Février 2008
Capteur	Riegl LMS-Q560	Optech 3100 EA
Surface (km^2)	1.1	20
Nombre de bandes	1	11
Nombre d'ondes	$2\ 342.10^3$	$70\ 135.10^3$
Hauteur de vol (m)	500	1000
Empreinte au sol (m)	0,25	0,8
Fréquence d'acquisition (kHz)	100	100
Largeur de l'impulsion laser (ns)	4	4
Echantillonnage des ondes (ns)	1	1
Densité d'impulsions ($/m^2$)	5	2
Projection cartographique	Gauss-Krueger Germany 9^o	Lambert 93

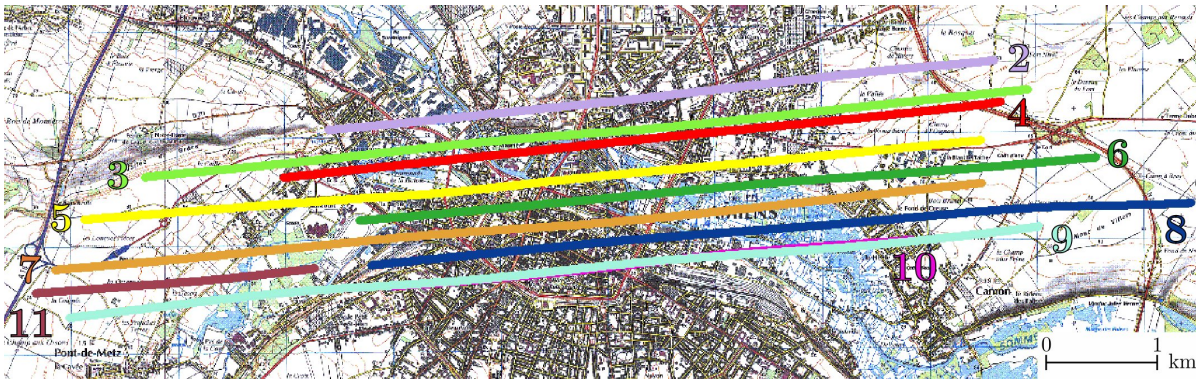
Tab. 1 – Spécifications des deux jeux de données

(a) Orthoimage sur la zone de Biberach.

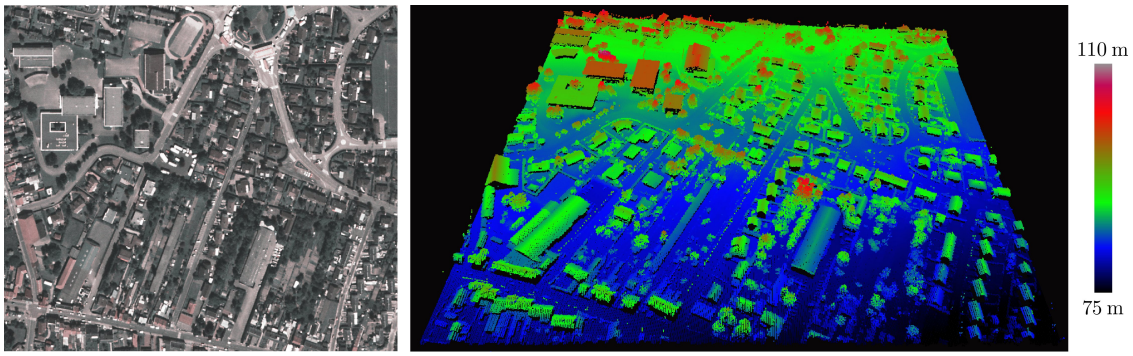


(b) Orthoimage et nuage de points 3D, colorié selon l'altitude des points sur la zone encadrée en rouge.

Fig. 3 – Jeu de données sur Biberach.



(a) Trajectoire de l'avion sur dix bandes laser acquises sur Amiens (Scan25 ©IGN)



(b) Orthoimage ©IGN et nuage de points 3D sur le centre-ville d'Amiens.

FIG. 4 – Jeu de données sur Amiens.

des maxima d'amplitude, que l'on appelle échos, dans chacune des ondes suffit à obtenir l'information de distance entre le capteur et les cibles et donc à générer un nuage de points 3D. Toutefois, dans notre objectif de classification de ce nuage, on cherche à confronter informations issues des ondes complètes et informations calculables uniquement à partir d'un nuage de points multi-échos standard. L'objectif de détection des échos se double de l'objectif de modélisation de ces derniers. La modélisation consiste à estimer la forme de chacun des échos via les paramètres d'une fonction analytique. Ces paramètres sont ensuite introduits dans l'étape de classification et de sélection d'attributs.
Les étapes de détection d'échos et de modélisation peuvent être menées conjointement ou séparément. Un grand nombre de méthodes de traitement du signal abordant le problème de mélange fini de modèles peuvent être adaptés à notre contexte, et deux approches distinctes sont proposées dans cette thèse. La première part de l'hypothèse communément admise dans la littérature lidar à retour d'onde complète que les échos des ondes sont bien modélisés par une fonction Gaussienne. La seconde considère que la reconstruction des signaux avec une unique fonction modèle symétrique n'est pas suffisante et on propose alors d'utiliser un processus ponctuel marqué pour ajuster les les échos des ondes lidar avec la meilleure fonction paramétrique prise dans une grammaire de modèles.

4.1 Décomposition et modélisation avec un modèle Gaussienne Généralisée

Il a été démontré dans la littérature qu'une très large majorité des échos des ondes lidar issues du capteur petite empreinte Riegl LMS-Q560 peuvent être modélisés par la fonction gaussienne. Celle-ci présente également l'avantage d'être simple (pas de problème pour estimer ses paramètres) et de fournir deux paramètres physiques, directement interprétables : l'amplitude et la largeur de l'écho. Notre propre analyse des ondes lidar issues à la fois des capteurs Riegl et Optech nous a montré qu'une telle hypothèse symétrique semblait valable. Toutefois, nous avons noté une légère distorsion des échos au sommet, apparaissant parfois plus pointus ou plus aplatis. Cela nous a conduit à proposer comme fonction modèle la Gaussienne Généralisée (GG), qui est une extension de la fonction gaussienne (cf. tableau 2). Son équation est la suivante :

$$f_{\mathrm{GG}}(x) = A \exp\left(-\frac{(x-\mu)^{\alpha^2}}{2w^2}\right) \tag{4.1}$$

A et w sont respectivement l'amplitude et la largeur des échos. μ est la position de l'écho. Le paramètre additionnel, appelé paramètre d'aplatissement est α et permet d'obtenir une plus grande flexibilité dans la modélisation.
La chaîne de traitement mise en place est la suivante :

1. le bruit de fond des ondes est éliminée puis ces dernières sont lissées par un filtre gaussien pour faciliter l'étape de détection des échos ;
2. la détection des échos se fait par une simple analyse du passage par zéro de la dérivée première du signal ;
3. l'algorithme de Levenberg-Marquardt est utilisée comme méthode d'estimation des paramètres de chacun des échos par Moindres Carrés Non Linéaires ;
4. on calcule les résidus entre les échantillons de l'ajustement obtenue et ceux de l'onde initiale :

- si au moins un résidu est supérieur à un seuil que l'on a défini, on ajoute un nouvel écho à la position du résidu maximal et on recommence les étapes 3 et 4
- sinon l'ajustement est terminé.

Comparée à la plupart des méthodes existantes, nous proposons la détection itérative des échos manquants via un ou plusieurs ajustements préalables du signal. Cela permet de détecter des échos dans des cas complexes (par exemple, deux échos se recouvrant) que ne peut résoudre une méthode simple de détection. En pratique, deux itérations suffisent, les ondes lidar n'étant pas très complexes.

Cette méthode d'ajustement itératif des ondes par Gaussienne Généralisée a été appliquée aux jeux de données de Biberach et d'Amiens. Les résultats sont évalués selon deux aspects : nombres d'échos détectés et qualité d'ajustement.
Tout d'abord, la méthode proposée permet de détecter 4,1% de points en plus pour Amiens, et 24% supplémentaires sur Biberach, si l'on compare aux nuages de points qui sont obtenus par une analyse temps-réel des ondes lidar (nuages fournis avec les données ROC). L'écart entre les deux zones s'explique par le fait que les conditions d'acquisition étaient différentes : vol "sans feuilles" pour Amiens et "avec feuilles" pour Biberach. Les ondes lidar d'Amiens sont moins complexes que celles de Biberach (ratio nombre d'échos par onde respectivement de 1,03 et 1,42 pour Amiens et Biberach). Cela représente tout de même près de 2,7M de points sur Amiens sur environ 63,8M d'ondes acquises. Cette différence se retrouve également dans le pourcentage de points additionnels détectés par notre approche itérative : 0,69% pour Amiens et 3,6% pour Biberach. Les échos supplémentaires sont principalement détectés dans les zones de végétation et sur les bords de bâtiments.
Dans un second temps, l'ajustement des ondes est évalué selon deux aspects : précision de détection des échos et qualité de reconstruction. Puisqu'il est très difficilement envisageable de procéder à une qualification exhaustive de notre détection des échos, une simple comparaison est faite avec les échos détectés en vol par le système lidar. Il s'agit d'une comparaison en distance suivant la direction du faisceau laser. Avec notre méthode, nous améliorons la détection des échos de 6 cm en moyenne, avec un écart-type moyen d'environ 38 cm pour Amiens. En pratique et après géoréférencement, cela correspond à générer des nuages de points avec une qualité altimétrique supérieure maximale de 15 cm et planimétrique supérieure maximale de 2,5 cm. Enfin, la qualité de reconstruction obtenues par notre méthode a été comparée à celles obtenues par une simple décomposition gaussienne et avec une décomposition GG mais sans détection itérative des échos. Les résultats montrent qu'à la fois l'approche itérative et l'ajout du paramètre d'aplatissement dans le modèle gaussien améliorent les résultats et valident l'approche proposée.

La décomposition des ondes par Gaussienne Généralisée permet d'extraire deux types d'attributs : les paramètres de la fonction (A, w et α) et les attributs que peuvent être calculés à partir de ces derniers. Il s'agit de la largeur à mi-hauteur (*Full-Width-at-Half-Maximum*), l'énergie de l'écho et la section rétrodiffusante σ (cf. figure 5. En particulier, σ est estimée avec une méthode de calibration fondée sur l'équation lidar. On se sert de cibles au sol de réflectance connue pour calculer σ. Des zones d'asphalte sont utilisées car elles représentent le seul candidat acceptable disponible sur tous les jeux de données malgré deux défauts principaux : faible réflectance et forte variabilité spectrale selon l'angle d'incidence.
Dans un objectif de classification, les histogrammes d'amplitude, de largeur, d'aplatissement et de section rétrodiffusante sont analysés pour les trois classes d'intérêt (sol, végéta-

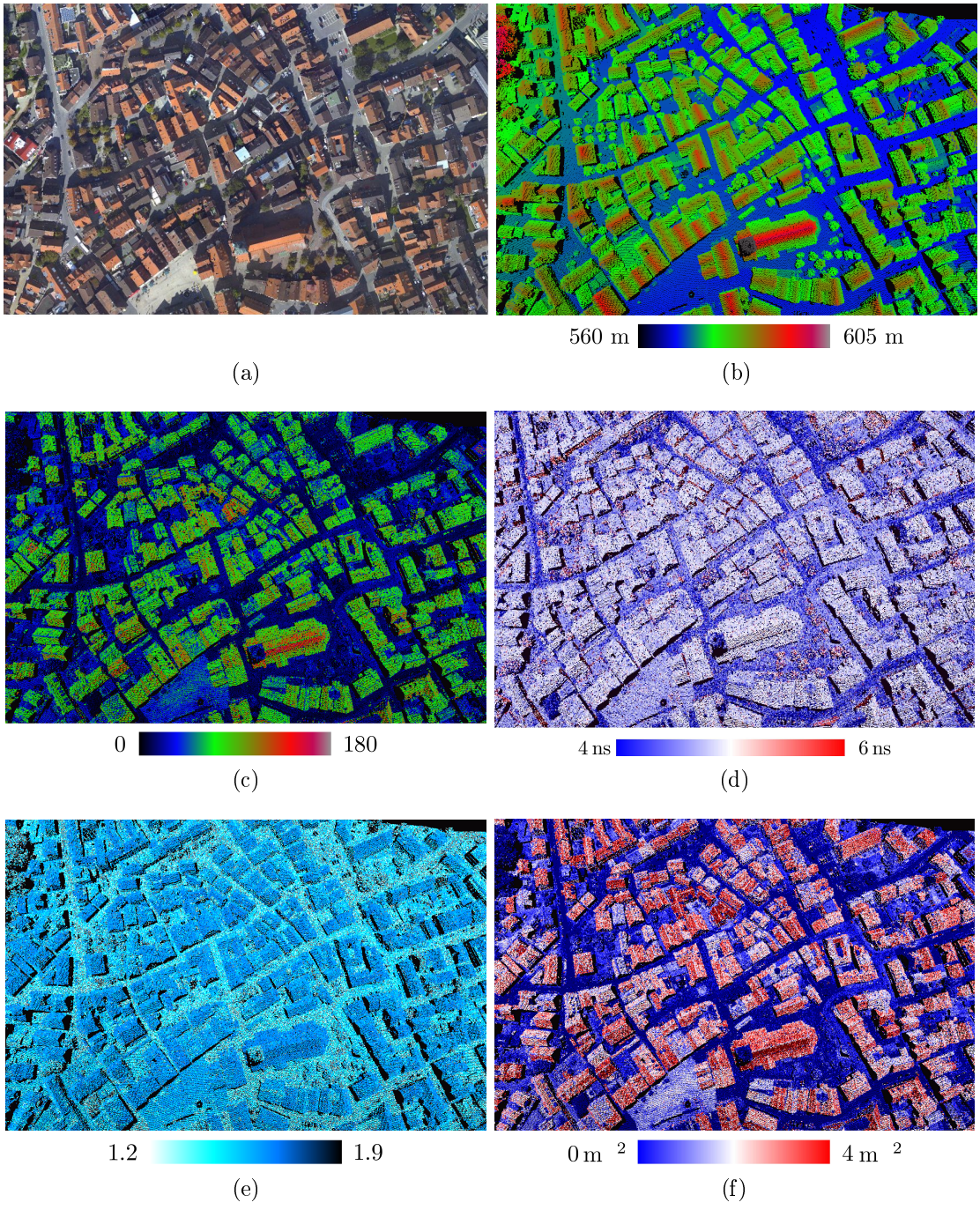


Fig. 5 – Visualisation des attributs ROC extraits d'une modélisation des ondes par Gaussienne Généralisée sur la ville de Biberach. (a) Orthoimage. Nuage de points 3D coloriée selon : (b) l'altitude, (c) l'amplitude, (d) la largeur, (e) l'aplatissement et (f) la section rétrodiffusante des échos.

tion, bâtiment), différents types de sol (asphalte, terre battue et herbe) et de matériaux de toits (métal, tuiles, ardoise). Ils montrent le fort potentiel de l'amplitude et de la section rétrodiffusante pour la segmentation des zones à faible réflectance (asphalte et métal) de celles à forte réflectance (terre battue et tuiles). Comme déjà présenté dans de nombreux articles, la largeur des échos semble discriminante pour les zones de végétation (largeur supérieure). Enfin, le paramètre d'aplatissement α ne montre pas un fort potentiel de discrimination parmi toutes les classes présentées. Il sera toutefois conservé dans l'étape de classification et de sélection d'attributs qui suit.

En conclusion, la méthode proposée s'est avérée très efficace. De plus, les améliorations proposées par rapport à la décomposition gaussienne traditionnelle ont permis la détection d'échos supplémentaires, l'amélioration de la qualité de reconstruction, l'amélioration de la qualité altimétrique des nuages de points 3D obtenus ainsi que l'extraction de paramètres morphologiques supplémentaires pour notre objectif final de classification. L'approche proposée n'est pas spécifique aux milieux urbains ni aux capteurs ROC petite empreinte. Elle peut être adaptée pour d'autres applications en changeant simplement la fonction modèle, tant que celle-ci reste estimable par Moindres Carrés. par ailleurs, il semble suffisant de jouer sur le niveau de bruit, le seuil de détection d'échos supplémentaires ainsi que sur le lissage de l'onde en entrée pour traiter des données issus de capteurs terrestres ou spatiaux. La qualité de reconstruction des ondes a été jugée satisfaisante avec le modèle GG. Cependant, une telle solution est trop restrictive. En effet, l'analyse de qualité d'ajustement a montré une certaine variabilité sur de grandes zones, y compris pour des ondes unimodales. Il semble donc qu'un certain nombre d'échos dans les ondes aient un comportement asymétrique et qu'il faille les ajuster avec une/des fonction(s) modèle spécifique(s). Enfin, dans un objectif de classification, il peut être intéressant de tester de nouvelles fonctions avec des paramètres distincts de ceux de la gaussienne. En conséquence, nous avons développé une approche originale qui prend en compte ces deux points d'amélioration.

4.2 Une approche par processus ponctuel marqué

4.2.1 Problème, approche et grammaire proposées

La deuxième approche proposée a pour objectif d'ajuster spécifiquement les échos asymétriques par une ou des fonctions autres que le modèle gaussien. Cela concerne certains échos d'ondes lidar et l'on ne sait pas a priori desquels il s'agit. Cette asymétrie apparaît dans plusieurs cas. Tout d'abord, la géométrie des cibles peut déformer les échos dans le cas de données petite empreinte : pour les zones planes, la rugosité des cibles peut entraîner un dissymétrie à droite comme à gauche, alors que pour les zones de végétation, la présence de plusieurs objets proches conduit souvent à une dissymétrie à droite, c'est-à-dire que l'énergie rétrodiffusée va en diminuant avec la pénétration dans la canopée. Etant donné que le modèle gaussien généralisé convient tout de même à une grande majorité des échos d'ondes petite empreinte, notre méthode doit également prendre en entrée ce modèle, en supplément du ou des modèles asymétriques choisis. L'avantage de mettre en place une telle méthode est également de pouvoir rester facilement différentes fonctions, fournissant ainsi différents jeux de paramètres qui pourront être ensuite intégrés et évaluer dans l'étape de classification. Mieux ajuster les ondes c'est enfin mieux les représenter et c'est donc proposer une meilleure méthode de compression des données ROC. Vu le volume supplémentaire de données issues de la numérisation des ondes, il s'agit d'un enjeu majeur

quoique peu abordé dans la littérature.

Il n'existe pas de méthode existante capable de résoudre notre double problème de sélection de modèle dans un mélange de composition inconnue. Il est possible d'envisager des méthodes de type Colonies de fourmis ou algorithmes génétiques mais notre choix s'est porté vers les méthodes stochastiques fondés sur les processus ponctuels marqués, qui ont montré un fort potential dans de nombreux problèmes de télédétection 2D ou 3D. Le modèle proposé présente alors un certain nombre d'avantages par rapport aux méthodes fondés sur le mélange de gaussiennes :

- Gestion de plusieurs modèles en même temps : la méthode prend en entrée une grammaire de fonctions. L'estimation de leur paramètres ne nécessite que l'évaluation des fonctions en tout point du signal. Toute fonction analytique est donc envisageable.
- Intégration de connaissances physiques sur le lidar : l'énergie de la formulation du modèle stochastique peut inclure des caractéristiques lidar physiques connues ainsi que des interactions entre échos voisins, permettant ainsi de bien gérer les ondes complexes.
- Exploration efficace de l'espace des configurations : un échantillonneur de Monte-Carlo par Chaîne de Markov couplé à des noyaux de propositions pertinents évite une recherche exhaustive dans de grands espaces de configurations. Cela est particulièrement intéressant quand le nombres d'objets est inconnu.

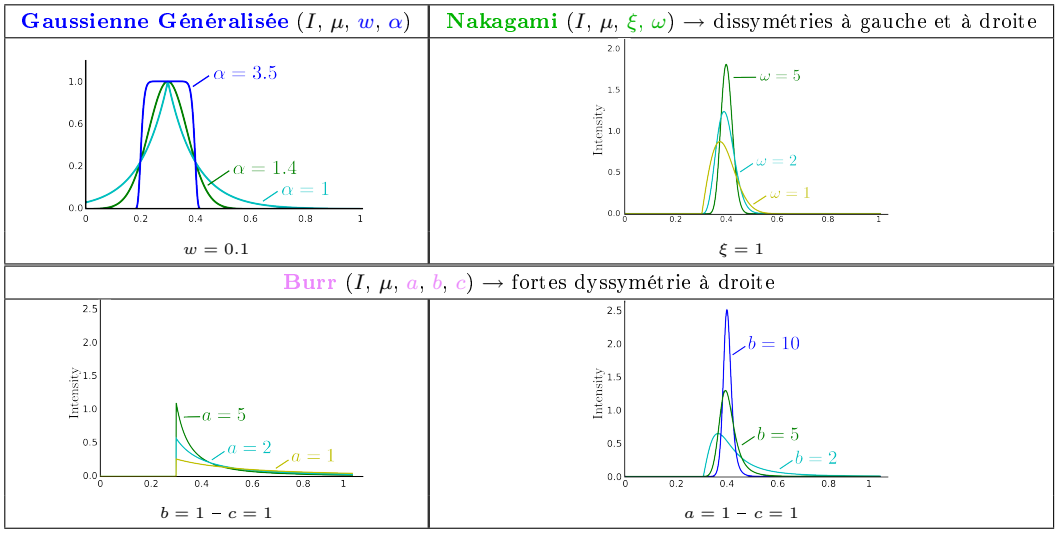
Il existe trois stratégies pour construire la grammaire de fonctions paramétriques. On a tout d'abord chercher à trouver une unique fonction capable de simuler tous les cas existants (symétrique, faible dissymétrie à gauche, plus ou moins forte à droite) mais la seule solution convenable (fonction Levy-skew α stable) ne possède pas de formulation analytique. Ensuite, on peut vouloir intégrer un grand nombre de fonctions simples (entre 5 et 10) pour couvrir tout le champs des possibilités. Toutefois, l'étape de sélection de modèles devient plus complexe et devant la faible variété de forme des échos, les résultats peuvent devenir difficile à qualifier. La solution s'est donc portée sur une grammaire minimale (2 à 4 fonctions) afin de rendre le processus plus robuste et de limiter les temps de calcul.
Trois fonctions ont donc été sélectionnées (cf. tableau 2) parmi un grand nombre de fonctions testées (Weibull, Gamma Généralisée etc.) :

- Le modèle Gaussien Généralisée est conservé (trois paramètres : A, w et α) car il convient à une grande majorité d'échos.
- La fonction Nakagami est choisie pour ajuster les échos dissymétriques d'un côté comme de l'autre. Elle possède trois paramètres : A, ω un paramètre d'asymétrie contrôlant la décroissance de l'écho et ξ un paramètre de forme contrôlant sa largeur.
- Enfin, pour les échos à la queue plus lourde (dissymétrie forte à droite), le modèle Burr est ajouté. Il possède quatre paramètres : A, deux paramètres de forme b et c ainsi qu'un paramètre d'échelle a.

La méthode proposée permet à la grammaire d'évoluer en fonction d'autres applications et d'autres formes de signaux.

4.2.2 Formulation énergétique du problème

Nous ne détaillerons pas dans ce résumé la théorie des processus ponctuels marqués. Nous renvoyons le lecteur vers [Lie00; Des11]. Notre cas dépasse le cas de comptage d'objets puisqu'il y a également une sélection de modèle à effectuer. On parle alors de processus



TAB. 2 – Grammaire de fonctions modèle sélectionnée, leurs paramètres et quelques comportements de celles-ci ($I = 1$ et $\mu = 0,3$ dans tous les cas).

ponctuel marqué multiple. Dans un cadre particulier, le problème de trouver la meilleure configuration pour notre problème dans un certain espace de recherche revient à minimiser une énergie donnée. On modélise donc chaque onde par une configuration d'objets, nos objets étant des fonctions paramétriques. Le modèle énergétique proposé permet d'évaluer ces configurations. La performance de notre méthode d'ajustement d'ondes lidar se fonde donc principalement sur la construction d'une énergie adaptée à notre problème. Cette énergie est composée d'un terme d'attache aux données (norme L_2 entre l'onde lidar et la configuration proposée) et un terme de régularisation qui permet d'introduire des connaissances a priori mais également des interactions entre objets d'une configuration. Le but est ainsi de favoriser ou de pénaliser certaines configurations possibles. Dans notre cas, il s'agit de limitations physiques liées à la rétrodiffusion de l'impulsion laser. Ces limitations sont modélisées par trois termes distincts :

- Limitation du nombre d'objets : un objet correspond à un écho. On cherche donc à favoriser les configurations ayant un nombre d'échos réaliste (jusqu'à 7), et en particulier celles avec un très faible nombre d'échos (entre 1 et 3).
- Limitation sur l'énergie rétrodiffusée : la connaissance de l'énergie de l'impulsion laser émise nous permet d'interdire toutes les configurations dont l'énergie n'est pas réaliste.
- Complexité de reconstruction : ce terme cherche à pénaliser les configurations dont deux échos sont plus proches que la résolution en distance du capteur lidar (souvent 2 ns c'est-à-dire deux échantillons). Cela favorise encore un reconstruction de signaux avec un faible nombre d'objets.

Trouver la configuration qui minimise l'énergie proposée est un problème d'optimisation non convexe dans un espace de grande dimension lui-même union de sous-espaces de dimensions variables. En effet, le nombre d'échos est inconnu et les trois fonctions choisies n'ont pas toutes le même nombre de paramètres. Une solution communément adaptée consiste à choisir un échantillonneur de Monte-Carlo par Chaîne de Markov à Sauts Réversibles (RJMCMC). Il s'agit d'un processus itératif : à partir d'une configuration initiale quelconque, une perturbation est proposée puis est acceptée ou refusée. L'échantillonneur ef-

fectue donc des sauts entre espaces de dimensions différents et ces sauts sont guidées par trois familles de mouvement appelés noyaux de proposition. L'acceptation ou le refus s'effectue selon l'algorithme de Metropolis-Hastings-Green.Le RJMCMC est couplé à un processus de relaxation, ici un recuit simulé, afin de converger vers la solution désirée. Cela permet d'atteindre sous certains conditions légèrement violées ici le minimum globale de l'énergie proposée quelque soit la configuration initiale proposée.

Le modèle proposé a été testé sur des données simulées, des ondes lidar large empreinte et sur les ondes des deux jeux test (petite empreinte, cf. Figure 6).
Les données simulées nous ont permis de qualifier le modèle mis en place et ses différents termes. Il a été ainsi vérifié le comportement de l'approche en cas de faible Rapport Signal à Bruit (moins bonne estimation des paramètres des échos), pour différents cas de recouvrement entre échos (tous bien gérés), en cas de mélanges de fonctions (sélection du bon modèle jusqu'à un taux de recouvrement critique). Tous les échos sont toujours détectés et leur position est précisément localisée. Il n'y a donc pas eu de perte de qualité quant aux nuages de points 3D générés par cette méthode.
Ensuite, les ondes issues des capteurs topographique moyenne et grande empreinte SLICER et LVIS ont été traitées. Les résultats montrent que la grammaire proposée permet relativement d'ajuster ces signaux plus complexes que ceux des capteurs petite empreinte. Dans ce cadre, les fonctions Nakagami et Burr sont privilégiées de manière équivalente par rapport au modèle GG. Les ondes traitées concernent en effet des zones de végétation tropicale dont la distribution altimétrique des cibles n'a pas de raison d'être gaussienne. Ainsi, l'atténuation du signal par le haut de la canopée est bien modélisé par la fonction Burr, alors que la Nakagami est utile pour le dernier "écho" du signal couvrant végétation basse et sol.
Enfin, les zones d'Amiens et de Biberach ont été traitées. Une analyse visuelle des ondes nous montre que l'ajustement est correct et que le processus itératif a bien le comportement souhaité : la première partie du processus cherche à se focaliser sur les modes du signal (détection d'écho) alors que la deuxième partie procède à l'ajustement local des échos avec la sélection du meilleur modèle et du jeu de paramètres optimal. Pour effectuer une analyse détaillée, six zones d'intérêt urbaines ont été choisies (bâtiments, herbe, cathédrale, zones mixtes etc.). Les principales conclusions sont :

- La détection d'échos est satisfaisante puisque nous obtenons des résultats similaires à ceux de la modélisation GG simple. Il n'y a pas de sur-ajustement ni de fausse détection d'échos.
- L'ajustement est de très bonne qualité et est amélioré par rapport à une modélisation GG simple.
- Le modèle Gaussien Généralisé est sélectionné pour plus de 80% des échos en zones urbaines. Il s'agit des zones planes (sol et pans de toits).
- Pour les zones de végétation, les trois modèles sont sélectionnés de manière équitables.
- Des échos asymétriques sont détectés sur des bâtiments, principalement quand leur géométrie se complexifie : faîtes de toits ou bordures. Il est également noté que le comportement symétrique ou non des échos est également lié à la radiométrie des objets puisque les échos de faible amplitude (*i.e.*, de réflectance plus faible) affichent une dissymétrie supérieure.

En résumé, les résultats montrent qu'il existe bien un certain nombre d'échos qui ne sont pas de forme symétrique mais qu'il en existe une assez faible proportion, et surtout que ces formes ne sont pas caractéristiques d'objets en particulier puisqu'il existe une influence

corrélée entre géométrie et radiométrie des cibles.

4.2.3 Conclusions et perspectives

En conclusion, l'approche originale proposée s'est avérée flexible à différents types de signaux et bien adaptée à notre double problème de détection d'échos et de sélection de la fonction paramétrique ajustant le mieux chacun des échos. On améliore ainsi la qualité d'ajustement des ondes et donc la compression des données ROC. On extrait également de nouveaux paramètres de forme pour les échos modélisés par la Nakagami et la Burr.
La principale limitation de l'approche stochastique concerne les temps de calculs. Les 70M d'ondes d'Amiens sont traitées en 58 jours contre 3,5 heures pour la décomposition en Gaussienne Généralisée. Les premières perspectives d'évolution de la méthode doivent donc se focaliser sur une amélioration du temps de traitement pour rendre le processus ponctuel marqué plus efficace à grande échelle. En supplément de la nécessaire amélioration des codes sources développés et de l'utilisation optimale des ressources des ordinateurs désormais disponibles, il paraît intéressant d'essayer de fusionner les deux méthodes de traitement des ondes proposées, c'est-à-dire de coupler ajustements local et global. Un ajustement par décomposition GG peut débuter ce processus puis un traitement spécifique pour les échos dont l'ajustement n'est pas satisfaisant, fondé sur l'approche stochastique pourrait être proposé. Il permettrait de démêler les cas de fonction modèle non adapté des cas de sous-détection d'échos.
Une dernière perspective de travail concerne le traitement des ondes non plus une par une mais de manière consécutive en gardant une cohérence temporelle, et donc spatiale. Des méthodes de type filtrage particulaire semblent bien adaptées à notre problème.

5 Classification de zones urbaines

5.1 Méthode proposée

Dans la dernière étape de notre processus, nous cherchons à classer les nuages de points obtenus par les étapes de décomposition des ondes décrites précédemment en bâtiment, végétation et sol. De plus, nous souhaitons évaluer la contribution des attributs issus des deux méthodes de modélisation (que nous appellerons attributs ROC) vis-à-vis d'attributs calculables uniquement à partir de systèmes lidar multi-échos (c'est-à-dire avec la connaissance unique de la géométrie 3D des points et de leur numéro d'écho). Nous cherchons à savoir si la numérisation des ondes lidar et les traitements supplémentaires nécessaires permettent une amélioration de la qualité de classification. Il existe un grand nombre d'articles traitant de la classification de données lidar à partir d'attributs ROC mais très peu se sont également penchés sur l'évaluation de leur pertinence dans le cas d'un grand nombre d'attributs lidar à disposition.
La classification s'effectue en 3D point par point. La stratégie adoptée est la suivante :

1. Calcul d'attributs fondés sur le nuage de points multi-échos.
2. Synthèse des attributs ROC obtenus par les deux méthodes décrites auparavant.
3. Classement groupé de ces deux types d'attributs par trois méthodes de sélection d'attributs.

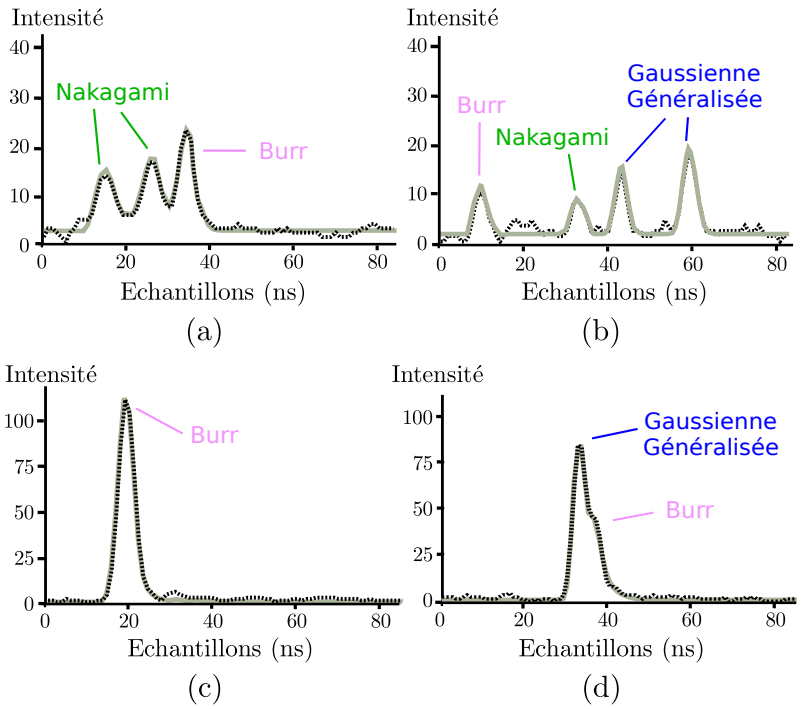


Fig. 6 – Décomposition et modélisation de quelques ondes par utilisation d'un processus ponctuel marqué. Une fonction correspond à un écho détecté (zone de Biberach). **(a-b)** Arbres. **(c)** Toit de bâtiment. **(d)** Buisson.

4. Classification supervisée des nuages de points lidar par Séparateurs à Vaste Marge (SVM) : différentes stratégies fondées sur les différents classements obtenues juste avant sont testées.

Les attributs sélectionnés sont décrits dans la section suivante.
La méthode proposée est flexible car l'ajout de nouvelles fonctions modèles dans le processus de modélisation des ondes lidar ne modifie pas l'étape de classification. De même, elle peut facilement être adaptée à d'autres problèmes de classification, puisqu'il suffit de modifier l'étape d'apprentissage du classifier SVM (c'est-à-dire la vérité terrain).

Trois méthodes de sélection d'attributs ont été sélectionnées dans la très grande littérature existante sur le sujet. Nous les avons choisies car elles sont facilement reproductibles, supervisées, et se fondent sur des caractéristiques différentes. Il s'agit du F-score, de ReliefF et de SVM-RFE (Recursive Feature Elimination). F-score et ReliefF mesurent directement le pouvoir de discrimination de chacun des attributs pour un problème de classification donnée. le classement des attributs se fait donc selon cette mesure. SVM-RFE utilise elle un classifieur SVM pour éliminer de manière récursive l'attribut qui contribue le moins à la décision d'étiquetage. Elle ne fournit donc qu'un classement.

Le choix des Séparateurs à Vaste Marge comme classifieur supervisée dans la large palette des choix possibles est assez objectif. Il s'agit d'une méthode non paramétrique qui possède une très grande capacité de généralisation (*i.e.*, qui ne nécessite pas une phase d'apprentissage exhaustive). Elle est également simple puisqu'il suffit de fournir un vecteur d'attribut et qui fonctionne très bien pour des données de grande dimension. Nous sommes dans ce cas-là puisqu'en entrée du SVM nous avons des nuages de points de plusieurs mil-

lions de points auxquels nous allons attacher un vecteur de 27 attributs. D'autres classifieurs tels que les Forêts Aléatoires et les Arbres de Classification et de Régression possèdent des propriétés équivalentes et sont également pertinents pour mener à bien cette tâche.
Nous ne détaillerons pas la théorie des SVM. Signalons simplement que notre choix s'est porté sur un noyau gaussien simple, l'estimation des deux paramètres du SVM se fait par validation croisée (largeur du noyau + facteur de pénalité C car nous nous trouvons dans un cas avec marges relaxées), et enfin puisque nous sommes dans un cas à quatre classes, la décision finale d'étiquetage se fait en "un contre un".

5.2 Attributs

Chaque point lidar 3D est associé à un vecteur de vingt-sept attributs. Ces attributs sont de trois types : spatiaux, fondés écho et ROC. Les attributs spatiaux et fondés écho sont calculables avec la simple connaissance de nuages de points multi-échos alors que les attributs ROC proviennent des deux méthodes de modélisation des ondes décrites auparavant. Les attributs non ROC sélectionnés ont été choisis pour deux raisons principales : (1) ils ont déjà montré leur pertinence pour la discrimination d'une ou plusieurs classes de notre problème et on souhaite les confronter aux attributs ROC extraits ; (2) ils nous ont semblé pertinents et étant donné que l'on vise une évaluation objective des attributs ROC, il semble en effet intéressant de les confronter au plus grand nombre d'informations calculables à partir de nuages de points lidar standard. Nous n'avons toutefois pas inclus tous les attributs existants de la littérature (par exemple, ceux de texture altimétrique).

Les attributs spatiaux (ou géométriques) sont au nombre de 15. Ils sont fondés sur la simple connaissance de la position 3D de chaque point, mais sont améliorés par le traitement des ondes lidar. On les classes en trois sous-catégories :

- Attributs altimétriques (3) : fondés sur l'altitude du point courant (par exemple, hauteur du point par rapport au point le plus bas dans un grand voisinage donnée).
- Attributs fondés sur le calcul des valeurs propres du tenseur de structure local (9) : l'analyse des voisins proches de chaque point (voisinage sphérique de 1,25 m) permet d'obtenir des informations sur la structure locale du nuage de points 3D. Les trois valeurs propres du tenseur de structure sont combinées pour fournir divers descripteurs locaux : planarité, sphéricité, anisotropie etc.
- Attributs planaires (3) : les voisins de chaque point servent également à estimer un plan local dont on se sert pour calculer divers indices tels que la normale au plan.

Les attributs fondés-écho se servent de la connaissance de la position du point courant dans l'onde à laquelle il appartient : numéro d'écho, nombre d'échos dans l'onde etc. Ils sont au nombre de quatre et peuvent être calculés à partir de la simple connaissance d'un nuage de points multi-échos.

Enfin, les attributs ROC (8) sont extraits des deux méthodes de traitement des ondes présentées dans la partie 4 :

- La décomposition simple par Gaussienne Généralisée fournit ses trois paramètres de forme (l'amplitude A, la largeur w et l'aplatissement α de l'écho) ainsi que trois paramètres issus de l'étape de calibration radiométrique des données ROC. Il s'agit de la section rétrodiffusante σ, de la section rétrodiffusante normalisée par la surface illuminée par le faisceau laser σ^0 et enfin la section rétrodiffusante normalisée à la fois de la surface illuminée et de l'angle d'incidence du faisceau laser sur la cible atteinte γ.

– L'approche par processus ponctuel marqué fournit un nombre de paramètres variable par écho puisque toutes les fonctions modèles choisies n'ont pas le même nombre de paramètres. On a donc choisi de synthétiser les informations de forme obtenues avec deux paramètres : le modèle sélectionné par le processus ($\mathcal{M}$ Nakagami, Burr ou GG), un coefficient d'asymétrie s de l'écho qui indique si ce dernier est symétrique ou possède un dissymétrie à gauche ou à droite.

5.3 Résultats

La zone de Biberach a entièrement été traitée (2,3M de points). La qualité des différents classifications testées se fonde sur le critère de précision moyenne (*Average Accuracy* – AA, qui est la moyenne des précisions par classe). Pour qualifier les résultats, faire la phase d'apprentissage du SVM et l'estimation de ses paramètres, une vérité terrain a été saisie pour environ 1/3 des points 3D de la zone. Le choix d'une saisie 2D, plus simple et plus facile d'accès pour un utilisateur final, empêche la bonne sélection des points en bords de classes et entraîne une sur-estimation de la qualité de classification.

5.3.1 Sélection d'attributs

Les trois méthodes de sélection d'attributs ont été lancées sur le vecteur de 27 attributs. On obtient donc trois classements différents que l'on synthétise également en faisant la moyenne des trois classements. Les attributs ROC sont segmentés en trois groupes, ce qui donne une première information de leur pertinence vis-à-vis des autres attributs en présence : les bien classés (dans l'ordre A γ et σ), les mal classés ($\mathcal{M}$, w, s et α) et celui dont le comportement est plus fluctuant en fonction des algorithmes (σ^0). L'ordre est ainsi :

$$A > \gamma > \sigma > \sigma^0 > \mathcal{M} > w > s > \alpha$$

Le mauvais classement de $\mathcal{M}$, de s et de α est assez logique. On trouve s et α dans des intervalles de valeurs assez restreints et on a déjà pu constater que l'asymmétrie et l'aplatissement des échos n'étaient pas spécifiques à des classes d'objets mais à des interactions faisceau laser/cible(s) particulières. De même, la sélection de modèle par l'approche stochastique fournit des résultats un peu bruités puisque deux jeux de paramètres pour deux modèles différents peuvent fournir des formes quasiment équivalentes. Enfin, la largeur de l'écho w est classée comme peu discriminante alors qu'elle est énormément utilisée dans la littérature. Ce classement peut provenir du fait que les méthodes sélectionnées ne gèrent pas correctement les variables corrélées. En effet, la largeur de l'écho est également présente dans σ, σ^0 et γ.
De manière similaire, pour les autres attributs spatiaux et multi-échos, nous retrouvons des attributs toujours très bien classés, telle que le nombre d'échos par onde ou la hauteur relative du point par rapport au sol, et d'autres toujours mal classés.

En supplément du simple classement des attributs, une manière plus robuste de qualifier l'apport d'une variable est d'évaluer sa contribution à l'amélioration de la qualité de classification. Pour cela, pour chacune des trois méthodes, nous introduisons de manière progressive dans le classifieur SVM un attribut par un en descendant dans le classement. Nous procédons alors à la phase d'apprentissage, de classification et de qualification à chaque étape. Une telle méthode nous permet de savoir pour chaque méthode (1) quand

est atteinte la meilleure classification ; (2) le nombre minimal d'attributs suffisant pour atteindre une qualité proche de la meilleure possible. On qualifie donc de pertinents les attributs inclus dans ces sous-ensembles minimaux (cf. tableau 3).

- F-score : une très bonne qualité est atteinte pour sept attributs, dont l'amplitude de l'écho A, alors que la meilleure qualité est obtenue avec les 27 attributs. Cette méthode s'avère donc peu performante.
- ReliefF : les 18 premiers attributs permettent d'obtenir la meilleure classification mais six attributs fournissent une qualité assez proche. A y appartient une nouvelle fois.
- SVM-RFE : la meilleure classification est obtenue avec un sous-ensemble de huit attributs. Cela semble donc être la méthode de sélection la plus efficace. Parmi ces attributs se trouvent trois attributs ROC : A, γ et σ.

5.3.2 Classifications

Différentes stratégies de classification sont adoptées pour de nouveau évaluer l'apport (ou non) des attributs ROC. On a en particulier comparé ainsi la classification des 27 attributs à celles des 8 attributs ROC, des 19 attributs non ROC, des sous-ensembles minimaux obtenus pour chacune des trois méthodes de sélection d'attributs ainsi qu'à celles des 27 attributs pondérés par les mesures de pertinence obtenues grâce au F-score et à ReliefF. On s'aperçoit que les attributs ROC ajoutés aux attributs non ROC améliorent la qualité de classification, que la pondération des attributs y contribuent aussi, et que la meilleure qualité est obtenue avec le sous-ensemble de huit attributs (dont trois ROC) sélectionnés par la méthode SVM-RFE (95.3%). Tous les chiffres sont synthétisés dans le tableau 3.

Stratégie	**Average Accuracy** (%)
Attributs ROC (8)	64,8
Attributs non ROC (19)	91,4
Tous les attributs (27)	94,35
F-score + sélection du meilleur sous-ensemble (7)	93,8
F-score + pondération des attributs	94,4
ReliefF + sélection du meilleur sous-ensemble (6)	94,6
F-ReliefF + pondération des attributs	94,7
SVM-RFE + sélection du meilleur sous-ensemble (8)	**95,3**

TAB. 3 – Qualité de classification pour les différentes stratégies adoptées.

Les principales erreurs rencontrées concernent des mauvaises classification ponctuelles liées à des comportements locaux non pris en compte lors de la phase d'apprentissage du classifieur. Il s'agit de points se trouvant sur les bords de bâtiments, sur les superstructures des toits ou sur les objets du bas-sursol (voitures, poteaux). Ils ont un comportement de végétation (rétrodiffusion multiple, faible réflectance) et sont donc étiquetés comme tels. La deuxième source d'erreur provient de la forte influence des attributs les plus discriminants comme la hauteur des points par rapport au sol ou leur amplitude. Dans les cas correspondant à des comportements non appris, l'étiquetage de zones entières peut être erroné. C'est le cas de certains bâtiments bas, aux toits souvent en métal, qui sont confondus avec la classe *sol*. Toutefois, une analyse globale de la scène classée permet de corriger ce type d'erreurs (Figure 7).

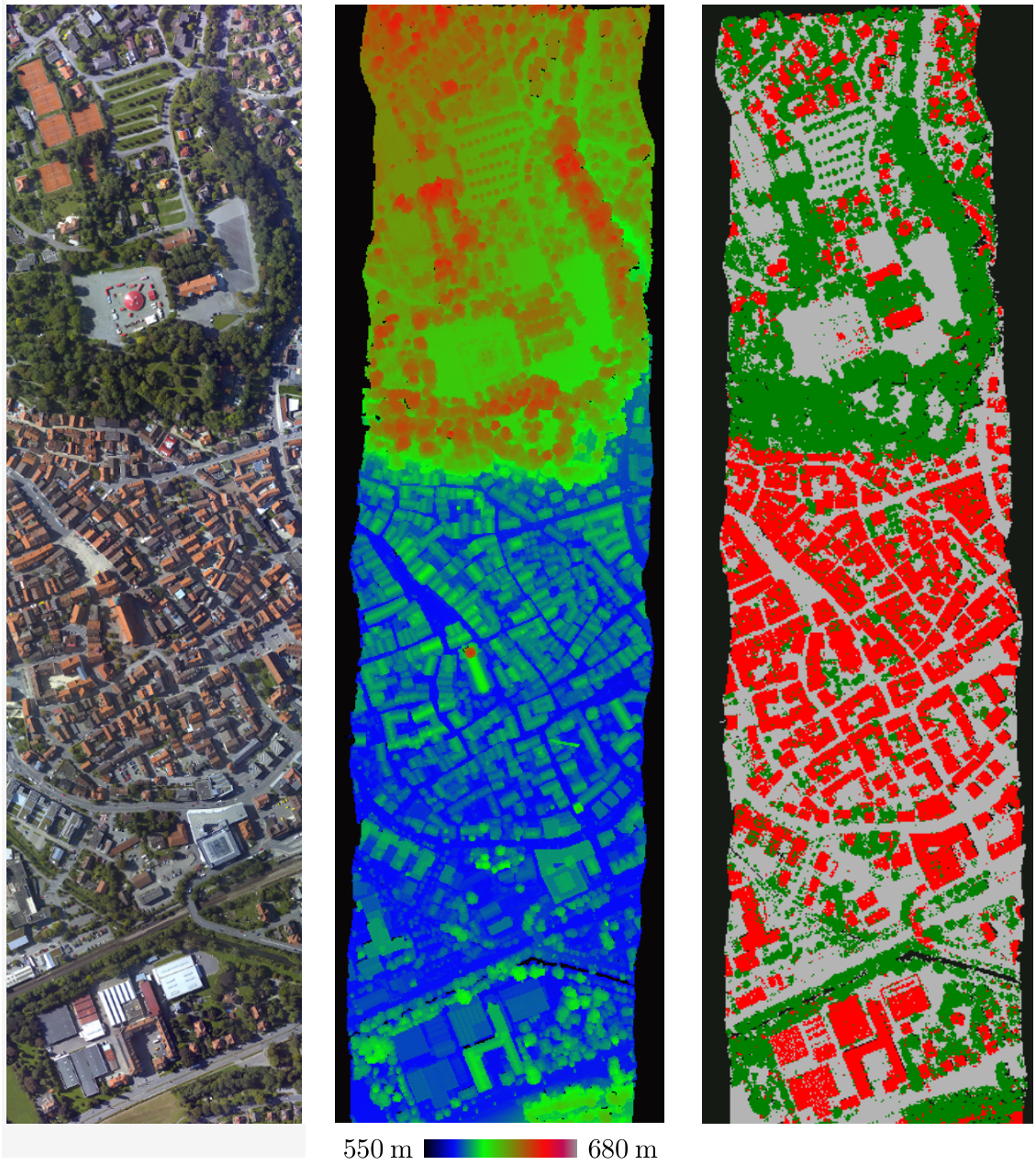


Fig. 7 – Résultats sur la zone de Biberach. Gauche : orthoimage. Milieu : nuage de points 3D généré par traitement des ondes lidar. Droite : classification obtenue (Rouge : bâtiments, vert : végétation, gris : sol).

5.4 Conclusions et perspectives

L'approche proposée a permis d'une part la classification *bâtiment/végétation/sol* de nuages de points 3D lidar en utilisant des attributs issus de plusieurs modélisations des onde complètes et d'autre part l'évaluation de la pertinence de ces derniers. Cette approche est flexible et peut accueillir d'autres attributs lidar si besoin ait. Nous avons réalisé une évaluation objective des attributs ROC en les confrontant à un grand nombre d'attributs calculables sans la connaissance des ondes complètes à travers trois méthodes de sélection d'attributs et diverses stratégies de classification. Les qualités de classification obtenues sont autour de 95%, ce qui montre également l'efficacité générale de la méthode.
L'analyse particulière du comportement des attributs ROC montre que :

- Trois attributs à savoir l'amplitude A, le coefficient de rétrodiffusion γ et la section rétrodiffusante normalisée par l'aire d'interception du faisceau laser σ sont parmi les plus discriminants.
- La dissymétrie et l'aplatissement des échos sont respectivement pas du tout et peu discriminants.
- Une fonction paramétrique symétrique suffit pour modéliser les échos des ondes dans ce cadre de classification *bâtiment/végétation/sol*.
- γ et σ sont obtenues via des méthodes de calibration radiométriques qui sont donc fondamentales pour notre problème.

Les perspectives de travail pour l'analyse des zones d'occupation du sol sont nombreuses. Tout d'abord, une étape de régularisation paraît primordiale. En effet, le classement se fait point par point sans tenir de la similarité avec les voisins. L'ajout d'informations contextuelles devrait améliorer la qualité de classification. Dans un second temps, étant donné que notre objectif était l'étude de la contribution des données ROC nous avons restreint notre problème à trois classes. Cependant, l'analyse des attributs extraits de la modélisation des échos par Gaussienne Généralisée nous a montré un fort potentiel pour discriminer matériaux de toits et surfaces au sol. Cela pourrait être testé mais nécessite également le développement de méthodes de calibration radiométrique adéquates. Enfin, l'enrichissement des thèmes d'occupation du sol passe également par le retour à la géométrie d'acquisition des ondes lidar. En effet, les ondes lidar mixtes, c'est-à-dire qui contiennent des points lidar étiquetés dans différentes classes, peuvent permettre, au moins en deux dimensions, l'apparition de nouvelles classes (bords de bâtiment, forêt dense, forêt éparse etc.).

6 Conclusions

Dans un contexte peu abordé dans le domaine du lidar aéroporté topographique, nous avons proposé une méthode complète, polyvalente et automatique pour le traitement et la classification de données issues de systèmes petite empreinte à retour d'onde complète. Elle permet de traiter une grande variété de signaux lidar en entrée tout en fournissant des résultats de classification de très bonne qualité. Dans un contexte d'évaluation de la technologie à retour d'onde complète apparue très récemment, nous avons également intégré des outils de comparaison vis-à-vis des données issues des systèmes lidar multi-échos classiques. La principale conclusion est que les informations que l'on est capable d'extraire sur la morphologie des échos laser sont discriminantes pour la classification des zones de sol, bâtiment et végétation. Nous nous sommes focalisés sur l'étude du milieu

urbain mais l'approche développé est totalement applicable aux milieux naturels et ne nécessitent aucun développement supplémentaire. Dans le cadre d'un travail mené au sein d'un agence cartographique national, l'IGN, cette thèse a également permis d'améliorer nos connaissances sur le vrai potentiel et les possibilités liés aux données ROC ainsi que la création d'outils pour leur manipulation et leur traitement.
La principale limitation de notre travail est la non-prise en compte des cohérences spatiale et temporelle entre ondes voisines qui devraient permettre à la fois de faciliter et d'améliorer les processus développés. Cela est nécessaire dans un objectif de passage à l'échelle des traitements pour de la cartographie à très grande échelle.

Les perspectives générales de recherche pour l'analyse des ondes lidar sont de trois ordres. Tout d'abord, des travaux sur la simulation d'ondes lidar pourraient permettre de mieux comprendre les influences particulières de la radiométrie ainsi que de la géométrie et de la micro-géométrie des cibles. L'objectif est d'améliorer à la fois les processus de classification en enrichissant le nombre de classes (matériaux de toits et de sol) et les méthodes de calibration radiométriques des données ROC. Dans un deuxième temps, la classification devrait être améliorée avec l'intégration de connaissances additionnelles telles que plusieurs points de vue lidar ou l'utilisation d'images aériennes. Enfin, si l'on s'écarte des chaînes de traitement classiques fondés sur les nuages de points 3D, il semble envisageable de mener à bien un traitement du volume 3D de données ROC pour une reconnaissance de formes directe sans analyse préalable.

Part I

Introduction

Chapter 1

Introduction

1.1 The analysis of urban areas

In recent years, the analysis of urban areas has become a challenging problem for the remote sensing community, due to advances in the technology. The ground sampling of satellite optical images is nowadays sub-metric. The most recent hyperspectral sensors can collect more than one hundred bands with also an increasing spatial resolution [Fau07]. Besides, aerial optical images with spatial resolution below 10 cm are available, opening new fields of research and potential applications with the detection of finer structures [Bré10]. Furthermore, in addition to data acquired in the nadir view, Synthetic Aperture Radar active imagery also contribute to the analysis of urban areas, especially for time critical events and when frequent datasets are required [Soe10].
A large range of applications actually appear when processing such high resolution data in urban areas [NRS05; Urb09]. The detection and characterization of urban structures such as man-made constructions, roads, or vegetated areas are necessary inputs for Geographical Information Systems *i.e.*, for generating databases of 3D buildings, trees, land-cover classes *i.e.*, finally for **automatic mapping purposes**. Depending on the resolution and the type of the data available over the area of interest, these databases may not be at the same scale, and will not be used necessarily for the same usage: change detection, disaster management, telecommunications, tourism, defense, navigation, simulations for detecting and solving ecological or social issues etc.

For such purposes, the airborne laser scanning (ALS) technology has many beneficial properties. Topographic data is captured with high density, allowing to reconstruct high spatial data over large areas. Furthermore, ALS systems deliver high resolution height data with an accuracy around 10 cm in a highly automatic way. We have a direct access to the geometrical properties of the objects without the need to first recover 3D features from optical images. The technique is therefore particularly suited for planning, monitoring and control purposes. In the last decade, the 3D representation of topographic objects has become more and more prominent, especially in consumer applications. The last ten years have gone to show that the airborne laser scanning community has actively taken part in the developments of new and efficient methodologies for the automatic acquisition of 3D urban objects (especially buildings), ground surface reconstruction (Digital Surface and

Terrain Models), as well as land-cover classification [ST09; Are09; Rot09; PY09; ZN10; VM10; Elb10].

1.2 Airborne laser scanning...

Airborne laser scanning (ALS) for topographic purposes is an active remote sensing technique providing direct range measurements between the laser scanner and the Earth topography. Such distance measurements are mapped into 3D point clouds. Such 3D points are reliable estimates of the location of the terrain surface and of the above-ground objects, but sample irregularly the acquired landscape. The altimetric accuracy of a topographic lidar measurement is indeed high (< 0.1 m), thanks to the direct georeferencing process, steadily improved in the two last decades with the introduction of the differential GPS technology and Inertial Measurement Units.
First active sensors carried by airborne or satellital platforms were designed at the beginning of the 1970's. They provided 1D profiles along the sensor track (nadir view) by sequences of single pulses. Modern sensors acquire many parallel overlapping strips. Such technology provides denser point clouds with a more regular distribution on the Earth surface: the point density can reach more than 100 pts/m^2 in some specific applications, *e.g.*, river dike monitoring.
ALS has become a well established technique for the acquisition of geospatial information, and is now widely established as a major tool for generating high quality 3D representations of the landscape. Due to the large number of devices and their yearly improvements, the growing number of surveying companies, the large variety of acquisition specifications, and the increasing number of available visualization and processing softwares, ALS data, also referred to **topographic lidar** data, has been adopted as standard input for many remote sensing applications.
The interrelationship between ALS and aerial photogrammetry has been intensively discussed in the last decade [Bal99c; Bre06]. Various comparison factors concerning both data acquisition (*e.g.*, coverage, weather conditions, costs, etc.) and surface reconstruction (*e.g.*, accuracy, redundancy, post-processing time, etc.) have to be taken into account to choose the optimal method for a certain purpose. They are commonly considered as complementary surveying techniques.

Depending on the geometry of illuminated surfaces, several backscattered echoes can be recorded for a single pulse emission. Such systems are called **multiple pulse** or **multi-echo** lidar sensors. This is particularly interesting in forested areas since lidar systems can measure both the canopy height and the terrain elevation underneath at once, contrary to photogrammetric techniques. Such a unique ability led to the immediate acceptance of ALS data for forest monitoring and management [HHL$^+$04; AMR05].
Furthermore, lidar data are known to be useful in many specific applications, in addition to 3D city modelling, such as bridge and power line detection [SV06; JS10], Digital Terrain Model generation [KPO$^+$07; BC10], corridor, coastal or opencast mapping [IL99], as well as metrology [FCH04] and cultural heritage [DBFJ08].

1.3 and the full-waveform technology

Topographic lidar data only provide a basic geometric representation of a scene: three-dimensional coordinates. The "intensity" feature is sometimes available in addition, but due to its low quality and the lack of knowledge on what it really is, it has been barely used.

The new technology of **full-waveform** (FW) lidar systems has emerged in the last twenty years, and has become popular in the last five years, with the emergence of airborne topographic commercial devices. It permits to record the received signal for each transmitted laser pulse, the result is called a *waveform*. Since the waveform is digitized at constant rate (GigaHertz resolution), FW data is thus a set of equally-spaced discrete samples of the amplitude of the echo signal.
Such a sample sequence represents the progress of the laser pulse as it interacts with the reflecting surfaces, and thus is constructed as an amplitude-against-time dataset (*cf.* Figure 1.1). Hence, FW lidar data yield more than a basic geometric representation of the Earth topography. Instead of clouds of individual 3D points, lidar devices provide connected 1D profiles of the 3D scene, which allows gaining further insight into the structure of the scene. Thus, in addition to range measurements, further physical properties of objects included in the laser diffraction cone may be revealed.

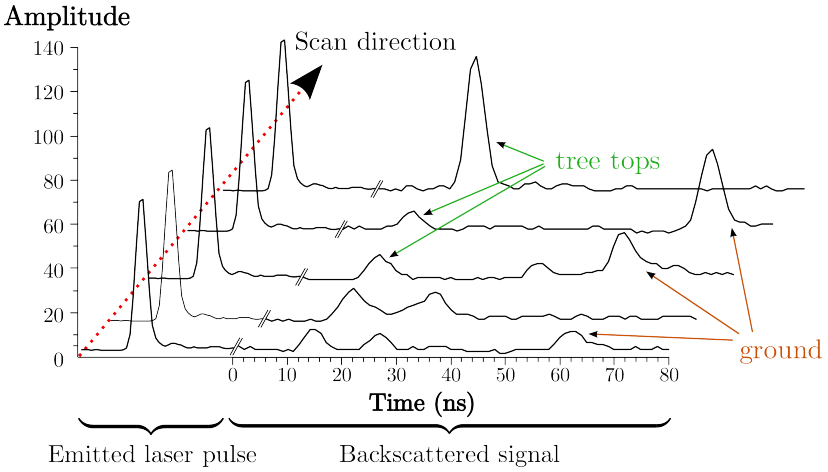


Figure 1.1: Successive lidar waveforms plotted in the laser beam direction plane (small footprint sensor).

Waveform analysis allows to set up advanced processing methods which increase pulse detection reliability, accuracy, and resolution: more 3D points can therefore be detected than multiple pulse systems. Digitized waveform data supplements the time-of-flight and intensity data that multiple-pulse sensors provide. This is especially beneficial for extended target ranging and for complex structures such as trees and vegetation. Furthermore, this new technology gives more control to the end user in the interpretation process of the physical measurement. It provides additional information about the structure and the physical backscattering properties of the illuminated surface (reflectance and geometry).

Full-waveform sensors have been first used for bathymetric purposes. Because of the complex shape of the signals in water areas, the seafloor depth cannot be correctly estimated on-board, during the survey. Consequently, a digitization process has been set up [GM88; GCRR00]. The first truly operational FW topographic system, LVIS, from NASA, appeared in 1999, and demonstrated the value of recording the entire waveform for vegetation analysis [BRH99]. This is the consequence of early efforts made since the beginning of the 1980's [Lin97]. The first commercial full-waveform lidar system, the LiteMapper 5600, appeared in 2004 [HUG04]. This is a small-footprint system (<1 m), in contrast to previously mentioned devices (>2.5 m), which has been designed without any dedicated application. However, several authors have shown the interest of such kind of data both for natural and urban areas, that is why major laser scanning companies have decided to add digitization terminals to their multiple-pulse systems. Since laser scanners with waveform digitizers are becoming increasingly available, many studies have been carried out during the last five years. In particular, the three last ones have witnessed a tremendous increase of research and literature on the subject.

Some recent works have presented original methodologies for processing full-waveform data [Jut07; Lin09; Rei10], but the contribution of such data for the analysis of urban areas has been, in fact, barely investigated.

1.4 Specific context of the thesis

The Institut Géographique National (IGN) is the French national mapping agency. The French State has entrusted the IGN with developing the Large Scale Reference system (RGE®) that incorporates data from its own databases or databases from other companies. Among its objectives is the production and the update of the topographic component of the RGE (BD TOPO® [TOP10]), which corresponds to CityGML's LOD0, LOD1, and LOD2. Indeed, LOD0 concerns the 2D geometric description of the elements with semantic attributes, whereas 3D building modelling attains LOD2.

In the operational context of national mapping, the IGN owns two digital cameras for orthoimage generation, and has purchased in 2007 an airborne laser scanner with waveform digitizer (Optech 3100 device). At first, the aim is improve the quality of the Digital Terrain Models (DTM), with an accuracy better than 0.3 m in altimetry, and better than 1 m in planimetry. The IGN has first focused on forested areas and regions vulnerable to flooding hazards (rivers and coastal areas). For such purposes, only the 3D point cloud (spatial information) is used. A fully operational workflow has been set up, from the strip registration task to the filtering process. The DTM are computed using the points labelled as *ground* during the latter step. Surveys are carried out with medium density point clouds ($\sim$ 2-3 points/m^2 per strip), which has been proved to be sufficient to reach the requested qualities.

The emergence of full-waveform data has arisen many questions about its genuine potential, both in forested and urban areas. Physical information about the scattering properties of the objects are now available, in addition to spatial features. The context of this thesis is therefore to assess the real contribution of small-footprint full-waveform airborne laser scanning data in urban areas [1]. The analysis will be carried out in the scope of land-cover

[1] A similar work on forested areas has be performed in the IGN in parallel to our thesis.

classification, which is the preliminary step of various kinds of algorithms for populating topographic databases. The development of more general land-cover classification processes is obviously at stake in national mapping agencies.

1.5 Objectives and limitations

The general objective of this thesis is to take as an input raw full-waveform lidar data *i.e.*, 1D signals (with all the metadata required for georeferencing them), process these waveforms to generate a 3D point cloud, and classify it with respect to the additional information retrieved during the processing step. In addition to this main workflow, our will is to understand the real potential of such recent data. From an engineer point of view, our goal is also to develop new tools to handle, display, and process lidar waveforms, as a first step for future developments.

Since we aim the analysis of urban areas with high resolution data, small-footprint lidar waveforms will be adopted. This means that the footprint size is inferior to 1m *i.e.*, inferior to the size of the objects of interest (except for vegetated areas). Consequently, each echo within the waveforms is likely to correspond to a specific target, and its shape analysis may reveal information about the scattering properties of this target.
When land-cover classification is at stake, multispectral or hyperspectral images should be involved because of their highly relevant spectral content. Nevertheless, the fusion of multisource data and the comparison of these two kinds of data will not be tackled. Our work will focus on ALS data, and the contribution of FW data will only be addressed with respect to multiple pulse sensors.

Three classes of interest are selected: *ground*, *vegetation*, and *building*. The ground class gathers roads and pavement, as well as the street items such as cars or poles. Vegetation include both trees and low above-ground vegetation such as hedges or bushes.
The discrimination of tree species, ground surfaces, and roof materials has not been addressed. We consider that tree species classification mainly relies on the 3D repartition of the points and is not FW specific. Conversely, opaque targets with different surfaces may be segmented using the scattering properties revealed by full-waveform data analysis but not with merely 3D spatial data.

The final goal of the thesis is the classification of the 3D point clouds generated from lidar waveforms. Therefore, several other topics of interest in urban areas will not be tackled. They are enumerated below, and will not be mentioned until the conclusions of the thesis.

- **Direct classification of the waveforms**, instead of classical 3D point labelling. Considering such classification still possible at the end of our workflow, this aspect has not been developed.
- **Waveform simulation** would be relevant for classification by adopting an analysis-synthesis strategy. However, we are not convinced to be able to reach the quality of simulation of small-footprint data that would be necessary for more advanced classification. For instance, discriminating grass and asphalt surfaces or slanted and

tilted roofs require sophisticated reflexion models that the traditional Lambertian assumption.

- **Pattern recognition on enhanced point clouds or on FW data**. Waveform processing allows to detect additional 3D points, compared to multiple lidar systems. Such points mainly lie on building edges, superstructures, and trees. Furthermore, the spatio-temporal data volume can be directly processed to extract objects of interest (planes, lines, building edges etc.), that would be relevant inputs for many 3D building modelling algorithms. The contribution of FW data may therefore be assessed from a pattern recognition point of view. Nevertheless, the IGN has already developed an efficient semi-automatic workflow for 3D city modelling from high-resolution optical images. Addressing such issues would therefore be redundant or inefficient without high density point clouds.

1.6 Overall strategy

Our strategy is composed of two main steps that are sketched in Figure 1.2. The inputs and the outputs of each of these two steps are illustrated in Figure 1.3.

1. **Waveform decomposition and modelling.** Raw data are 1D signals. To generate a 3D point cloud, we need to extract the locations of the modes within these signals, since each mode (called echo) is assumed to correspond to a specific object (*i.e.*, a waveform is the superimposition of several signals considered as independent). This task is called **decomposition**.
   The shape of these echoes depends on the properties of the hit target. Characterizing this shape by a parametric function is interesting since more than simple spatial features can be derived, and fed into a classifier. This is the **modelling** step.
   These two parts are traditionally tackled successively. We will first adopt such strategy and select one relevant modelling function, whose parameters will be estimated with a Least-Squares algorithm. Secondly, another methodology will be presented. Instead of being limited to a unique model, the development of a stochastic approach using a marked point process will allow us to find the best-fit model in a given library, and merge the decomposition and modelling steps. A better adjustement leads to improved signal reconstruction and data compression.

2. **Classification.** The first part has allowed us to derive, for each waveform, several range values. A georeferencing process turns such values into a 3D cartographic point cloud. Along with spatial information are associated various shape features, named full-waveform features, extracted from the two methods mentioned above. Besides, a local neighbourhood analysis is performed for each point, providing other spatial attributes relevant for our classification purpose.
   These two kinds of features are merged into a single feature vector, input of the classification task. A **supervised classifier** is selected, namely Support Vector Machines. In order both to choose the most relevant feature set, and evaluate the relevance of full-waveform features in comparison to geometric attributes, a **feature selection** step is added. Two options will be considered: performing the selection before making the decision, and wrapping the selection into the classification. Consequently,

conclusions may be provided, independently or not of the selected discrimination method.

The methodology for land-cover classification in urban areas by full-waveform lidar data analysis is designed to be flexible. Depending the theoretical understanding of pulse propagation in such regions or the sensor specifications, the modelling functions may vary. With the proposed approach, it can be changed without modifying the feature selection process and the classifier rules. Furthermore, the same conclusion can be drawn when adjusting the level of detail of the classification: one simply has to modify the training samples of the classifier.

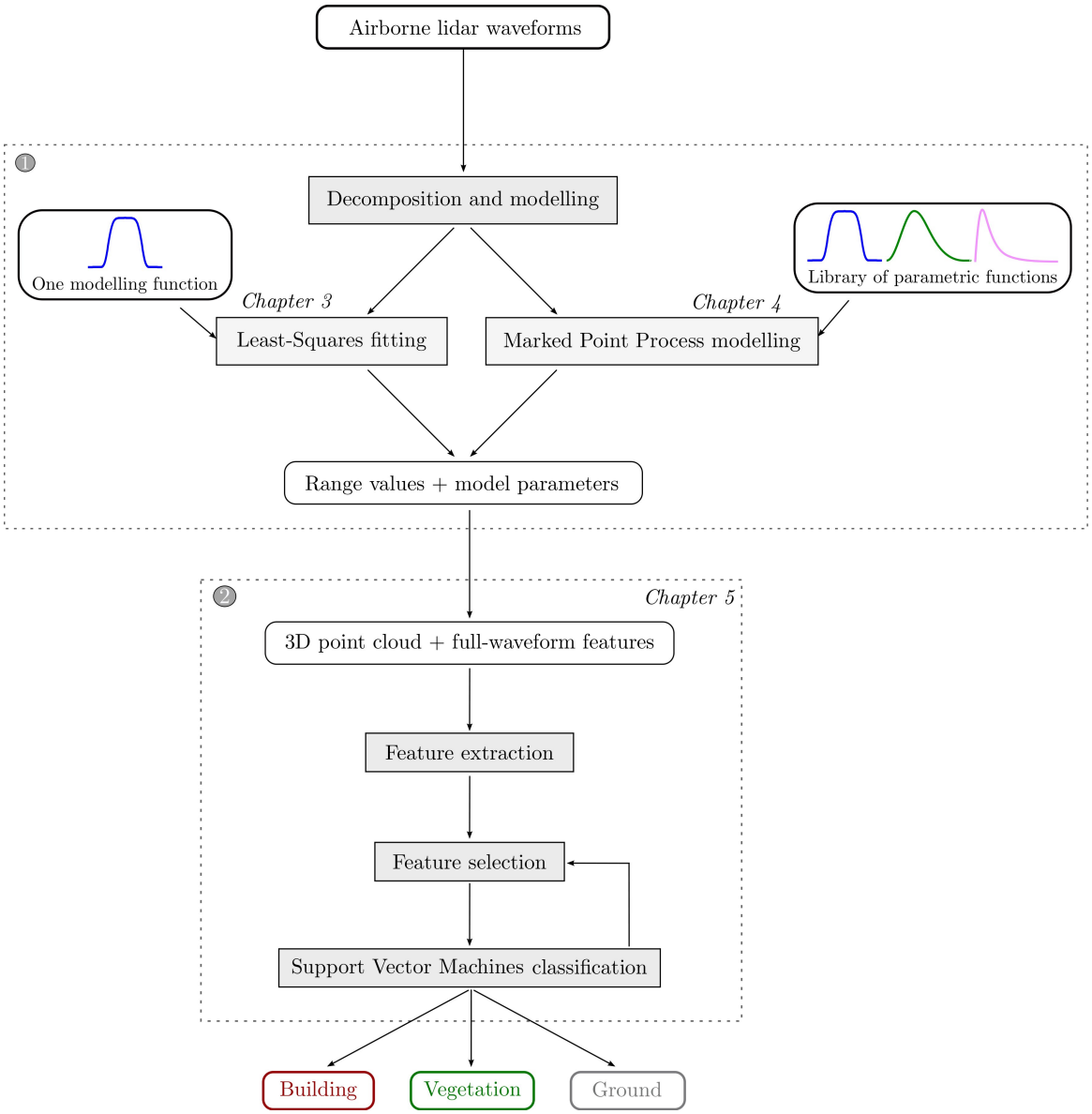
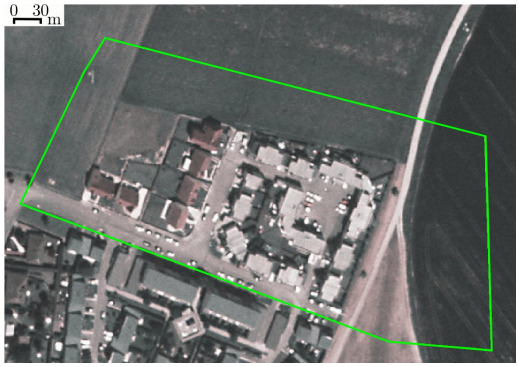
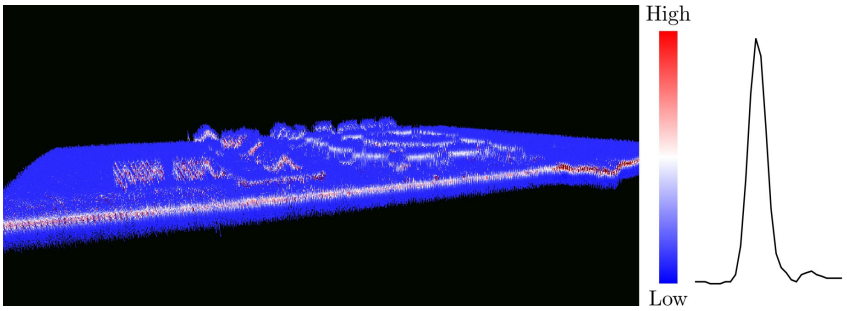


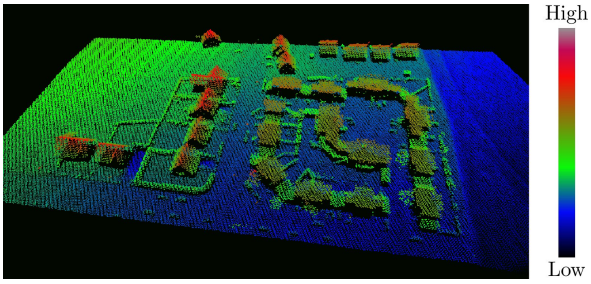
Figure 1.2: Overview of the flowchart of the thesis.



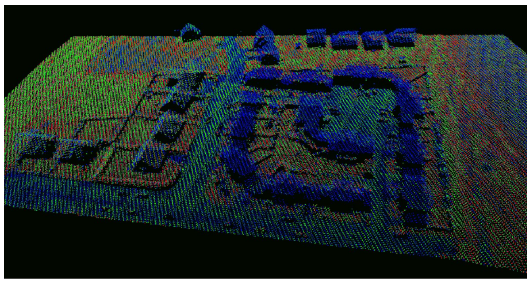
(a) Ortho-image ©IGN.



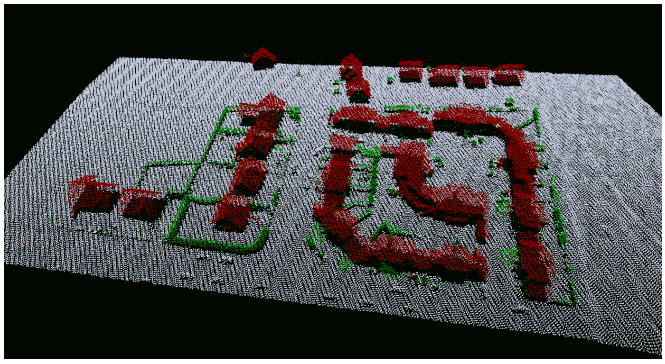
(b) Raw 3D full-waveform volume.



(c) Estimated 3D point cloud (height coloured).



(d) Estimated 3D point cloud (amplitude coloured).



(e) Classification: Building – Vegetation – Ground.

Figure 1.3: (a) Illustration of the proposed strategy over a small area of interest. (b) Our input data are 1D signals, here georeferenced and coloured with respect to the amplitude of each bin of each waveform. (c) These signals are decomposed into echoes, which generates a 3D point cloud. (d) These echoes are adjusted by a model, which provides shape features. (e) These features along with spatial ones are used as input in a supervised classifier.

1.7 Structure of the document

This work is divided in four main chapters, organized in three parts. The first part introduces full-waveform data and existing process. The second one deals with waveform decomposition and modelling, whereas the last part addresses the issue of 3D point cloud classification.

- Chapter 2 presents airborne laser scanning and the specificity of the full-waveform technology. The first aim of this chapter is to provide information about the recording process so that an end-user may know what these "raw" signals really represent. Furthermore, the main FW sensors are described, whatever their applications: pioneer bathymetric sensors, NASA experimental systems, and eventually current available topographic devices. Finally, an as comprehensive as possible state-of-the-art of the main existing FW processes is provided. The main signal processing methods is firstly presented. Secondly, the subsequent applications on generated 3D point clouds are described, both for urban and forested areas.
  This chapter also aims to introduce general models and notions that will be taken up in the following chapters, *e.g.*, the lidar equation or the echo amplitude calibration and correction.

- For generating 3D point clouds with echo shape features, a waveform processing step is mandatory. In Chapter 3, we propose a variation of the standard Gaussian decomposition method. The echoes are modelled with the Generalized Gaussian function, introducing a new feature, namely the echo shape. The objective is to improve signal fitting, and therefore better estimate the morphological echo attributes: amplitude and width. Such model coupled with an enhanced echo detection method allows to retrieve additional 3D points compared to traditional multiple pulse data. The waveform decomposition and modelling stages are quantitatively and qualitatively evaluated, and, finally, the extracted FW features are analyzed through the perspective of land-cover classification.

- An alternative approach of the standard Gaussian-based decomposition methods is proposed in Chapter 4. The problem is addressed using a marked point process. Such stochastic framework offers the possibility to fit each echo by the most suitable function of a predefined library (in our case, three models), and adversely with far less restrictions that using a non-linear least square method. The signal processing task is moved towards a higher level since it is now mainly ruled by physical lidar-based parameters. Furthermore, a better signal reconstruction is expected which would lead to better data compression and representation. The performance and the effectiveness of the method are assessed both on simulated and various real waveforms. The relevance of the model library is discussed for several urban areas of interest. Finally, the fitting process is evaluated and compared to the Gaussian method.

- Subsequent to waveform decomposition and modelling, Chapter 5 deals with the classification of the generated 3D point clouds. In addition to the 3D point geometry and other ensued pure spatial features, the modelling step provides echo shape features that may be discriminant for land-cover classification. The Support Vector Machine (SVM) classifier is presented as a supervised classifier suited to our problem: it is robust to the dimensionality of the data and has good generalization performance,

even if the situation of a limited training set. Because of the fusion of geometrical and FW attributes for discriminating *ground*, *vegetation*, and *building* areas, an attribute ranking approach is set up to assess the relevance of the newly derived FW features, the SVM classifier is coupled with a feature selection step. Two main procedures are investigated, and the results compared in term of accuracy. This chapter ends with a brief presentation of an alternative approach involving Random Forests for FW data relevance assessment in urban areas.

- Eventually, the conclusions about the contributions of our analysis are drawn in Chapter 6, and interesting perspectives of research are also presented for the analysis of urban areas. An outlook on the real potential of FW data is finally sketched.

Chapter 2

Full-waveform topographic lidar

The aim of this chapter is threefold. Firstly, Section 2.1 briefly recalls the principle of airborne laser scanning. Then, the full-waveform technology and resulting data are presented in Section 2.2. The objective is also to gather the knowledge that we have acquired at the engineering level with the processing of the raw data, so that an end-user may really know which kind of information is really provided. Secondly, this chapter may also be considered as a state-of-the-art of the existing sensors (Section 2.3) and methods that have been developed so far to process lidar waveforms (Section 2.4). This comes in useful as an introduction of modelling and classification issues that will be addresses in details in the corresponding Parts. Furthermore, several general models and notions such as the lidar equation or the amplitude feature calibration and correction are introduced in this Chapter, and will be used afterwards in various parts of the following chapters. Eventually, the two datasets that have been processed in this thesis are presented and their main specifications detailed in Section 2.7.

2.1 Principle and technology of laser scanning

Airborne topographic laser scanning (ALS) is an active remote sensing technique providing direct range measurements between a laser scanner device and the Earth surface. ALS physical principle consists in the emission of laser pulses from an airborne (aircraft or helicopter) or spatial (satellite) platform at a high repetition frequency. The distance is derived from the measured round-trip time of the backscattered signal between the sensor and the target [Bal99b]. Then, 3D point clouds are obtained with a direct georeferencing processes:

- An hybrid system using both GPS (differential measurements with a ground station located near the survey area) and inertial measurements (IMU) is used to optimally calculate supporting vector attitudes, the absolute orientation of the laser sensor and by post-processing after the survey, the smoothed best estimated trajectory [HJW02].

- The embedded hardware system perform real-time signal processing tasks: range values are retrieved by finding the location of the modes of the backscattered signal. Such modes are called *echoes*.

- Afterwards, a georeferencing process turns the point value expressed in laser device polar coordinates {range, scan angle } to a {x, y, z} triplet within a given geographic datum. Laser scanning companies traditionally deliver 3D point clouds to end-users.

Basic airborne lidar (Light Detection And Ranging) systems are therefore composed of a laser transmitter and a receiver (rangefinder unit which receives the reflected pulses and measures the distance), a mechanical scanner, a hybrid GPS/INS positioning system, a storage media, and an operating system for signal digitization and on-line data acquisition. This unit monitors and synchronises measurements, and processes data in real-time to extract range values. The georeferenced points are not computed on-board but afterwards [Bal99a; TW04].
The first active sensors carried by airborne or satellital platforms were designed at the beginning in the 1970's. They provided 1D profiles along the sensor track by sequences of single pulses nadir view. By forward motion of the sensor carrier and an additional scanning mechanism in across-track direction (e.g., a rotating mirror, a prism...), strips of 150 m to 600 m swath width are covered by modern sensors, depending on type of device and carrier altitude. Such technology provides denser point clouds with a more regular repartition on the Earth surface.
Airborne topographic systems have now proved to enable fast, reliable, dense, accurate, but irregular mapping of terrestrial landscapes. The accuracy of the measurement is high (typically < 0.1 m in altimetry and < 0.4 m in planimetry).

2.1.1 Physical principles

Both pulsed and continuous wave lasers are being used. Pulsed systems measure the round-trip time of a short light pulse from the laser to the target and back to the receiver. Continuous wave systems carry out ranging by measuring the phase difference between the transmitted and received signal. Most of airborne topographic lidar systems are pulse-based.
Laser pulses are emitted with very high energy, short pulse duration, high firing rate, and narrow beam width. The laser light is coherent with monochromatic radiation. The measurement of the time-of-flight t, created by light travelling in a given medium, allows to simply evaluate the slant range d between the sensor and the target using:

$$d = \frac{v_g}{n}\frac{t}{2} \tag{2.1.1}$$

where v_g is the group velocity of light and n the refraction index of the medium ($n \backsimeq 1.00025$ in the air, and often assumed as equal to 1).
Laser ranging devices actively emit pulses of short duration (typically a few nanoseconds) in the infra-red domain (wavelength between 800 nm and 1550 nm) of the electromagnetic spectrum. The target sizes are therefore much larger than the wavelengths. Depending on the wavelength, the emitted electromagnetic wave interacts with atmospheric particles (absorption or scattering, known to have negligible influence if water vapour and aerosols are absent), but mainly with the illuminated natural or man-made objects belonging to the Earth surface. Thus, the propagation velocity of the laser pulse is most of the time assumed to be constant.

The Pulse Repetition Frequency (PRF) depends on the acquisition mode, determines the flying altitude, and is typically superior to 100 kHz. A pulse release is done when the previous pulse recording is effective. In fact, the PRF is constant. For instance, with a 50 kHz pulse rate, the laser system is not obliged to wait the previous pulse return under a flying height of 3 km. With ALS systems, such borderline cases never occur. However, the latest systems have even the ability to fire a second laser pulse before the recording of the previous pulse [RT08], thus increasing the point density without additional strips.

Due to diffraction, the laser beam inevitably fans out; a typical value for the beam divergence lies between 0.4 and 0.8 mrad. Therefore, a single emitted pulse may cause several echoes from objects located at different positions inside the conical 3D volume traversed by the pulse. The number of multiple returns therefore depends on the site characteristics (see Section 2.1.3), but also on the flight specifications. Indeed, the laser spot diameter D is (assuming a fronto parallel flat target):

$$D = 2H \tan \frac{\beta}{2} \tag{2.1.2}$$

where H is the distance to the sensor, and β the beam divergence. For instance, D=0.5 m at a distance of 1000 m with β=0.5 mrad. In the ALS context, the laser beam once it intersects with the surface (H is then the flight height) produces a spot diameter with known dimension, called the laser footprint, which varies according to the surface (slope and material characteristics), and may be deformed as an ellipsis for flat surface. With increasing beam divergence, the footprint and the 3D conical region become larger, and the probability to obtain multiple echoes per emitted pulse increases. This is particularly interesting in forested areas since lidar systems can measure simultaneously both the tree canopy height and the terrain elevation underneath.

A typical pulse width of τ_p=4 ns corresponds to a length of approximately 1.2 m at the speed of light. This is an important characteristic of any laser system, since it has a major impact on how multiple returns may be detected or differentiated. Two echoes can be discriminated if their distance d is larger than half of the pulse length l_p:

$$d \geq \frac{1}{2} l_p \qquad \text{with } \; l_p = \frac{c}{n} \tau_p \tag{2.1.3}$$

This means that for τ_p=4 ns, targets may be discriminated if their distance is at least larger than 0.6 m.
More details about topographic laser scanning principles (from laser fundamentals to system components) and existing systems may be found in [ST09; VM10].

2.1.2 Lidar measurement formulas

The standard lidar equation is derived from the radar equation. It describes the measurement process by taking the detector and target characteristics into account. It also relates the power of the transmitted and return signals [Jel92]. More detailed formulas have been proposed to model real world constraints [DRC97; CSL01]. Nevertheless, they are only valid for single sources or for flat surfaces. In case of targets that are distributed in space,

the reflected signal is the superposition of echoes at different distances.
The received power P_r can be expressed as an integral:

$$P_r(t) = \frac{D^2}{4\pi\,\lambda^2} \int_0^H \frac{\eta_{\text{sys}}\,\eta_{\text{atm}}}{R^4}\, P_t(t - \frac{2R}{v_g})\,\sigma(R)dR \tag{2.1.4}$$

where t is the time, D is the aperture diameter of the receiver optics, P_r is the received power, P_t is the emitted power, λ the wavelength, H the flying height, R the distance from the system to the target, η_{atm} and η_{sys} respectively the atmospheric and system transmission factors, v_g the group velocity of the laser pulse, and $\sigma(R)dR$ the apparent effective differential cross-section [WUD+06]. The cross-section is called "apparent" since an object reflecting the signal at a given distance may occlude an object further away.
The power of the received signal can also be considered as the sum of the contribution of N targets with their own characteristics:

$$P_r(t) = \sum_{i=1}^{N} (P_{r,i} \star \eta_{\text{sys}} \star \eta_{\text{atm}})(t) \tag{2.1.5}$$

where $\star$ is the convolution operator and $P_{r,i}(t)$ is the echo of the i$^{\text{th}}$ object expressed as:

$$P_{r,i}(t) = \frac{D^2}{4\pi\,\lambda^2} \int_{R_i-\Delta R}^{R_i+\Delta R} \frac{1}{R^4}\, P_t(t - \frac{2R}{v_g})\,\sigma_i(R)dR \tag{2.1.6}$$

where R_i is the mean distance, $[R_i - \Delta R, R_i + \Delta R]$ is the spatial spread, and $\sigma_i(R)$ is the effective differential backscattering cross-section of the object. The reflected signal can be seen as the convolution between the transmitted pulse and the effective differential cross-section. Typically ΔR lies between 0.4 and 1.5 m. Thus, $\Delta R \ll R$, and as a consequence, we have:

$$P_{r,i}(t) \approx \frac{D^2}{4\pi\,\lambda^2\,R_i^4}\,(P_t \star \sigma_i^{'})(t) \tag{2.1.7}$$

where $\sigma_i^{'}(t)$ is the apparent cross-section of illuminated areas within each range interval (in m^2). The power of the received signal can finally be expressed as:

$$P_r(t) = \sum_{i=1}^{N} \frac{D^2}{4\pi\,\lambda^2\,R_i^4} \underbrace{(P_t \star \eta_{\text{sys}})(t)}_{\text{system contribution}} \star \underbrace{(\eta_{\text{atm}} \star \sigma_i^{'})(t)}_{\text{environment contribution}} \tag{2.1.8}$$

The received waveform is therefore the convolution between the emitted pulse and the surface response $\sigma_i^{'}$, that is known to be the backscatter cross-section (BCS).
For a given local survey, η_{atm} is almost constant with the time but is mainly affected by the range value R. $\eta_{\text{atm}} = \exp(-2Ra)$ for [Ste00; JSE02] or $\eta_{\text{atm}} = 10^{-2Ra/10000}$ for [HP07], where a is the atmospheric attenuation coefficient expressed in dB/km and as a function of the laser wavelength and the visibility distance in [Ste00] (empirical spectral approximation based on Koschmieder's formula). In practice, η_{sys} is assumed to be constant over time and for a given lidar sensor.
The apparent cross-section $\sigma_i^{'}$ gathers all the target characteristics and is defined as:

$$\sigma_i^{'} = 4\pi\frac{\rho A_t}{\Omega} \tag{2.1.9}$$

where Ω is the solid angle of the target ($\Omega = \pi$), A_t is the target area ($A_t = \pi R^2 \beta^2/4$), and ρ is the target reflectance. Assuming a Lambertian scattering target $\rho = \rho_m \cos\alpha$, where ρ_m is the material reflectance and α is the angle between the laser beam and the target normal. Therefore, we have:

$$\sigma_i^{'} = \pi \rho_m \cos\alpha \, \beta^2 R_i^2 \tag{2.1.10}$$

Finally, Equation 2.1.8 may be rewritten as follows:

$$P_r(t) = D^2 \eta_{\text{sys}} \rho_m \cos\alpha \sum_{i=1}^{N} \frac{\exp(-2R_i a)}{4R_i^2} P_t(t) \tag{2.1.11}$$

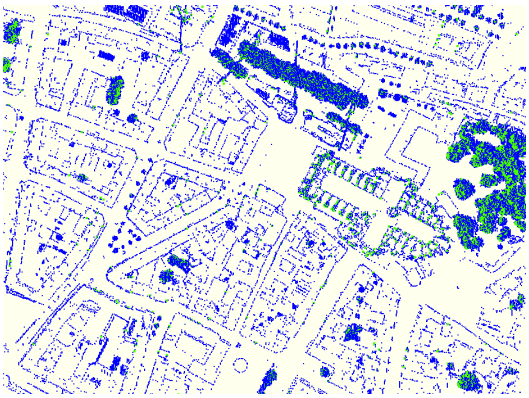
2.1.3 Limitations of multiple pulse systems

The first commercially available airborne laser scanners provided a simple backscattered echo per emitted pulse. The recording of a single echo is sufficient if there is a simple target within the diffraction cone. However, even for small laser footprints (0.2 - 2 m), there may be many objects within the travel path of the laser pulse: individual scattering contributions are generated for each illuminated object. **Multiple pulse** or **multi-echo** laser scanning systems are designed to record more than one echo. The earlier systems typically collected first and last pulses only. The more recent systems are now able to discriminate up to six individual returns from a single pulse [TW04]. Multiple reflections occur on vegetated areas. When the vegetation is not very dense, it is often assumed that the first echo belongs to the canopy top and the last pulse to the ground. In reality this is not always the case. For urban areas, in a particular viewing angle, when the laser beam hits a building edge, two echoes can be generated. The first pulse corresponds to the spot illuminating a part of the roof, while the second one to the spot hitting the ground (see Figure 2.1). The same behaviour may be noticed within building roofs when superstructures (mainly chimneys) of sufficient height are present. Large laser footprints enhance this phenomenon.

For multi-echo systems, pulse detection is performed on the flight on the backscattered signal. The hardware system detector turns a continuous waveform to several time-stamped pulses, giving the position of individual targets. Many peak detection methods exist, but lidar manufacturers do not provide any information and qualification about the method implemented in their hardware systems. The number and the timing of the recorded pulses are critically dependent on the detection method, since discrete systems provide only a range value at a certain rise time of the echo and no information about the echo shape [WUM+04]. Standard peak detection methods are detailed and analyzed in Chapter 7 of [ST09].

The returned signal depends on the emitted pulse, the incident angle between the emitted pulse and the target, the target roughness and reflectivity behaviour (expressed with the Bidirectional Reflectance-Distribution Function – BRDF), the distance to the target (R^2 attenuation), and the occlusion and influence of previous hit targets in more complex situations. Consequently, distance computation methods based on simple trigger echo detectors suffer from several drawbacks in the following critical measurement situations:

(a) Orthoimage ©IGN.



(b) Location of lidar multiple echoes. Blue and green points correspond to second and third echoes, respectively.

Figure 2.1: Illustration of the multiple pulse property of lidar systems in urban areas (Amiens, France). Building edges and vegetated areas are clearly enhanced with a 0.8 m footprint size (Optech 3100 sensor).

- **Misestimation of pulse location**: depending on the target characteristics and laser incidence angle, the laser pulse may be attenuated, stretched or skewed. This may have an impact on the performance of the detection method, even for flat areas [WUM+04]. The resulting noise in the final 3D point cloud is mainly altimetric (more than 0.3 m in some cases).

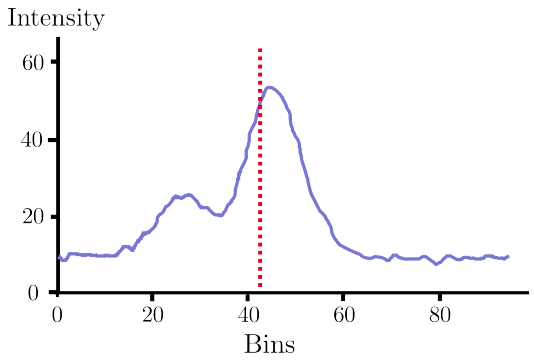
- **Missing echoes**: lidar systems may be limited to detect a maximum number of peaks (*e.g.*, 4). When this upper bound is exceeded, echoes are inevitably missed. It happens in vegetated areas.

- **Overlapping pulses**: standard detection techniques may not resolve signals into separated echoes when several modes are closely located. Firstly, signal modes can be missed (Figure 2.2.b). Secondly, the pulse location may be estimated between these two pulses and then generate a wrong 3D point [PSTA05]. In urban areas, such cases can be:

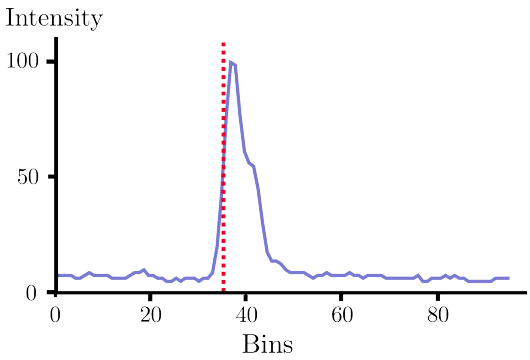
  - low ground vegetation, low above ground urban items, and building superstructures: they may give rise to a small bump on the leading edge of the ground or building echo;
  - vegetation areas, due to the tree canopy reflection (top, branches, and stem).

- **Weak echoes**: The two first echoes contain a large majority of the total reflected signal power. Real-time detection of more than three or four pulses requires thus the detection of low intensity signal within noise. Weak echoes may therefore be missed (Figure 2.2.a). These situations correspond to:

  - echoes within trees and lying on the ground below;
  - roof echoes attenuated by strong material absorption and/or significant incident angle.



(a) A two-mode lidar waveform with one weak echo (Optech 3100 system).



(b) A lidar waveform with two overlapping echoes, *e.g.*, a hedge (Riegl LMS-Q560 system).

Figure 2.2: Limitation of on-board waveform processing. In both cases, a single mode is found, and inaccurately. The vertical lines symbolize the position of the echo pulse extracted by the hardware system.

2.2 Full-waveform lidar data

Full-waveform topographic laser scanning permits to record the received signal for each transmitted laser pulse. Such sequence is called **a waveform**.

2.2.1 What does a waveform represent ?

A waveform represents the laser backscattered energy as a function of time. Full-waveform lidar systems differ in sampling rate, in scan pattern and in footprint size. Most commercial systems are small-footprint (0.2 - 1 m diameter, depending on altitude and beam divergence) and high pulse rate. They gear to high resolution terrain mapping by providing a high point density and an accurate altimetric description within the diffraction cone. They illuminate a single or a few surfaces within the footprint, yielding waveforms with distinct return pulses corresponding to specific surfaces. A small laser footprint is therefore required to collect a high number of geometrically well defined terrain echoes, such as urban areas. With very small spot sizes in urban areas ($<$0.5 m), the waveforms are bound to be "single-mode", except at discontinuities. Such cases correspond to hard targets such as ground and building roofs (see Figures 1.1 and 2.3). For distributed targets like trees, the number of modes is not foreseeable but may be solved with suitable signal processing algorithms (see Figure 2.4).

In natural areas, the main drawback of such systems with small footprints is that they often miss the top of the trees. It is difficult to determine whether or not ground under dense vegetation has been reached. Consequently, ground and tree heights cannot be estimated accurately [DB00]. Large footprint systems (10-70 m diameter) are intended for characterization of canopy structure. Their advantage is to increases the probability of the laser beam to hit the ground and avoid the biases of small-footprint systems. Moreover, they sample a representative canopy volume in a simple footprint since the return waveform gives a record of the vertical distribution of intercepted surface within a wider

area (Figure 2.3). Consequently, widening the footprint decreases the point density and aggregate the altimetric profiles, providing an average signal.

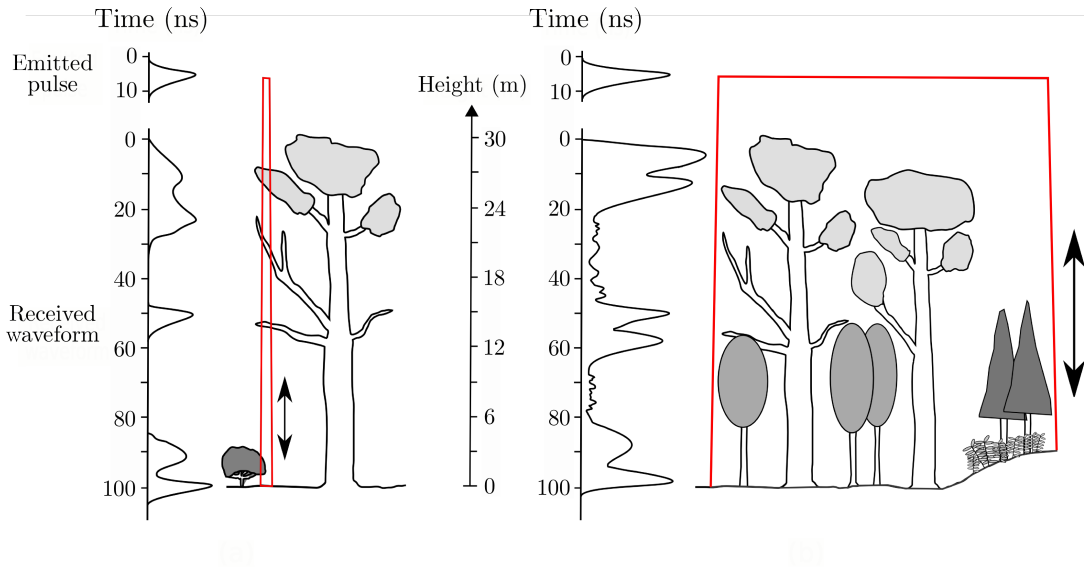
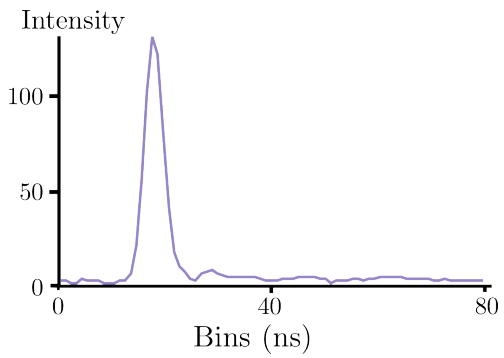


Figure 2.3: The laser beam is displayed in red. **Left**: Emitted and received signals in a forested area with a small-footprint lidar (spot size $\sim 0.5\,\mathrm{m}$). With a small-sized footprint, all targets strongly contribute to the waveform shape, but the laser beam has a high probability of missing the ground. **Right**: Emitted and received signals in a forested area with a large-footprint lidar (size $> 5\,\mathrm{m}$). When considering large footprints, the last pulse is bound to be the ground, but each echo is the integration of several targets of identical range at different locations and with different properties.

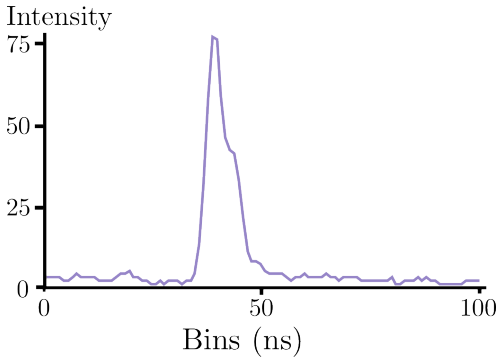
2.2.2 Recording full-waveform data

The two main techniques for recording the signals using an analog-to-digital converter are described in [JS03]. First, the received optical signal (which is the backscattered laser radiation) is collected by the receiver optics and converted to an electrical signal by an avalanche photodiode. In a second time, the output photocurrent is amplified and then digitized using a n-bit quantizer and a sampling period (of 1 ns for commercial devices). The waveform amplitude values are the output of this process. They have no units, and are often expressed in DN (Digital Number).

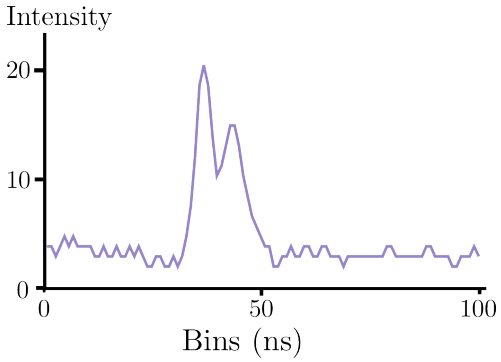
To record the waveform, lidar manufacturers have added digitization terminals to their systems and hard disks with high storage capacity. The waveforms are usually digitized on 8, 12 or 16 bits. The volume of data is bound to very superior to the 3D point cloud over the same area (more than five times for the Amiens dataset, see Section 2.7). The main limitation of surveying areas with a full-waveform lidar system is subsequently the storage capacity. One current issue thus deals with data handling and management since much larger data volume are now recorded.
The digitization sampling period is constant and varies between 1 and 10 ns. To avoid massive storage problems, the entire waveform (*i.e.*, the time history of the laser pulse from the sensor to the target and coming back) is not integrally recorded but only for



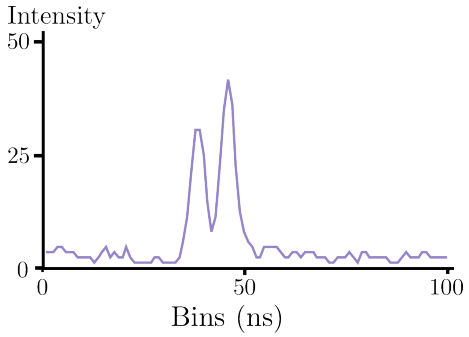
(a) Building roof (one echo).



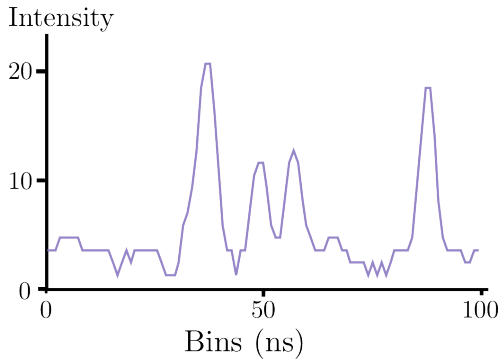
(b) Superstructure (one echo for the superstructure, one for the roof).



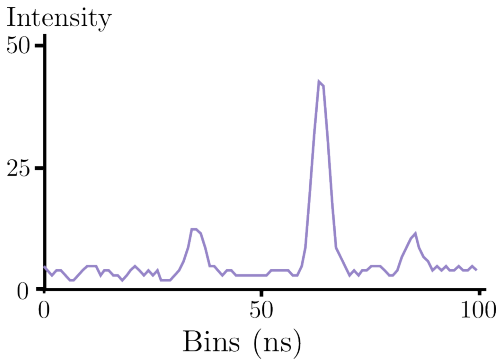
(c) Hedge (one echo for the hedge, one for the ground).



(d) Dense vegetation.



(e) Sparse vegetation (three echoes for the canopy, the last one for the ground).



(f) Vegetation with two main levels (last pulse for the ground).

Figure 2.4: Several examples of small-footprint waveforms over various objects of interest.

a given length after crossing a fixed minimum value. Such recording is ruled by sensor specifications and some of them may be tuned depending on the surveying area:

- A range gate exists, during which the digitizer does not respond to a signal, in order to prevent spurious return signals from the atmosphere from triggering the digitizer (*e.g.*, 100 m for the Optech 3100).
- In the case of a drop-out, the backscattered signal never crosses the threshold and no segments are recorded.
- The record is limited to a predefined maximum number of samples. For example, Optech ALTM systems can store up to 440 samples for each pulse. This is equivalent to a discrete vertical section of 66 m (440 $\times$ 0.15 m per sample). The TopEye MarkII system saves 128 samples according to a predefined mode which is either "*first pulse and later*" (127 samples after the first) or "*last pulse and earlier*". It means that full-waveform systems will not record, within a given waveform, both echoes from the canopy and from the ground, if the trees are taller than the maximum "recording length" of the system.
- To overcome this drawback, systems may assemble several sequences for a single laser shot. If the amplitude of the last samples of the waveform are above a fixed threshold, a new segment of fixed length is recorded. The number of waveform segments thus depends on the terrain. Flat, open terrains produce one segment, while more complex targets may produce more segments. The upper bound is 2 for the Riegl LMS-Q560 system and 7 for the Optech 3100 sensor. In practice, two sequences suffice.
- In addition to the data above the threshold, the digitizer also records a number of samples before the first sample crosses the threshold, and after the last one drops below the threshold.

The digitizer does not detect weaker signals any better than a traditional multi-echo device. In fact, the signal-to-noise ratio (SNR) of the digitized signal is lower because of quantization and other noise introduced by the digitizer. Greater sensitivity can only be realized by averaging multiple waveforms. All waveform data are captured in a single shot with a sensor operating from a moving platform. Thus, this is only possible by post-processing techniques (see Section 2.4.2).

Finally, when a pulse is emitted by the laser, a tiny part of its energy is diverted by a beam splitter onto a photodiode whose signal or start pulse triggers the timing device. It is often called the "emitted/transmitted pulse" or "T0 pulse". The proportion of the real emitted laser energy that is recorded in unknown to an end-user.

2.2.3 Advantages and limitations

We detail here the advantages and limitations of using full-waveform data for a beginner end-user. These conclusions are bound to be objective, and are also based on practical issues that have been noticed when displaying the data, without any assumption on which processes would be carried out on waveforms.

Advantages

- No minimum pulse separation limitation: FW systems overcome the inevitable limitations of the rangefinder electronics *i.e.*, the inability of discrete-return systems to resolve features separated by less than a few meters in the range direction. Overlapping peaks may be found, resulting in denser 3D point clouds. The remaining limitation comes from the digitization rate (Nyquist-Shannon sampling theorem).

- Detection of ground reflections: In presence of ground-clutter and of significant background noise, weak echoes may be missed by sensor hardware algorithms. Waveform processing can enhance the ability to detect such echoes.

- Additional information retrieval: Off-line processing allows to foster information extraction from the raw signal. In addition to peak extraction, a morphological analysis of the waveform modes may be carried out.

- Data accuracy improvement: By designing an appropriate peak detection algorithm, a end-user can improve range estimation, and may qualify the extracted 3D point cloud.

- Metadata availability: For each emitted pulse, the GPS time stamp, scan angle, sensor trajectory, position and attitude, and rangerfinder data (range+amplitude) are available in addition to the emitted and backscattered signals. These information are necessary to generate a 3D point cloud but also offer several opportunities.

- Amplitude data exploitation: Metadata knowledge allows to make the peak amplitudes more homogeneous for the corresponding survey (see Section 2.6.2).

- Sensor geometry processing possible: We call *sensor geometry images* raster data which topology is based on lidar metadata. One axis of the image corresponds to the sensor acquisition angle, and the other represents the scan line number (which may be considered as the acquisition time) [DMB08]. Each pixel refers to a unique lidar waveform (see Figure 2.5). This efficient image-based organization allows a direct access to waveform metadata and neighbours, which is algorithmically interesting and not possible with traditional 3D unstructured point cloud (*e.g.*, image processing algorithms can be used). Full-waveform data can therefore be considered as 3D data with a depth distance.

Limitations

- Lower pulse repetition frequency: Waveform digitizing and recording forces the lidar system to decrease the pulse repetition frequency. High density FW surveys thus require more overlapping strips which has two drawbacks: the survey costs increase and additional strip registrations are necessary.

- Lower Signal-to-Noise Ratio (SNR): The waveform SNR is lower than for discrete-return systems. This can lead to poor signal dynamics and sometimes to holes in datasets. Weak pulses may have amplitude values under the predefined threshold and thus waveforms may not be recorded (see Figure 2.6).

- Entire waveforms? To avoid recording hundreds of useless samples, the recording process is constrained (see Section 2.2.2). For high building edges and vegetated areas without understory, waveform recording may be stopped before the laser pulse reaches the ground since no sample of sufficient amplitude is present in between (the so-called *dead*

zone).

- Huge datasets: Since full sequences of emitted and backscattered signals as well as metadata are now recorded, FW acquisitions lead to five to fifteen times more bulky files that those provided by multiple pulse systems, even in binary formats. A tedious engineering task is therefore compulsory to extract the waveforms and the information to georeference them. [LZT10] have proposed a wavelet-based lossy compression scheme for lidar waveforms, allowing a data reduction rate of 21%.

- Limited acquisition time: With overwhelming datasets, FW lidar surveys are currently limited by hard-drive capacities. A 140 GB hard-drive corresponds to 1.6 hours of data acquisition time at a 50 kHz PRF for the Optech 3100 system.

- No dedicated software: At the moment, there are no standard FW file formats nor dedicated softwares that allow an end-user to visualize in 1D and 3D, process and georeference raw waveforms. The ASPRS Las file format version 1.3 allows waveform data storage but has not yet been adopted. Moreover, lidar companies offer display softwares to their customers, even with processing tools, but they are limited to their own file formats (*e.g.*, Riegl's RiAnalyze). Several free or open-source softwares exist but they are not yet fully operational [CBD+09; MOK+09].

Lidar sensor companies are now making efforts to reach a PRF close to the one of multiple pulse systems. In order to allow end-users to handle, display and easily process FW data, one solution that is now proposed is to directly generate a 3D point cloud by on-board waveform analysis. Additional features are derived by partial waveform modelling that also overcomes the problem of pulse separation of multi-echo devices [UT10].

	Additional properties of FW ALS	**Evaluation**
Advantages	No minimum pulse separation limitation	+++
	Detection of ground reflections	+++
	Additional information retrieved	+++
	Data accuracy improvement	++
	Metadata available	++
	Amplitude data exploitable	++
	Sensor geometry processing possible	++
Limitations	Lower PRF	-
	Lower Signal-to-Noise Ratio	--
	Entire waveforms ?	-
	Huge datasets?	--
	Limited acquisition time	-
	No dedicated software	---

Table 2.1: Advantages and limitations of full-waveform laser scanning.

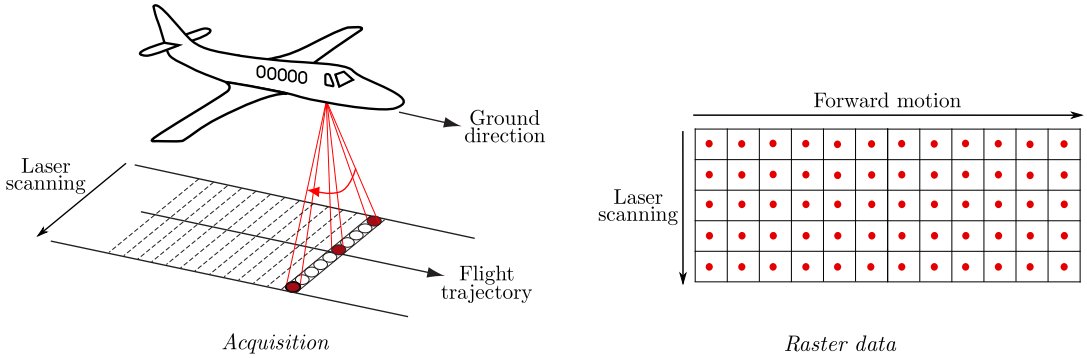


Figure 2.5: Illustration of the generation of the so-called *sensor geometry* (courtesy of Adrien Chauve).

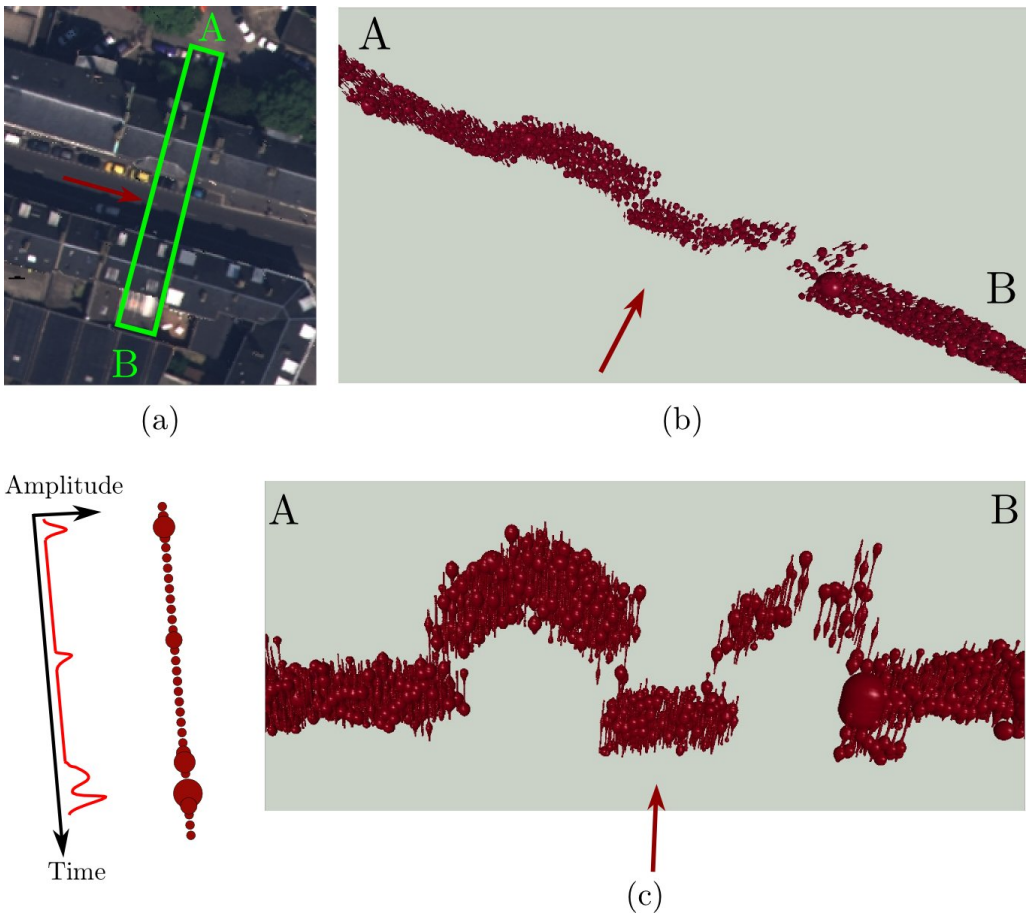


Figure 2.6: In case of the high SNR and low amplitude values, waveforms can exhibit very low values, and therefore will not be recorded: example over roof tops of Amiens. (a) Orthoimage ©IGN. (b) and (c) two points of view of the same section in the waveform data volume. The left roof has been acquired. The right one exhibits holes in one side.

2.3 Typology of full-waveform systems

The first full-waveform systems were designed in the 1980's for bathymetric purposes [GM88]. Topographic devices appeared in the mid-1990's with spatial and airborne experimental systems from NASA. Aerial systems have been commercially available since 2004. Both bathymetric and experimental systems have been developed for specific applications, whereas lidar companies have added digitization terminals without targeting particular purposes. In contrast to the pioneer systems, commercial systems exhibit smaller footprints. They also feature higher sampling rate.

2.3.1 Bathymetric systems

Designed for accurate sea-depth determination, they are composed of two beams, one green, and one infrared. The green beam (532 nm) traverses the air-water interface and propagates in the water until the sea bottom with a minor attenuation. The infrared beam (1064 nm) is reflected by the water and measures the range from the plane to the sea surface. Although the absorption coefficient is high in near infra-red wavelengths, the footprint of the beam is sufficiently large (> 2.5 m) to result in a return from small waves which cause the signal to be reflected back to the receiver (a Raman return is used as a secondary surface detector). Bathymetric waveforms are therefore composed of two peaks. Processing these waveforms consists in finding the two main signal maxima and deriving the range values. We will not further detail bathymetric lidar systems in this dissertation. More information are available, for instance, in [GCRR00].

There are currently several bathymetric lidar systems: **LARSEN-500**, the very first bathymetric system, **LADS** (Laser Airborne Depth Sounder), the **SHOALS** series (Scanning Hydrographic Operational Airborne Lidar Survey), fully operational since 1994, **Hawk Eye**, developed in Sweden on a model similar to SHOALS and **EAARL** (Experimental Advanced Airborne Research Lidar), developed by NASA in 2002. Their main characteristics are described in Section 2.3.4.
Particularly, the EAARL system has been designed to map near-shore bathymetry, topography, and vegetation structure simultaneously [NBWO06]. It can accommodate a large signal dynamic range and has been built to consider significant variations in the vertical complexity of the surface target. Thus, it enables an automatic adaptive acquisition of dramatically different surface types, which suits very well to coastal areas.

2.3.2 Experimental topographic systems

The earliest efforts to use full-waveform digitization topographic lidar systems have been shown by NASA, mainly in forestry research. The following prototypes have been designed to assess the characteristics of woodlands or land cover. They aim at mapping large areas to provide data at a resolution of several meters and a swath width up to 1-2 km. More details can be found in [ST09], Chapter 2.

- Scanning Lidar Imager of Canopies by Echo Recovery (SLICER): It was designed to characterize the vertical structure of the canopy. This medium-sized footprint airborne device demonstrated that full-waveform systems could be used to assess the characteristics of woodlands, distinguish tree ages and species, and characterize the structure

of extensive areas [LHC+99]. SLICER data may be downloaded on line [SLI10].

- Shuttle Laser Altimeter (SLA): This satellite sensor was designed to cover seas, clouds, and land (glaciology, tectonics, hydrology, geomorphology, etc.). Two versions were produced, SLA-01 and 02 (1996-1997), for a feasibility study for the future MBLA and GLAS systems. SLA-02 was used to verify the accuracy of a global 1 km resolution DTM and thus characterize some systematic biases [HGCL99].

- Laser Vegetation Imaging Sensor (LVIS): This improved version of SLICER was used to test and provide data for developing algorithms, calibrating instruments and evaluating the performance of measurements to assess the ulterior Vegetation Canopy Lidar (VCL) mission (*cf.* MBLA system). It also demonstrated the potential of full-waveform data to characterize woodland areas and measure the Earth topography, even below the canopy [BRH99]. For surfaces with similar reflectance, the amplitude of the waveform, corrected by attenuation effects, is a measure of the amount of canopy material. It was mainly used to develop a processing scheme to determine biomass features as well as a real-time algorithm for classifying ground points by analysing the return waveform. Sample data from this system are public [LVI10].

- Multi-Beam Laser Altimeter (MBLA): The MBLA system was part of the VCL mission (Vegetation Canopy Lidar). VCL is an active space-based lidar remote sensing system consisting of a five beam instrument with 25 m contiguous along track resolution. VCL aimed at providing datasets for understanding major environmental issues (climatic change, sustainable land use), and improving global biomass and carbon stocks estimation. VCL's core measurement objectives were canopy top heights, vertical distribution of intercepted surfaces and ground surface topographic elevations [VCL10]. This program was due to be launched in 2003 but was abandoned.

- Geoscience Laser Altimeter System (GLAS): The five year ICESat satellite mission, carrying GLAS sensor, was launched in January 2003 to study the evolution of land and sea glacial masses in the Antarctic and Greenland, the roughness and thickness of sea ice, the topography (using a 1064 nm laser) and the vertical structure of clouds and aerosols (532 nm laser) [Let04; GLA10]. ICESat classifies the return waveform in real-time into land/ice and icesheet/sea by analysing the return waveform and recognizing Gaussian distributions from which the main characteristics are extracted [BZB+03]. GLAS data has also been processed for forest parameter retrieval, hydrological and change detection purposes [USN08]. Datasets are available on the mission Web site [ICE10].

- Forthcoming sensors: The **ICESAT-II** mission (NASA) aims to measure vegetation canopy height as a basis for estimating large-scale biomass and changes. It will use a multi-beam approach, in contrast to ICESAT, and is scheduled for launch in late 2015. The **DESDynI** (NASA) mission combines an L-band Interferometric Synthetic Aperture Radar (InSAR) system, and a multiple beam lidar with 25 m spatial resolution and 1 m vertical accuracy. The goal is to study surface deformations, ecosystems (terrestrial biomass structure), and ice dynamics. DESDynI aims to take advantage of the precision and directness of the lidar to calibrate and validate the InSAR, especially in ecosystem types where field campaigns have not occurred, and to measure the distance between the canopy top and bottom elevation. It should be launched in 2016 or 2017 [DES10]. The **Excalibur** project (ESA) should be operational near 2018 with a lidar system that should sample every 1 m the signal backscattered from the Earth surface over the last 300 m until the ground.

2.3.3 Commercial lidar systems

The first commercial airborne full-waveform laser scanner system was introduced in 2004 (Lite Mapper-5600 based on the Riegl LMS-Q560 laser scanner), without any dedicated application [HUG04]. The manufacturing companies are Riegl (LMS-Q560/680, VQ-380/460), Toposys (Falcon series), TopEye/Blom (MarkII), Optech (ALTM 3100), and Leica (ALS-50II and 60). They have added digitization terminals to their latest sensors and reduced their pulse rate, even if now efforts are made to reach PRF close to multiple pulse systems.
Companies such as Riegl also propose full-waveform systems for terrestrial laser scanning and mobile mapping systems. At the moment, waveform processing is only performed on-board and there are no more digitization [DBS10]. A solution may be to use airborne systems in a terrestrial context [BM07], while preserving eye-safety conditions.

2.3.4 Technical specifications of the main existing systems

The following two tables summarize the main characteristics of the full-waveform lidar systems mentioned above. Commercial sensor specifications will vary with the time, depending on the spring of new sensor generations. For instance, Riegl has recently launched the LMS-Q680 with much higher PRF (up to 400 kHz), thanks to multiple-time-around processing algorithms.

Notes:

►*Final year*: blank if the system is still in use.

►*Wavelength*: when two wavelengths are given (typically 1064 and 532 nm), this means that the system includes two lidar systems, each with its own wavelength. These are either bathymetric applications or satellites with dual coverage (land and sea).

► From *Flying height* to *Range accuracy*: the characteristics are given as ranges of values. These are manufacturer's data and are limited by the system flying height. The formulas for determining the exact values for a given height may be found in [Bal99b].

►*Range accuracy*: this is the accuracy given by the manufacturers after on-board peak detection in the return waveform (telemeter accuracy). It is distinct from the along-track minimum distance between two consecutive peaks and from the altimetric accuracy of lidar data.

► In a cell, "-" means that the information is unknown or not available.

	System	Company - manufacturer	Platform	Beam deflection	Beginning - final year	Wave-length (nm)	Flying height (km)	Pulse rate (kHz)
Bathymetric	**LARSEN 500**	Terra Surveys Optech	airborne	rotating mirror	1983 -	1064/532	0.5	0.02
	MarkII	LADS TopEye	airborne	fibers	1989 -	1064/532	0.37 - 0.5	0.9
	Hawk Eye	Saab Optech	airborne	oscillating mirror	1990 -	1064/532	0.05 - 0.8	0.2
	SHOALS 1000T	US Army Optech	airborne	oscillating mirror	1994 -	1064/532	0.2 - 0.4	0.4
	EAARL	NASA	airborne	oscillating mirror	2002 -	1064/532	0.3	3
Experimental	**SLICER**	NASA	airborne	oscillating mirror	1994 - 1997	1064	< 8	0.075
	SLA-02	NASA	satellite	none	1996 - 1997	1064	285	0.01
	LVIS	NASA	airborne	oscillating mirror	1997 -	1064	< 10	0.1 - 0.5
	GLAS	NASA	satellite	none	2003 -	1064/532	600	0.04
	MBLA	NASA/Univer--sity of Maryland	satellite	oscillating mirror with 3 beams	none	1064	400	0.01/0.242
Commercial	**LMS Q560**	Riegl	airborne	polygon	2004 -	1550	< 1.5	≤ 100
	Falcon II	TopoSys	airborne	fibers	2005 -	1560	< 2.5	50 - 125
	MarkII	TopEye	airborne	Palmer	2004 -	1064	< 1	≤ 50
	ALTM 3100	Optech	airborne	oscillating mirror	2004 -	1064	≤ 3.5	≤ 70
	ALS60	Leica	airborne	oscillating mirror	2010 -	1064	0.2 - 6	≤ 50

Table 2.2: Main technical specifications for full-waveform lidar systems

System	Pulse energy (mJ)	Pulse width (ns)	Scan rate (Hz)	Scan angle (°)	Beam divergence (mrad)	Footprint size (m)	Range accuracy (cm)	Digitizer (ns)
LARSEN 500	-	12	20	30	4	2@500m	30	1
LADS MarkII	7	-	18	27	-	-	15	2
Hawk Eye	2/15	7	0.3 - 7	0/40	2 - 15	1-7.5@500m	30	1
SHOALS 1000T	2/15	6	0.3 - 7	0/40	2 - 15	0.8-6@400m	15	1
EAARL	0.07	1.3	25	22	0.03	0.15@300m	3	1
SLICER	-	4	80	-	2	10@5km	11	1.35
SLA-02	40	8	-	-	0.3	85@285km	150	4
LVIS	5	10	500	14	8	40@5km	30	2
GLAS	74/36	6	-	0	0.11 - 0.17	66@600km	5 - 20	1
MBLA	10	5	-	-	0.06	24@400km	100	4
LMS Q560	0.008	4	5 - 160	45	0.5	0.5@1km	2	1
Falcon II	0.005	5	165 - 415	28	0.7	0.7@1km	0.35	-
MarkII	-	4	<50	14/20	1	1@1km	2-3	1
ALTM 3100	< 0.2	8	<70	50	0.3/0.8	0.3/0.8@1km	1	1
ALS60	< 0.2	5	<90	75 usually	0.22	0.22@1km	2	1/2

Table 2.3: Main technical specifications for full-waveform lidar systems (*second part*)

2.4 Processing the backscatter waveforms

Two approaches have been explored for processing the lidar waveforms (see Figure 2.7). On the one hand, it consists in decomposing the waveform into a sum of components (the so-called *echoes*), so as to differentiate and characterize the different targets along the path of the laser beam. The aim of this approach is to maximize the detection rate of relevant peaks, to generate a denser 3D point cloud and, finally, to extend the waveform processing capabilities by fostering information extraction from the raw signal. Increasing the number of 3D points is of interest for forestry applications (*cf.* Section 2.6), and for retrieving building edges in urban areas. Extracting more information may be useful for segmentation and classification purposes, in both forested and urban areas (see Section 2.6.3 and Chapter 5).
On the other hand, the whole 1D signal is preserved. A spatio-temporal analysis is applied to find features within a 3D waveform space. This approach is suitable for urban areas where the geometry is regular (*cf.* Section 2.6.4).

Most research on full-waveform analysis are focused on the former point cloud enhancement approach. The latter approach has barely been investigated. This section therefore deals with waveform decomposition. First, waveform may be deconvolved from the fluctuating transmitted pulse in order not to be sensor dependent (Section 2.4.1). If advanced techniques focus on extracting strong and weaks echoes and on improving range determination (Section 2.4.2), other approaches model the waveforms using analytical functions (Section 2.5) subsequent to the decomposition process.

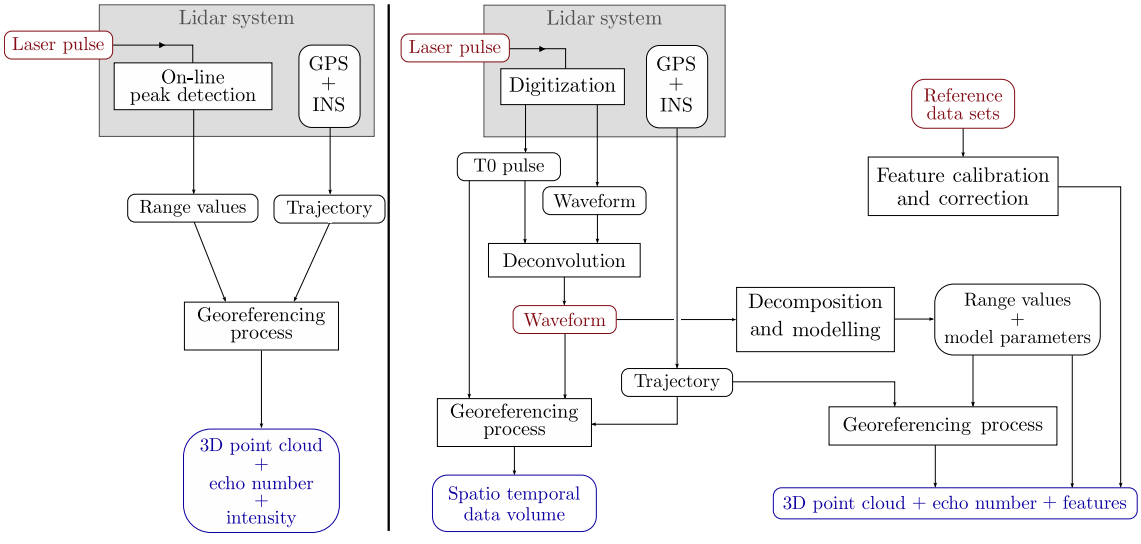


Figure 2.7: From data acquisition to final products: comparison between (a) Multiple pulse systems, and (b) Full-waveform sensors.

2.4.1 Deconvolution

When the transmitted pulse is known, deconvolution can be carried out to remove the influence of its shape on the recorded data. The main issue is that this is an ill-posed problem if no *a priori* knowledge is introduced to constraint the problem. Standard signal processing approaches, such as the Fourier-quotient method, therefore yield unsatisfactory results. The following paragraphs present the approaches that have been proposed to address this problem in the lidar waveform context.

Wiener filtering

Based on a physical understanding of the pulse propagation and its interaction with the illuminated surface, [JS06] propose a relevant algorithm to discriminate different surface responses which are very closely located in range (< 0.15m). First, the received waveform of the backscattered pulse is computed using the lidar equation. It depends on the transmitted waveform (modeled by a Gaussian function randomly modulated by a Gaussian noise), on the spatial energy distribution of the emitted pulse (which models the laser device), on the surface response (with given reflectance and geometric properties, here with two differently elevated specular plan plates), on the atmospheric transmission, and receiver efficiency. Echoes are first detected within the received waveform with a noise dependent threshold. Then, deconvolution in the Fourier domain of the transmitted waveform with the received waveform is carried out, and the surface function is estimated using the Wiener Filter (the Wiener Filter is a real function estimated from the modulated transmitted waveform and the background noise, it is mainly applied here for noise reduction in the frequency domain). Finally, waveform fitting is performed with the Levenberg-Marquardt technique (see Part 2.5), using the surface function as the modelling function. Experiments on different kinds of surfaces show that surfaces with a distance corresponding to less than 0.15 m may be resolved.
An alternative to the Wiener filter may be the Richardson-Lucy algorithm [Nor06]. The author reports that the latter one converges slower, is much more computationally heavy, and may amplify the noise during its iterative procedure.

Expectation-Maximization based approach

An Expectation-Maximization (EM) based deconvolution strategy adapted from [FN01] is adopted in [Par07; PN09]. The problem is formulated as a Bayesian estimation problem where the maximum a posteriori (MAP) of a given energy function is estimated. This function incorporates prior knowledge of the waveform and does not require any assumption on the waveform noise (explicitly or implicitly as for the Wiener Filter approach). The iterative EM algorithm is used to find a close solution of the MAP estimator. One salient feature of this method is the provision of a parameter that tunes the trade-off between a high number of detected peaks and a high probability of false detection. This method was developed for airport obstruction surveying, in order to improve object detection (such as poles, antennas, etc.).

Least-squares method with regularization

[WZR+09] have proposed a method for deconvolution in the time domain. The three variables (emitted and received waveform + backscatter cross-section – BCS) are considered as piecewise linear functions that allow to turn the deconvolution problem into the resolution of a system of linear equations with a least-square method. Moreover, due to the presence of noise, the solution may be unstable and a regularization technique is necessary. A smoothness constraint is therefore applied on the least-square problem. This technique is efficient but requires a good knowledge of the noise level within the data. Results on synthetic simulations and real waveforms from various sensors are shown.

B-splines

A method based on uniform B-splines has been proposed by [RBP10]. This also allows to solve the problem in a linear least-square approach, without any initialization step and assumption on the shape of the emitted pulse and the scatterers. Indeed a B-spline of degree 3, corresponds to a function with continuous curvature ($\mathcal{C}^2$ class). The emitted and received waveforms are modelled as B-splines. The BCS can therefore also be estimated as a B-spline curve. Fitting a B-spline curve to observations is a linear least-square problem. The authors show good results both on synthetic datasets with known cross section and real waveforms.

2.4.2 Advanced echo extraction methods

The main reason for decomposing the waveform is to extract more points in a more reliable way. Several methods have been carried out so far. Waveform fitting algorithms, described in Section 2.5, also permit to find peak location and echoes undetected by conventional multiple pulse systems due to hardwired internal thresholds for peak detection. However, they often require to model the echoes with an appropriate analytical function.

Correlation-based methods

[TWH05] detect pulses using both pulsed and continuous wave ranging systems. They have noticed that a peak detection using a single threshold leads to constant errors that depend on the SNR, whereas an adaptative threshold may be used when the emitted pulse is known. Measuring the correlation between echoes coming from both systems allows a significative noise reduction and improves peak detection.
For estimating the number and position of echoes within the waveforms, [RWMU08] have adopted the Average Square Difference Function (ASDF) method. This is a time delay estimation technique based, here, on the correlation between the emitted laser pulse and the recorded waveform. A function is defined to describe the shift between the two waveforms. The single echoes are located at the minima of this function, that can either be local or global minima, since multiple echoes may appear.

Improved range determination

For urban landscapes, [KJS08] propose an iterative method to improve the point density, the range accuracy as well as the segmentation between partially penetrable objects and impenetrable surfaces. The goal of this approach is to fill gaps that can appear in partly occluded surface regions. Assuming that the laser beam hits a planar surface with a given slope, the surface response is modelled (transmitted and received waveforms are known) using a Matched Filter [Tur60]. It produces range values and generates a 3D point cloud. Points are segmented according to a given feature and those expected to belong to impenetrable surfaces are used to estimate surface primitives. A new surface response is finally computed and used as prior knowledge at the beginning of the algorithm.
This method (as well as the deconvolution approach) allows to determine the range of each echo without any pulse shape assumption, and to detect weak echoes corresponding to partially occluded and partly illuminated regions.

Detection of weak and overlapping pulses

[SYJ07] show it is possible to detect weak pulses corresponding to partially occluded targets or objects with poor surface backscatter properties. A waveform stacking technique is performed by establishing neighbourhood relationships between consecutive waveforms. Mutual information is therefore accumulated to produce a "global" scattering for such targets. This technique predicts new echoes left undetected by standard algorithms. The method has been applied on terrestrial waveforms but would be efficient as well on airborne data.
[LMSV10] have also tackled this problem with a special emphasis on overlapping peaks. They have developed an algorithm that detects echoes within pulse shapes that exhibit asymmetric or non Gaussian behaviours (*e.g.*, low above ground vegetation). They aim to find inflexion points on both sides of primary visible peaks, with sufficient amplitude and distance to the main peak location. Then, a Gaussian fit is applied and missed peaks are iteratively retrieved in the residuals maxima. Finally, a qualitative analysis on motorway barriers shows that overlapping pulses may be resolved very well and thus close target differentiated ($<$0.6 m in this case).

2.5 Waveform modelling and echo fitting

The advantages of waveform modelling are three-fold. First, as for methods described above, new echoes are found. On forested areas, it can, for instance, provide up to 100% more pulses than a real-time system [RKS08; CVB+09]. Secondly, waveform processing improves object range determination, even over complex surfaces [ZSA+02]. For instance, in forested areas, both canopy and ground height estimates may be improved [DPLV08] but this result depends on the survey specifications and the landscape [CVB+09]. In urban areas, range accuracy of solid opaque targets can also be improved [KJS08]. Thirdly, modelling the echoes provide additional parameters that may be useful for classification purposes (see Section 2.6.3).

When modelling the echoes within a waveform, a parametric approach is always chosen in

current state-of-the-art approaches. Parameters of an analytical function are estimated for each detected peak in the signal. These parameters provide additional information about the target characteristics (shape and reflectance) and extend the waveform processing capabilities. Statistical elements extracted by signal processing techniques are the number of significant peaks, their range to the sensor, and the parameters of their modelling function. A single function is always used to model all echoes of the waveforms.

One wishes to decompose a waveform $y = f(x_i)$ into a sum of n components:

$$y_i = \sum_{k=1}^{n} \phi_k(x_i) + b \tag{2.5.1}$$

where ϕ_k the analytical function with a set of parameters $\mathbf{\Theta}$, f is the waveform model ($f = \sum_k \phi_k$), $\{x_i\}_{i=1,..,N}$ is a sequence of N uniformly-spaced points, $y = \{y_i\}_{i=1,..,N}$ is the sampled waveform, and b is the background noise.
A relevant echo model is particularly suitable so that related parameters should be used for segmenting and classifying the 3D point cloud. A large body of literature addresses the issue of echo fitting with a given parametric model. The two main steps and associated references are described below.

2.5.1 Modelling the waveforms with Gaussian mixtures

A waveform is the convolution between a laser transmitted pulse (assumed to be of Gaussian shape with a calibrated width) and a "surface" scattering function (backscatter cross-section – BCS), often considered as a Gaussian function [WUD+06]. The received signal is then assumed to be a mixture of Gaussian distributions. The Gaussian function is then the unique adopted function to model full-waveform data. Its analytical expression is:

$$\phi_k(x) = A_k \exp\left(-\frac{(x-\mu_k)^2}{2w_k^2}\right) \tag{2.5.2}$$

where μ_k is the pulse location, A_k is the pulse amplitude, and w_k is the pulse width. Thus, $\boldsymbol{\theta}_k = \{\mu_k, A_k, wa_k\}$.
The Gaussian model is sufficient for most applications, especially for large-footprint lidar data [ZSA+02; WUD+06]. It has been noticed for GLAS data that the Gaussian decomposition method is inaccurate when a waveform is deformed due to atmosphere forward scattering [BZB+03]. Besides, for small-sized and medium-sized footprints, this model is not always justified. In urban areas, many peaks are distorted: indeed, most of the return waveforms are subject to the mixed effects of geometric (*e.g.*, roof slopes) and radiometric object properties (*e.g.*, different kinds of streets and roof materials). Hence, the characteristics of return peaks may differ significantly.

2.5.2 Fitting the waveforms

Two main methods have been carried out to fit the waveform with a single modelling function: Non-linear optimization techniques (*e.g.*, using a least-squares approach or the Trust Region algorithm, Maximum likelihood estimate with *Expectation-Maximization* algorithm.

Non-linear least-squares approach

[HMB00] provide a general description of non-linear least-square methods which are used in bathymetric [WA91], satellite [BZB+03], terrestrial [JS06], and airborne laser scanning systems [DPLV08; RKS08].
This problem is a system of N observations with $m \times n$ unknown parameters. m=card $\boldsymbol{\theta}$ is the number of parameters of the modelling function and n the number of echoes. n is estimated using standard peak detections methods or approaches described in Section 2.4.2. The quality of the results is evaluated by a variable ξ. One aim at fitting the data with a prescribed accuracy ϵ.

$$\xi = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f(x_i) - y_i)^2} \; < \; \epsilon \tag{2.5.3}$$

The system is solved using a non-linear least-squares method, the Levenberg-Marquardt (LM) technique [Mar63]. There is no algorithm for solving the problem directly since the $\{\phi_k\}_k$ functions are not linear. The LM algorithm is known to be robust but requires a good initialization step. Further details are provided in Chapter 3.
Initial values are provided by traditional pulse detection methods, often improved to take into account weak and overlapping pulses undetected by such methods. For instance, to overcome the problem of a wrong initialization step, [HMB00] add progressively peaks in the least-squares fitting algorithm according to their amplitude (from highest to lowest peaks) until ξ is greater than a given threshold.
For applying constraints on parameter estimates, [LMSV10] prefer using the Trust Region algorithm as non-linear optimization technique.

Maximum likelihood approach: the EM algorithm

[PSTA05] have developed a pulse detection method based on the Expectation-Maximization algorithm (EM) [DLR77]. The EM algorithm is a two-stage iterative optimization technique for finding maximum likelihood solutions. EM alternates between performing an expectation (E) step, which computes an expectation of the likelihood, and a maximization (M) step, which computes the maximum likelihood estimates of the parameters by maximizing the expected likelihood found on the E step. It consists in computing, for each sample, the probability of belonging to one of the k^{th} distributions which decomposes the signal. Parameters found during the M step are then used to begin another E step, and the process is repeated. They assume that the return waveform is a sum of Gaussians. Nevertheless, it is possible to choose functions to fit the return waveforms. Since there the number of echoes is not known beforehand, a model selection step is required. A modified version of the Akaike's Information Criterion (AIC) has been implemented to both take into account the closeness of the reconstructed signal to the waveform and a reconstruction complexity term.

2.6 FW point clouds: quantitative analysis and processing

The after mentioned studies made great strides towards demonstrating the relevance of FW lidar data. Most of them are subsequent to the decomposition and modelling steps.

2.6.1 Pulse shape analysis

The elements which could modify the shape of the waveform have been studied so far by several authors. These works have been carried out mainly with simulated or close-range data. The FOI (Swedish Ministry of Defence Research Institute) proposes a comprehensive laser data simulation model. It includes the detector characteristics, the target geometry and reflectivity, the atmospheric attenuation, etc. The main assessment is that the influence of target geometry and radiometry are mixed and cannot be decorrelated simply.

- Emission angle: variations in the shape are noticed with changes in the angle of incidence. The smaller the angle, the narrower and more symmetrical the peak.

- Roughness: different responses are noticed between smooth surfaces (one or two echoes if there are discontinuities) and porous surfaces (multiple echoes at different depths, equivalent to the behaviour of trees and vegetation) [VAM04]. [HH10b] have noticed that the pulse echo width allows to derive a surface roughness criterion similar to a standard roughness criterion computed from a much more denser point cloud.

- Target geometry: [KKL07] measured for various urban materials the dependency of the amplitude from the incidence angle. Furthermore, [Ste00] and [CSL01] have carried out experiments on targets with simple shapes to study the combined effect of the geometry and target reflectance on the return waveform. A significant spread in the pulse and a decrease in its amplitude are observed over flat surfaces when the incident angle increases. It may lead to an erroneous estimation of the distance, especially with significant incidence angles. Besides, similar waveforms are observed between plane surface and corners. Consequently, two plane surfaces can be assimilated to a single one. Eventually, simulations were carried out on step targets to investigate whether it was possible to resolve two close planar surfaces, and on different canopy shapes.
A comparison of waveforms reflected by various shapes shows the potential advantage of classifying objects by analysing the full waveform from a single pulse. However, it must be pointed out that these conclusions are based on the major assumption that the target exhibits a Lambertian reflectance.

- Geometry + radiometry: [JS03] set up experiments to assess the effect of various urban materials on the return waveform. Images of range, amplitude, and pulse width are calculated for pebbles, corrugated iron, slanted slate plate and flat roof tiles. The four types of roof material tested behave in different ways. This confirms that it is not possible to classify waveforms simply as vegetation/buildings/roads. Similar responses are recorded between objects of different classes. Besides, there is significant deviation within the same class. This study is similar to that described for example in [LMBM05] on the study of the variation in reflectance of various urban materials (asphalt, tar, concrete, granite, etc) using very high resolution aerial images.

- Range: the received power is function of $1/R^2$ (see Equation 2.1.11) where R is the distance from the sensor to the target. The survey specifications and the local topography have thus an impact on the received waveforms.

- Other factors: [CSL01] have noticed that the influence of the atmosphere and specular effects on the target are negligible on the range measurement compared with the other effects.

2.6.2 Calibration and correction of the intensity feature and the target cross-section

The intensity of the returned pulse has been found to be of low quality and to be highly dependent on the incidence angle, and thus, in the past, it was concluded to be of little use. However, with the spring of FW sensors, improvements are possible.
Intensity is not yet a clearly defined term [WHU+08]. The echo amplitude is most commonly referred to as *intensity*. However, the intensity should be associated to the total energy of the echo (*i.e.*, $I = \sqrt{2\pi}\, A\, w$ for a pulse of Gaussian shape). In this dissertation, amplitude will refer to the value of the echoes at the maximum location and intensity will refer to the echo energy.

The echo amplitude depends on many factors: target characteristics, lidar system, scan geometry, etc. Fluctuations can be noticed on large datasets between surveys, for instance, due to different atmospheric conditions, and even between flight strips. The intensity/amplitude values provided by commercial lidar systems as well as those extracted from waveforms processing are neither calibrated nor corrected. Few studies have been carried out so far on intensity calibration and correction. They are described below. They aim at converting intensity to a relative but comparable measurement for different epochs with different conditions, *e.g.*, for multi-temporal analysis or classification. Three main methods exist: calibration using artificial reference targets or natural targets, and correction of specific influencing factors.

Calibration

The Finnish Geodetic Institute (FGI) has first proposed methods using reference targets with calibrated reflectance to study the effect of the albedo on the shape of the return waveform and their directional properties [KAHS05]. Results show that reflectance plays a predominant role on the amplitude and width of the peak. The influence of the albedo cannot be separated from other target properties. However, the calibration protocol makes it possible to use intensity as data in its own right. Besides, a laboratory study was carried out on various targets [KHM05]. The materials tested are sand with various grain sizes, asphalt, trees, lichen and moss with small angle variations. Calibration is possible but it proves to be difficult to define a reliable method to classify into various components based only on the intensity characteristics. The authors recommend studying much larger samples in the hope of drawing up an initial classification methodology.

Calibration may also be performed using the lidar equation [WUD+06; WHU+08] and natural targets of known reflectance. For that purpose, the echo amplitude and width are needed over reference targets. Therefore, only full-waveform sensors allow an accurate calibration for all target classes. A Gaussian behaviour of the emitted pulse and the scatters are hypothesized. We have for a given target $\sigma' = C_{\text{cal}} R^4 \hat{A} \hat{w}$, where $\hat{A}$ and $\hat{w}$ are respectively the estimated echo amplitude and width. Asphalt streets are considered as Lambertian targets with reflectance ρ_{asphalt}=0.2. Thus the calibration constant can be estimated, assuming a high stability of the laser pulse and constant atmospheric conditions:

$$C_{\text{cal}} = \frac{\pi\, \rho_{\text{asphalt}}\, \beta^2}{R^2\, \hat{A}_{\text{asphalt}}\, \hat{w}_{\text{asphalt}}} \tag{2.6.1}$$

It leads to the following calibration formula for target cross-section:

$$\sigma' = \frac{\pi \, \rho_{\text{asphalt}} \, \beta^2 R^2 \, \hat{A} \, \hat{w}}{\hat{A}_{\text{asphalt}} \, \hat{w}_{\text{asphalt}}} \tag{2.6.2}$$

This process is improved in [BHL+08] by measuring accurately the backscattering properties of natural surfaces with a reflectometer, and using Spectralon targets from homogeneous reference targets.

Correction

Correction techniques have been developed to adjust amplitude variations due to the influence of varying atmospheric conditions, flying altitude, and topography [CS06].
[AKHS06] show a strong correlation between the reflectance of surfaces and intensity for any flight height. Thus, the intensity value has to be corrected according to the scanning distance, the angle of incidence, the atmospheric transmission, the attenuation and power transmitted by the system. Certain flying heights are preferred (from 200 to 1000 m) for a suitable ground spot, with scan angles less than 10^{o}.
Two methods for intensity correction are proposed in [HP07]. A data-driven correction allows to globally correct datasets with different flying heights thanks to an empirical mathematical model. The function parameters are estimated using extended targets. Another correction is possible point by point, based on the lidar equation, assuming the same knowledge as in [AKHS06].

A number of hypothesis are made in these papers. Using full-waveform lidar data may overcome several issues of amplitude correction [HP07].

- Target geometry: First, the angle of incidence θ between the laser beam and the surface is accurately known, and the surface slope can be better estimated than with a conventional 3D point cloud [KJS08]. The apparent surface of the target in the sensor direction decreases with a $\cos\theta$ factor if the target is assumed to be an ideal Lambertian scatterer.

- Range factor: The range effect is corrected by applying a R^2/R^2_{strip} factor on the intensity value, where R^2_{strip} is an arbitrary value set for the whole survey area. However, it is shown in [PDHF07] that this range dependent inverse-square model might be insufficient to estimate the accurate amplitude.

- Pulse fluctuations: Since the emitted pulse is recorded, it permits to take emitted pulse intensity fluctuations into account. It is no longer needed to assume it constant, and fluctuations up to 10% have been noticed between scan lines. This may be corrected by applying a normalizing factor which is the ratio between the mean amplitude of all the emitted pulses ($\overline{A_e^{\text{strip}}}$) and the estimated amplitude of the current emitted pulse (A_e) [BCB+09].

The final amplitude correction formula is:

$$A_{\text{final}} = \frac{\hat{A}}{\cos\theta} \frac{R^2}{R^2_{\text{strip}}} \frac{\overline{A_e^{\text{strip}}}}{A_e} \frac{1}{T^2} \tag{2.6.3}$$

Where T is the atmospheric transmission (assumed constant). The impact of amplitude correction over a dense urban area may be noticed on Figure 2.8.

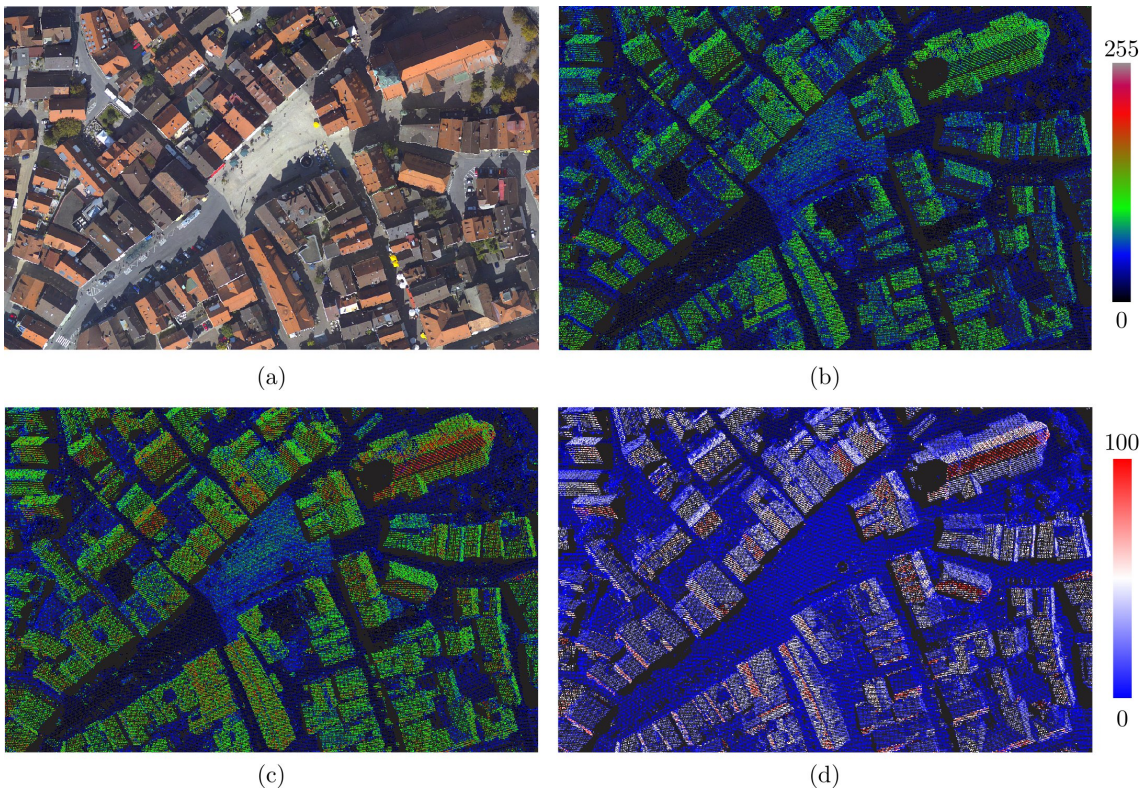


Figure 2.8: Amplitude correction based on geometric and full-waveform information (Biberach). The trajectory of the plane goes from left to right, close to the northern part of the church. (a) Orthoimage; (b) Amplitude before correction; (c) Amplitude after correction; (d) Difference between both values.

[JG09b] have noticed that the intensity inside a region exhibit a significant variance even for a constant incidence angle. This may be due to material features or local surface effects. The latter ones are difficult to model. The authors have investigated the use of the Phong model instead of the traditional Lambertian model (specular characteristics in addition to the diffuse backscatter properties), but no improvements have been yet noticed, mainly because of the lack of various roof materials in the dataset and reference targets with known backscattering characteristics, and because of the low variability of the angle of incidence.

2.6.3 Parameter behaviour and subsequent classification

It is possible to define some particular behaviour for several objects. The echo is wider on the canopy or ploughed fields compared to roads or meadow areas. High amplitudes are noticed on grass and bare earth and variable amplitudes on the roofs of buildings, depending on the roof materials. Moreover, weak echoes are wider and the intensity of additional extracted points is, predictably, lower than for the first echoes. FW datasets thus allow the gathering of relevant features for segmentation and classification applications.
However, a point cloud obtained by one of the methods described in Section 2.5 cannot be classified categorically, even with additional features and hints about the knowledge of

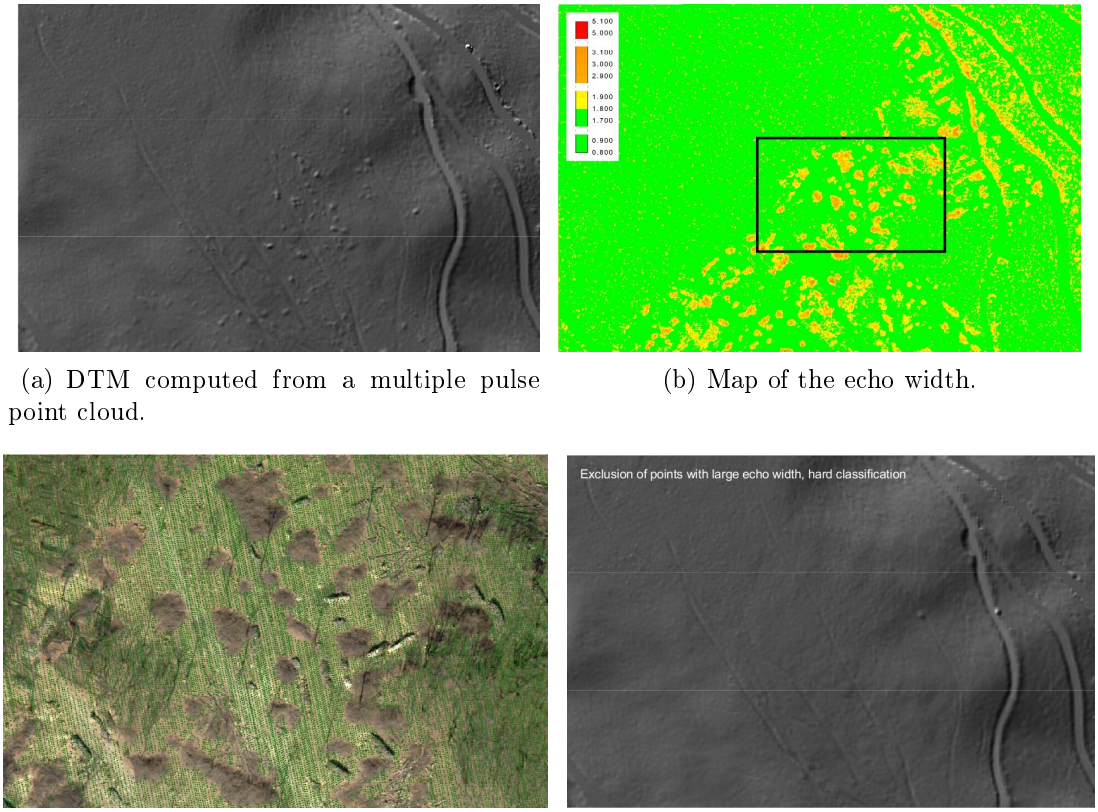
the influence of each object. A wide echo with low amplitude does not necessarily come from vegetation. Roads and building roofs are made of different types of material and, therefore, have different characteristics. The features currently extracted from a modelling step often have values similar to natural objects [GJT07]. Consequently, simple classification algorithms, *e.g.* classification tree with empirical thresholds, lead to a high rate of incorrect classification [DPL06; DHU+06]. A reliable approach to detect vegetated areas is proposed in [GJT07], based on the eigenvalues of the covariance matrix computed for each point with the intensity values in a cylindric environment. However, geometric features are used in addition to waveform features for the discrimination. Other features and methods have to be found to classify full-waveform point clouds. Attributes such as the BCS or the backscattering coefficient, which describe the scattering properties of the targets, seem valuable to retrieve these physical characteristics and improve segmentation/classification results [WHU+08].
Moreover, in order to exploit pulse shape parameters, we have to bear in mind that the shape of the n^{th} pulse of a waveform depends on the scattering characteristics of the $n-1$ first echoes. It is known that the cross-section of the n^{th} target is expected to decrease depending on its "rank" in the waveform, but it is not straightforward to predict how [WHBD08].

The problem of vegetation discrimination is also tackled in [RHHP08], in an object-based approach. The 3D point cloud is first slightly over-segmented using a seeded region growing algorithm based on the echo width. Then, for each segment, basic statistics (minimum, maximum, standard deviation, etc.) are computed for the selected point features: amplitude, echo width and geometrical attributes. These features are fed into a Classification Tree algorithm with distinct complexity parameters. It allows to select automatically the best discriminative features and the associated thresholds, and to reach accuracies slightly superior to those reported in [DHU+06] (>90%). An improved echo ratio feature is computed in [HH10a] allowing a better vegetation discrimination with lower computational point cloud analysis (especially when dealing with very high point densities, *e.g.*, 50 pts/m^2). Finally, the rule-based classification is refined using, in particular, echo amplitude and width to discriminate non vegetation objects such as building walls, roof edges, and power-lines.

[NMT09] evaluate the discrimination performance of the Gaussian parameters and waveform metrics (see Section 2.6.5.1) for land cover classification. With a supervised Bayesian pairwise classifier, eight features are used to discriminate seven classes (grass, trail, dead trees, oaks, junipers etc.). The method allows to select class dependent features based on their incremental contribution to a given relevance function. All the pairwise classifiers are computed and the ultimate class is selected within the outputs based on a maximum Bayesian posterior probability rule. The most frequently selected features are the peak amplitude and the energy ratio, and the rise time and fall time of the waveforms are also discriminant for different tree types and densities. The classification results are compared favorably to the same workflow based on Quickbird multi-spectral high resolution images. The main reason is that the lidar-based input data are structural parameters and these features improve the classification of labels that are spectrally similar but structurally different, such as trees of different ages.

2.6.4 Filtering and surface reconstruction

The pulse width is an important FW feature since it holds information on the vertical spread of the object surface. For small-footprint systems, the echo width is mainly affected by the presence of low vegetation (ferns, small bushes etc.) and low objects lying on the ground (street items or hedges). They feature a pronounced height variation within a footprint. Therefore, objects closely located in range will contribute to a single echo but with an echo width superior to the emitted echo width. Conversely, this is not influenced strongly by the terrain slope because the height differences due to slope are typically much smaller than the length of the pulse (a 4 ns pulse corresponds to 1.2 m in height). Echo width can therefore be used to discriminate echoes that cannot originate from the ground. The following authors have consequently improved Digital Terrain Model (DTM) algorithms by taking this property into account.



(a) DTM computed from a multiple pulse point cloud.

(b) Map of the echo width.

(c) Green: final points conserved for the new computation.

(d) Final DTM, using FW information.

Figure 2.9: Improvement in Digital Terrain Model generation using the echo width extracted from a Gaussian decomposition of full-waveform data [DBFJ08].

[DBFJ08] look after archaeological features under vegetated areas. They remove from a last echo point cloud all points that exhibit an echo width superior to an empirically-based threshold. It corresponds in practice to the understory vegetation. Then, a traditional filtering process is carried out on the remaining points. This step improves the computed DTM by removing small bumps that are not subtle topographic changes (Figure 2.9).

However, a pre-elimination step is not mandatory. Multiple-pulse DTM algorithms have also been improved to wrap the echo width as ground/off-ground discriminant feature. For instance, [LM09b] introduce the width feature into the progressive TIN densification approach developed in [Axe00] (also used with FW data in [LM09a] for archaeological purposes). They first label the point cloud depending on the echo width: class "Overlapping pulses" correspond to close echoes where width estimates are unreliable: class "Type B" is rough surface pulses with a width superior to a given threshold and class "Type A" corresponds to smooth surfaces (width inferior to the same threshold). This class is split into "Open smooth" and "Obscured smooth" points depending on their position within the waveform (single or multiple returns). Such labelling allows a target specific tuning of the parameters of Axelsson's algorithm. Finally, the authors demonstrate that their approach is able to remove low vegetation and may correct local under and overestimation of the terrain surface. The "Type A/Type B" threshold has been determined in [LM10], where a ground/off-ground point segmentation is subsequently performed. The Decision Tree segmentation is refined by regularizing the point label depending on the labels of the neighbourhood points. The authors also noticed that pulse width must be used with caution for weak and very strong amplitude echoes.

[Müc08; MBH10] prefer introducing a priori weights into the hierarchical robust interpolation algorithm presented in [KP98]. The aim is to attract the terrain surface on points that have a high probability to belong to the ground. The method has the advantage of keeping all the lidar points for the DTM generation process (thus the false negative ground points may still influence, even weakly, the surface). The weights are set up by fitting a distribution function on the histogram of the pulse width with respect to the pulse amplitude. The results show that it both improves and fastens the process.

2.6.5 Applications in forested areas

Full-waveform lidar data have been widely used for forest analysis. The waveforms are decomposed to produce dense 3D point clouds in the canopy which are then used mainly to estimate forest parameters at the scale of the stand. Most of the literature on full-waveform systems deals with this topic. Firstly, one tries to benefit from a denser point cloud to improve forest parameter estimation. Indeed, it is especially beneficial in biomass and forest structure estimation techniques, where conventional multiple pulse methods may be biased due to their limited understory tree detection capabilities. Secondly, using jointly lidar geometric information and estimated radiometric properties (backscatter cross-section) has been tackled for tree segmentation and species classification. Finally, modelling is performed to understand the influence of forest parameters on the waveform shape.

2.6.5.1 Estimating forest parameters

Many studies have already been carried out to estimate forest parameters using multi-echo lidar data: high point density can be used to extract trees in small areas, their height and crown diameter [PHS02], their volume, to classify them according to species [HP04], to estimate their particular characteristics [AMR05] and even to measure the growth of the

forest and detect trees that have been felled [YHKM04]. Woodland parameters may be estimated at large scale: density of population, coverage, biomass, etc [MAH+99].

Multi-peak waveforms have a longer rise time due to the energy scattering and pulse broadening. Rise time corresponds to how quickly energy reaches the 90% amplitude on the leading edge. For single Gaussian waveforms, it is proportional to amplitude. Thus, for large-footprint systems, rise time is correlated with the amount of crown variability and crown depth. [DB00; LCPH02; RAM05] present the main studies with multi-echo and full-waveform airborne lidar for forestry applications, pioneered with NASA experimental sensors. Full-waveform lidar metrics are used to estimate the following woodland parameters:

- **Canopy height**: modeled from the measurement of the difference between the height of the first and last echoes, for different types of forest (temperate, boreal and tropical), at the tree or the stand levels [LCA+99; KRSB06; RNSL08]. It is generally underestimated by at least one meter.
- **Vertical distribution of canopy material**: essential to determine other canopy features such as the above-ground biomass, predicting the state of the forest and determining the age of a plantation [LCA+99].
- **Canopy height profile**: directly derived from the vertical canopy distribution for any deciduous forest. It defines the occlusion rate of a plantation [HLP01].
- **Canopy cover**: obtained directly. This is the fraction of the signal reflected by the target corrected by the estimated ground reflectance [MAH+99].
- **Canopy volume profile**: obtained by modelling. It can show the qualitative and quantitative differences between different ages of a given species [LCA+99]. It may also provide information on vertical leaf profiles [HLP01].
- **Above-ground biomass**: modeled from the tree height measurements. This correlation was shown in mixed coniferous/deciduous areas, in mountains and in dense boreal and tropical forests [DDC+02; HDP+05].
- **Basal area**: cross sectional area of the trunk, at DBH (Diameter Breast Height, *i.e.*, 1.37 m) [LHC+99; MAH+99].
- **Mean stem diameter**: tree height is strongly correlated to the stem diameter [DDC+02]. Allometric equations allow to derive the stem diameter according to the canopy height and the tree species.
- **Crown and stem volume**: these features are inferred. The crown volume (tree parameter) is computed knowing the canopy volume (stand parameter), the tree density and species. The stem volume is inferred according to the mean stem diameter and the tree height.
- **Waveform metrics**: canopy energy, ground energy, total waveform energy, ratio between canopy and ground energy, rise time to the first peak, fall time of the last peak, and height of median energy are computed in addition to Gaussian features in [NMT09] to discriminate between tree species.
- **Other parameters**: the density of large trees can be inferred. Other additional data (thermic, optical, radar) is required for a satisfactory determination of the Leaf Area Index (LAI) and to have a knowledge of the tree species. It appears possible to classify species [RKS08] but a reliable algorithm has yet to be implemented.

Certain variables are obtained using allometric equations and vary according to the type of forest and their main characteristics [HDP+05]. Some parameters cannot be obtained or

with low accuracy. It is therefore difficult to obtain a comprehensive estimation of forest parameters (even for the main parameters such as the tree height and the crown diameter), and consequently parameters inferred as well as general relationships between structural forest variables for a given type of biome (and *a fortiori* for all types). Metrics derived from full-waveform data are not always significantly correlated with forest structural characteristics at the tree level, even if it works well for some forest types [AMS+06].
First works dealing with forest change parameters have been carried out with large-footprint lidar data [DPLV08]. Differences, in maximum canopy height, for instance, (0.5 m), have been noticed between summer and winter data.

Finally, forest parameter retrieval on steeped areas has to be mentioned. The height of the upper layer of the tree canopy, located in thalwegs, for instance, can be the same as the higher ground parts of the terrain. In case of erroneous DTM, several forest parameters, such as the canopy height, crucial for modelling and inferring other features, will be incorrect. That is the reason why waveform features have been introduced to improve ground/off-ground segmentation algorithms and derive more reliable DTMs. [NRSL10] state that the Gaussian decomposition of the large-footprint GLAS waveforms is a reliable approach for identifying the ground surface within the waveforms. Moreover, an increasing slope leads only to a minor negative offset in ground estimation, which assesses the relevance of large footprints for coarse DTM generation.

2.6.5.2 Tree detection and tree species classification

Other works consider full-waveform data at the scale of the tree. For instance, tree stems can be detected. [RKS07] benefit from a high point density extracted from waveform data to have enough points reaching the stems (25 pts/m^2). The authors first separate low above ground points that belong or not to the tree crown by a clustering algorithm and then perform a RANSAC-based 3D line adjustment to find the stems. Moreover, in [RSKS09], the authors segment in 3D the tree crowns using first a watershed algorithm on the Canopy Height Model. The Normalized Cut algorithm is then employed, integrating point amplitude, point echo width and estimated stem positions as feature vector for clustering. The authors finally state that a high point density (>10 pts/m^2) is beneficial for the crown segmentation but does not improve the stem detection results.

In case of surveys with high scan angles, [LRLL07] have noticed that waveforms with two strong close echoes correspond to signals that first hit the tree foliage and then the stem. They thus use the second strong echo to determine the trunk location. Eventually, features provided by the waveform decomposition may be used in combination with geometrical attributes to classify deciduous and coniferous trees [LRLL07; RKS08]. A K-means clustering approach is adopted in [RKS08], where a high overall accuracy is reached under leaf-off conditions (96%).

The tree species classification problem has also been tackled in [HHL+08; HMH+09]. [HHL+08] describe the method to segment the forest canopy and provide reliable regions for the area-based classification carried on in [HMH+09]. [HHL+08] also states that statistical values of geometric and radiometric lidar features extracted from full-waveform data

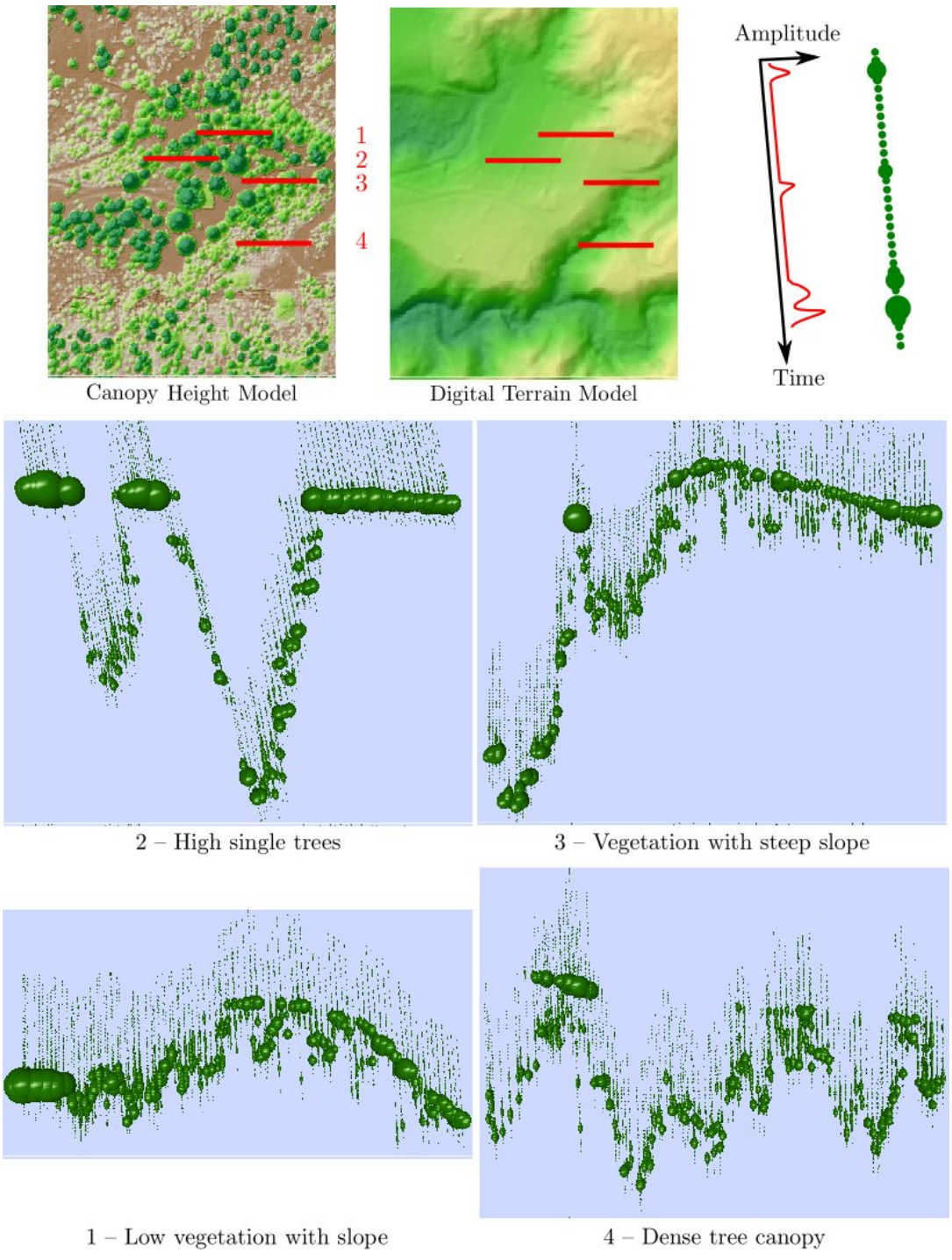
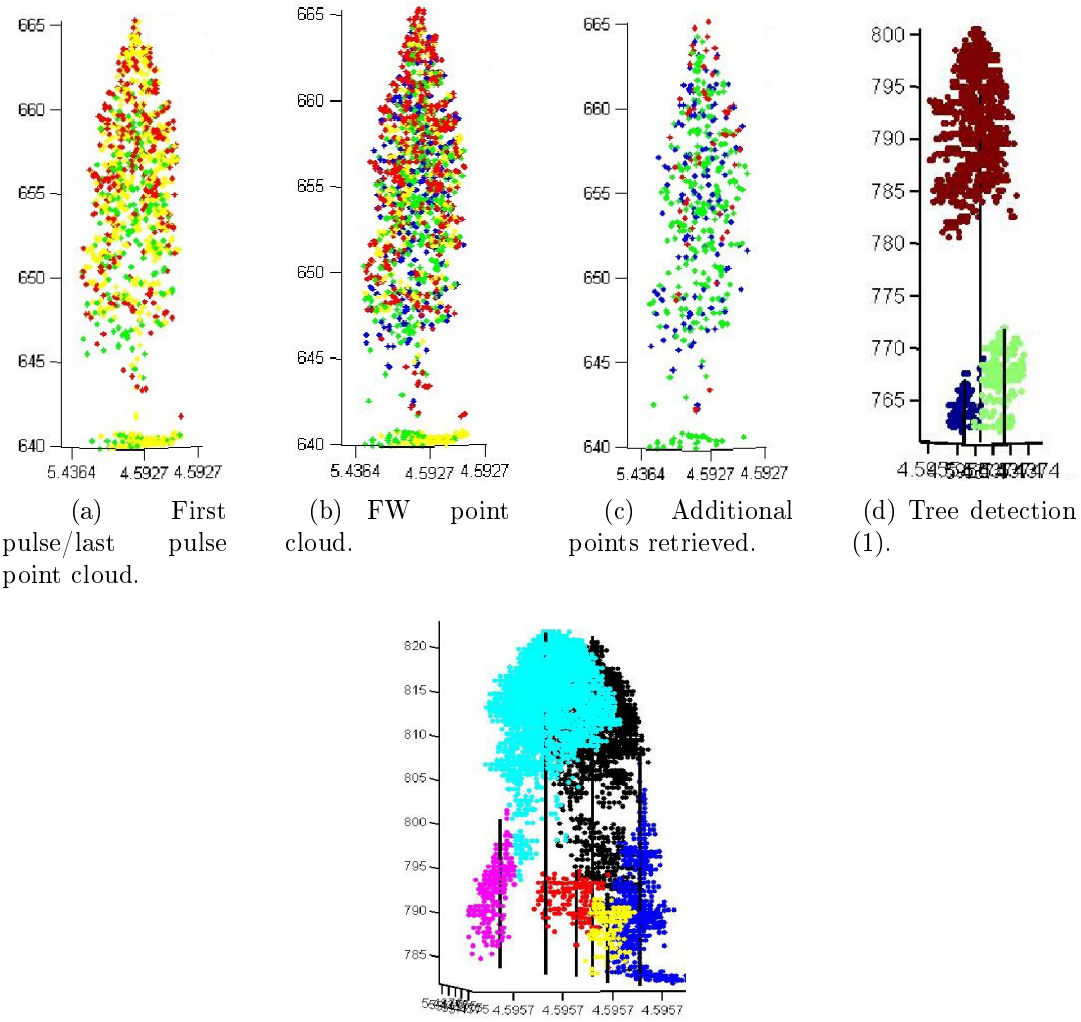


Figure 2.10: Waveforms displayed in 3D along 1D profiles. Each bin is represented by a sphere which radius is proportional to the bin amplitude. The area of interest is a forested mountaneous area in the French Alps. The data have been visualized using the software FullAnalyze [CBD+09] (courtesy of Adrien Chauve).



(a) First pulse/last pulse point cloud.

(b) FW point cloud.

(c) Additional points retrieved.

(d) Tree detection (1).

(e) Tree detection (2).

Figure 2.11: Additional points retrieved in vegetated areas, and tree detection using a Normalized Cut approach [RSKS09; Rei10]. Red: first pulses – Green: last pulses – Blue: Middle points – Yellow: Single points.

have great potential for species classification. For that purpose, the dataset has to be acquired under leaf-off condition: this allows geometric features to better reveal the physical backscattering properties of the canopy (not only the top layer). Consequently, geometric criterion (a canopy density feature, the 50^{th} height percentile d_{50}, and the echo width) and the BCS are jointly used to discriminate both at pixel and segment levels beech, spruce, and larch trees in [HMH+09]. First, a fuzzy rule set based on d_{50} and the backscattered cross section is employed to separate deciduous (beech) and coniferous trees (spruce and larch). Then, segments with a standard deviation of the echo width superior to a threshold are assigned to the larch labels. However an overall accuracy of 75% is reached, the ground truth is not significant enough to assess the transferability of the method to other areas.

2.6.5.3 Modelling forested areas

Modelling forested areas is particularly difficult due to the strong geometric complexity of the internal structure of the trees. Several studies have been carried out on this topic mainly with large-footprint lidar data. Waveforms are generated over large areas (footprint superior to 10 m) in order not to take small tree elements into account. A tree model with a high level of details (leaves) is therefore not necessary.

[BH99] give one of the first methodological developments for modelling forest scenes by simulating full-waveform lidar data. They simulate waveforms by breaking down the surface hit into small surfaces with their own backscattering characteristics but with the same reflectivity (typical of dense forest). A strong correlation between this data and that from the LVIS sensor is obtained. They show that the unmodelled effects such as multiple retrodiffusions do not make a significant contribution to the shape of the return waveform.

[SR00] propose a more comprehensive model linking full-waveform data to the spatial structure and to optical properties of the vegetation. They simulate a forest where each tree can be parameterized by its height, its species and its maximum diameter. The return waveforms are simulated by dividing the 3D scene into small cells, with specific characteristics. Full-waveform data are then simulated with different tree species and ages and then compared to terrain data and SLICER samples. The simulations show that the model may be used to find main lidar signatures. It also shows that a lidar signal provides an indication of forest populations for both horizontal and vertical structures. It has also been noticed that the age and species of the trees have a considerable effect on the waveform shape.

A different study of the relationship between the waveform and the canopy structure parameters is proposed in [NMJD01]. The arrangement of the structure on this relationship is taken into account to determine the 3D vegetation parameters. This is an adaptation of a hybrid geometric optical radiative transfer model (called GORT), to describe the effects of these parameters on the radiation environment [NMLW+97]. GORT describes the waveform as a function of the canopy parameters, which was checked against SLICER data. The lidar equation is used to determine the probability that a gap above a given height in the canopy appears. Besides, the authors propose a Directional Gap Probability Function to represent the probability that a laser beam reaches a certain point of height without being reflected. Finally, the results of the model are confirmed by comparing SLICER and LVIS data [PNMB+01].

[KMS+06] use the 3D model proposed in [SR00]. The radiative transfer model is inverted

to determine the forest biophysical parameters. Some parameters are set using in situ spectrometric measurements or approximations. Simulated data are used to establish the feasibility of the inversion and show that the parameters can, potentially, be deduced: the tree height is accurately estimated, unlike the other parameters such as the LAI and the vegetation coverage.

These methods do not take multiple scattering into account which also affect the return waveform. The method described in [Kot05] considers possible model errors, in particular for a dense canopy. To solve this, a 3D radiative transfer model is applied to simulate the propagation of photons across the vegetation. The predictions obtained on various mixed forest sites are compared to SLICER data. The simulated signal does not always match the real backscattered signal, regardless of the tree species. The main limitation of the model comes from the formulation of the probability functions that quantify the canopy structure.

Finally, the FLIGHT radiative transfer model presented in [NRSL10] may also represent multiple scattering of light (at different wavelengths) within the canopy and with the ground surface. In addition, a topographic surface can be introduced in order to model the effect of the topography on the simulated waveforms. They are compared with GLAS waveforms over a complex forest structure. Good agreement is found for several forest metrics, but some differences between both waveforms have been noticed without having been explained so far.

2.6.6 Applications in urban areas

The contribution of full-waveform data is less obvious in urban areas than in woodlands since multiple pulses only appear when the laser beam hits building edges. Most of the literature dealing with FW data in urban areas concern 3D point cloud classification (*cf.* Section 2.6.3). However, the problem of enhanced pattern recognition in raw FW data has also been tackled. [JS03] show that it is possible to visually distinguish between different urban materials hit by a laser beam and concludes that in-depth processing is required to recognise objects. They use the concept of neighbourhood between waveforms and extracted points to interpret their data. Henceforth, the interpretation of a 3D point cloud can be improved by spatial and temporal cross-correlation of the peaks in successive waveforms.

The well-known problem of detecting the building edges is covered in [JNS05]. A scene is segmented in "boundary zones" and "internal zones" by a region growing technique using the number of pulses, amplitude and distance to the target. The results are refined by modelling the amount of the laser beam power that fails in each region (inner and outer parts). This allows to locate where the beam center is and helps finding where the boundary line is. Precise results are obtained accurate to 1/10 to 1/12 of the point spacing. However, these results are based on simulated data, assuming uniform reflectance and flat "internal" areas, simplifying the modelling problem.

Eventually, by analysing the parameters of each echo and taking the relationships between adjacent waveforms into account, they distinguish vegetation and man-made structures and then extract characteristic lines for the objects found in the scene [JS05]. A spatio-temporal measurement irregular voxel space (x, y, t) is first generated to show neighbourhood relationships x for the forward motion direction, y for the swath direction, and t for

the time). Such 3D volume is analyzed in 2D in order to retrieve the sensor geometry (see Section 2.2.3). The volume is therefore sliced in vertical sections in the $(x-t)$ plane to give 2.5D information and the required relationships. The aim is to detect 3D segments within this cube, assuming a scene with a flat surface perpendicular to the incident signal (*e.g.*, , a facade). Intensity images are obtained and are processed by line detectors (in this case, Hough transform).

2.7 Data sets

Two full-waveform datasets are available for our purpose.

▶ The first dataset has been acquired over the city of **Biberach** an der Riß (Bade-Wurtemberg, Germany). The survey has been carried out by Toposys on September 2006, with the Riegl LMS-Q560 system, covering approximately 1.1 km^2 (leaf-on condition). One strip is available, with a point density close to 5 points/m^2. It results on more than 2.2 M waveforms. The strip covers different kinds of landscape: a residential area, a wooded area with significant slope, a semi-dense downtown, and an industrial area (see Figure 2.12).

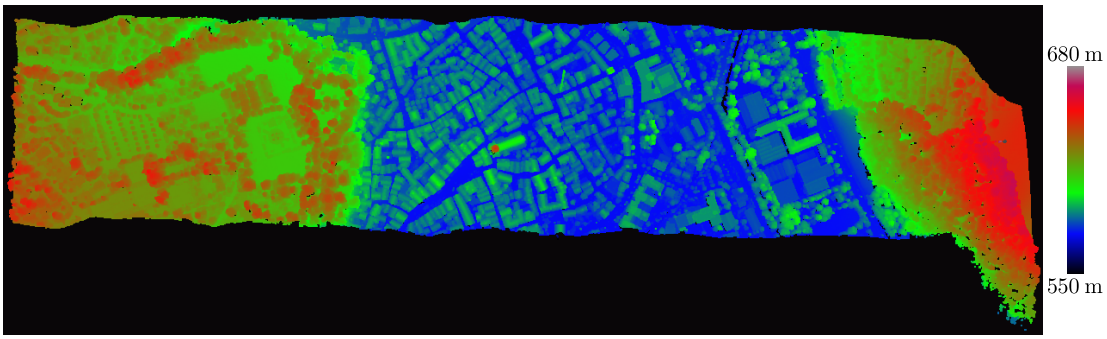
▶ The second dataset covers the city of **Amiens** (France) in February 2008 (leaf-off condition). 11 strips have been acquired with a Optech 3100 EA device (approximately 20 km^2), directly by IGN. The point density is approximately 2 points/m^2. The North and the West sides of the city correspond to fields and crops. The East side is mainly composed of woods and fields between small water channels (see Figure 2.14). Amiens downtown is not very dense but is composed of building with various kinds of shapes. The overlapping ratio between two strips is superior to 50%. However, we have decided not to benefit from such an overlap for the modelling and classification steps (see Section 1.6). All the strips are therefore considered as independent. More than 70 M waveforms are available over the city of Amiens.

Orthoimages are available for the two areas of interest. 0.25 m orthorectified aerial images cover the city of Biberach (except for the industrial area), and Amiens downtown. 0.5 m orthoimage tiles, with lower radiometric and geometric qualities compared to the latter ones, are available for the rest of the Amiens area. Since it has been decided not to use aerial image data for classification (see Chapter 5), the orthoimages are mainly used for two practical issues. First, they are mandatory data for checking the quality of the 3D point cloud georeferencing process (visual inspection). Besides, they are used to create manually 2D polygons of the classification ground truth. Such reference data are necessary for both training the supervised classifier and for assessing its accuracy.
We have carried out the whole waveform processing chain (see Figure 2.7). Firstly, from the raw full-waveform files, multiple pulse range values, T0 pulses, backscattered signals, as well as navigation metadata are extracted. Secondly, waveform processing and georeferencing is performed to turn range values into 3D point clouds expressed in the cartographic projection of the area of interest (Gauss-Krueger Germany 9^o, and Lambert 93 for Biberach and Amiens, respectively).

Area / Metadata	**Biberach**	**Amiens**
Date	September 2006	February 2008
Sensor	Riegl LMS-Q560	Optech 3100 EA
Surface (km^2)	1.1	20
Number of strips	1	11
Number of points	$2{,}342.10^3$	$70{,}135.10^3$
Flight height (m)	500	1000
Footprint size (m)	0.25	0.8
PRF (kHz)	100	100
Pulse width (ns)	4	4
Temporal sampling (ns)	1	1
Pulse density ($/m^2$)	5	2

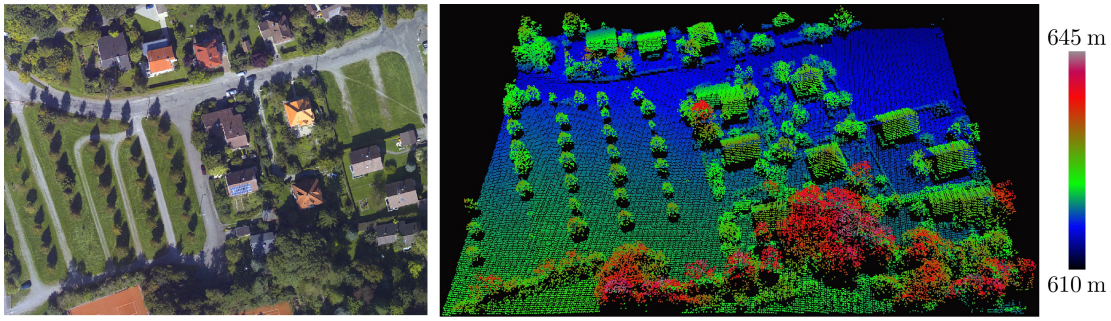
Table 2.4: Full-waveform datasets

(a) Orthoimage (0.25 m resolution).

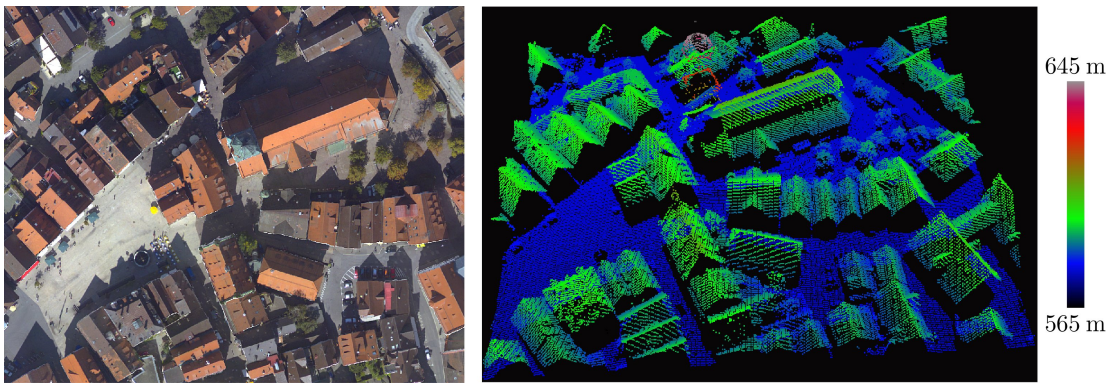


(b) 3D point cloud, coloured with respect to the altitude.

Figure 2.12: Biberach dataset: one strip across city downtown.



(a) Orthoimage and 3D point cloud of a residential area (Area 1).



(b) Orthoimage and 3D point cloud of the downtown (Area 2).

Figure 2.13: Focus on two areas of Biberach. See Figure 2.12.a for their location on the whole strip.

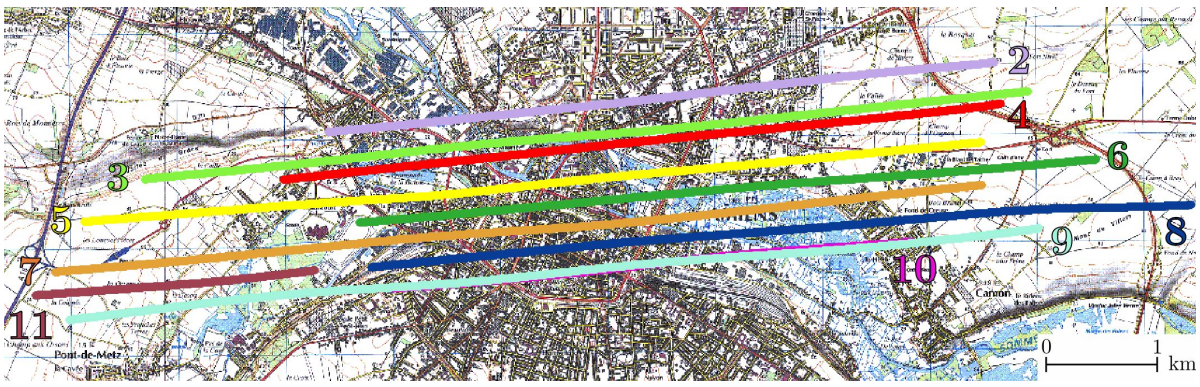
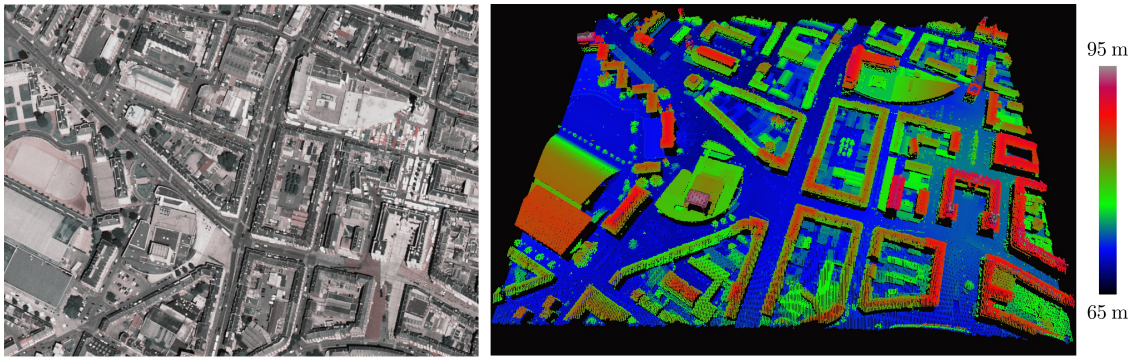
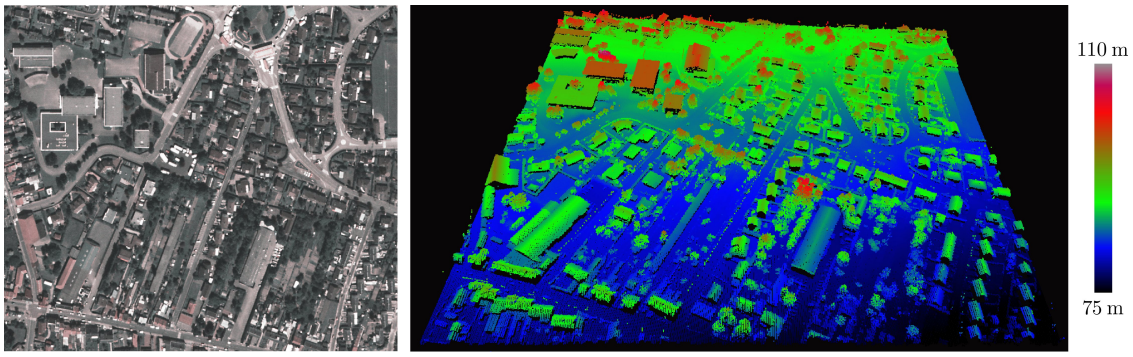


Figure 2.14: Amiens dataset. The trajectory of each strip is coloured separately. Strip 1 has not been displayed. This is the calibration strip and the lidar system has operated far more in the North. Scan25 ©IGN.



(a) Orthoimage ©IGN, and 3D point cloud of a dense urban area.



(b) Orthoimage ©IGN, and 3D point cloud of a residential area.

Figure 2.15: Focus on two areas of Amiens.

Part II

Processing lidar waveforms

Chapter 3

A Generalized Gaussian model for decomposition and modelling

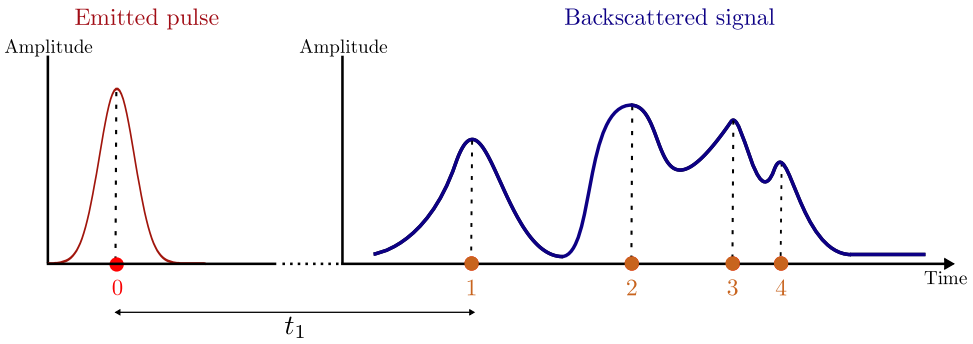
3.1 Introduction

The full-waveform laser scanning technology consists in recording a significant amount of consecutive 1D signals backscattered from the Earth surface. No additional process is performed on the data. As introduced in Chapter 2, data provided by such sensors yield more than a basic geometric representation of the Earth topography. The objective of waveform processing is to extract more information from raw lidar data than provided by multi-echo lidar systems or basic processing methods. Theoretical advantages and limitations of such data have been underlined in Section 2.2.3.
On the one hand, end-users have now access to the raw data: T0 pulses, waveforms, and georeferencing metadata. On the other hand, a pre-processing step is now mandatory to turn raw signals into 3D points clouds, so as to subsequently foster information extraction. The advantage of waveform processing is twofold (see Figure 3.1). By designing his own signal fitting algorithm, an end-user can:

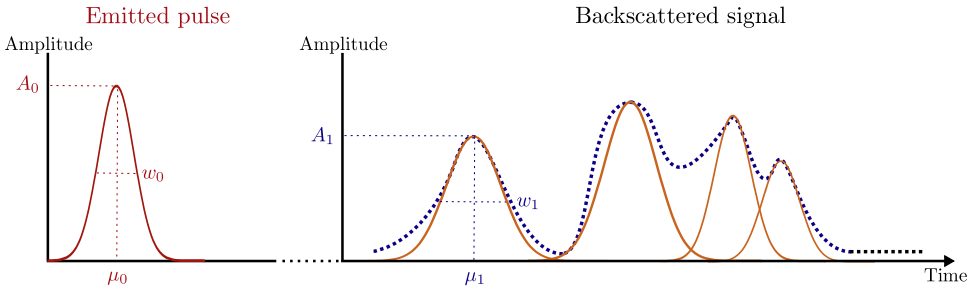
1. Maximize the detection rate of relevant peaks within the waveforms. **Waveform decomposition** allows to find the echo positions: a waveform is a series of components assuming that the position of each component within the waveform can be used to calculate the mean elevation of a specific object within the laser footprint. More points can be extracted in a more reliable and accurate way. Therefore, maximum locations should be better determined, and close objects better discriminated.

2. Model each echo with a suitable parametric function. The echo shape can be retrieved, providing relevant features for subsequent segmentation and classification purposes. This step is called **waveform modelling** or/and **waveform/echo fitting**.

These two steps can be carried out successively or jointly. The number of echoes within the waveforms can be estimated with an advanced algorithm beforehand (and then the best-fit functions can be estimated for each echo), or directly wrapped into the modelling step. A large body of the literature address the problem of lidar waveform modelling. Most of the papers that deal with FW data have proposed their own methodology



(a) Principle of waveform decomposition. The number of modes of the received signal is found and their position estimated. Then, the range between the sensor and the objects is estimated by computing the elapsed time between the emitted pulse and the retrieved echoes (here shown with object 1).



(b) Waveform modelling. For each waveform echo and for the transmitted signal, parameters of a modelling function are estimated (Gaussian in this case). Their specific shape is retrieved and the waveform reconstructed.

Figure 3.1: The two steps of waveform processing: decomposition and modelling. For each waveform, the output of such process is a series of range and model parameters values, specific to each retrieved echo.

to extract 3D points with additional features. They have in common that they use a superimposition of Gaussian functions, which is fitted to the measured waveforms with least squares, and mainly differ in the waveform decomposition step, often designed to fit to a specific problem (vegetated areas acquired with large footprint data, close target discrimination with small footprint sensors etc.) [HMB00; DPLV08; RKS08; RWMU08; LMSV10]. One can notice that such works are barely focused on urban areas [JS06; WHBD08]. Since our motivation is to analyze 3D point clouds generated from FW data in urban areas, waveform processing is mandatory and will be tackled in this chapter.

Methods based on wavelets [HD02], neural networks [BMWW04], splines [Uns99], kernel-based density estimation techniques involving, for instance, Parzen windows [BV03] or Support Vector Machines [WGS+98] are known to fit 1D signals with large flexibility and efficiency. However, they do not fulfill our requirements since the problem is not addressed by adopting a finite mixture model (FMM) [MP00]: each mode of the waveform is not segmented and modelled with a given analytical function. Such approaches are therefore not suitable. The problem of FMMs, and especially when dealing with a Gaussian Mixture Model, has been widely tackled in the image processing community, and may be relevant

to process lidar waveforms: pixel colour histograms are multimodal signals that are aimed to be decomposed and modelled, *e.g.*, for background/foreground segmentation of images and videos. Historically, estimates of the parameters of the class probability densities in mixture densities have been retrieved via the Expectation-Maximization (EM) algorithm [DLR77]. The maximum-likelihood-based method either requires knowledge of the number of components or must be coupled with model selection (as adopted, for instance, in [PSTA05]); many authors have proposed improvements and extensions to this algorithm [FJ02]. The efficiency of such algorithms relies on good initial model parameter values that is why a large body of literature also deals with this issue. For instance, EM can be coupled with simplex search, particle swarm optimization [FL07] or genetic algorithms [BBM07]. A widely adopted solution, especially when dealing with complex modelling functions, consists in adopting a stochastic version of the EM algorithm (SEM, see [CCD96]). Thanks to the stochastic sampling involved in this scheme, numerical tractability is gained along with the better exploring capabilities as compared to EM scheme and higher chances of missing the local maxima. Nevertheless, such advanced methods are not necessary in our context since for waveform decomposition it has already been shown that simple detection methods are sufficient [WUD+06] and that suitable initial parameter estimate can be achieved using lidar sensor physical knowledge (see Section 3.2.3). In case of complex waveforms with overlapping peaks, state-of-the-art image histogram thresholding methods may be helpful but still require the integration of physical knowledge.

Alternatives to EM exist such as Bayesian methods, Kalman filtering, the minimum-distance algorithm, optimization techniques (using, for instance, the gradient descent or the Levenberg-Marquardt algorithm), or the method-of-moments [Cra46; KZ04]. When dealing with parametric functions yielding more complicated analytical expressions than the standard Gaussian function, the classical statistical estimation methods fail because their moments do not exist or are too complicated to compute. For instance, in a context close to the lidar remote sensing community, new approaches have been developed for SAR data to deal with this problem combining the method of log-cumulants and the Mellin transform [TNTM04; KMSZ08].

Since seminal papers on airborne waveform processing [HMB00; WUD+06] have proved that Gaussian decomposition using non-linear least squares optimization techniques give good results, firstly for generating 3D points and secondly for extracting additional parameters for subsequent classification, we have followed the 2-step workflow adopted in the literature, with some improvements for urban areas. There is no specific need to design a dedicated algorithm for Gaussian-based waveform modelling. In case of non symmetric echoes, this is however relevant and will be tackled in Chapter 4.

Waveform decomposition is first performed with a standard first derivative analysis of a smoothed version of the waveforms. Besides, assuming echo shape to be symmetric but with slight distortion, we model them with the Generalized Gaussian function, which is an extension of the standard Gaussian model. This leads to a nonlinear optimization problem, which is performed adopting a non-linear least squares method, the Levenberg-Marquardt algorithm. Finally, the fitting step is repeated until no additional echoes are found in the residuals between the raw waveform and the proposed adjustment. The results of the decomposition are compared qualitatively and quantitatively both with the real-time

sensor detection methods and the standard Gaussian decomposition approach. Finally, extracted Generalized Gaussian parameters are analyzed over areas and classes of interest in a urban context.
However focused on urban areas for small-footprint data, the developed approach remains valid over vegetated areas [CVB+09] and could be adapted to large-footprint waveforms with minor modifications.

The proposed approach is developed in the following Section, whereas Section 3.3 presents the decomposition and modelling results for our two datasets. The fitting process is also qualified. Then, the full-waveform features, namely the parameters of the Generalized Gaussian modelling function, as well as derived attributes are analysed in the perspective of land-cover classification in urban areas. Eventually, the main conclusions and interesting outlooks are discussed in Section 3.5.
Figure 3.2 gives an overview of this 2-step approach.

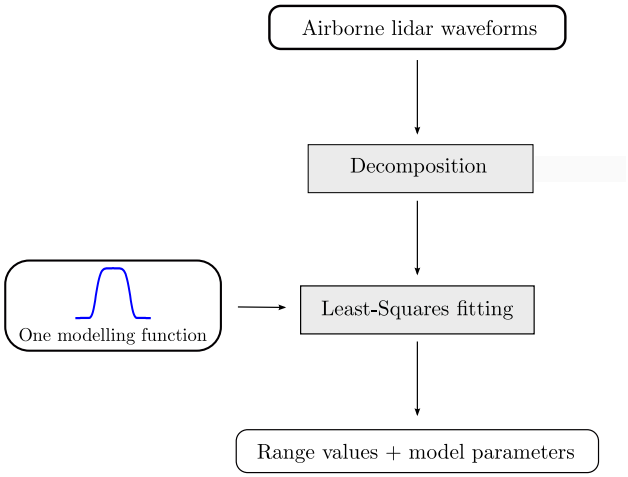


Figure 3.2: Overview of the waveform decomposition and fitting process.

3.2 Methodology

3.2.1 Presentation

The first goal is to extract as many peaks from the signal as possible, but also shape information for each echo. The approach is parametric *i.e.*, we estimate parameters of a mathematical model for each echo. We assume that the returning laser pulse is composed of a series of potentially-overlapping reflections similar in shape to the impulse response and that can be fitted with the same analytical function. To improve signal fitting, a new waveform model has been proposed and is described in Section 3.2.2. As mentioned above, waveform processing consists in two main steps: first, the number of components and initial values are estimated (Section 3.2.3). Then the parameters are optimized. The optimization process is well-known and it has been demonstrated that either the Expectation-Maximization algorithm (Maximum Likelihood estimates) [PSTA05] or the Levenberg-Marquardt algorithm (non-linear least-squares method) give good results (Section 3.2.4). Nevertheless, optimization relies strongly on initial parameters. They

therefore must be estimated very carefully to avoid erroneous results. As waveforms may be complex, the 2-step process is carried out iteratively. Once the model parameters have first been optimized for each echo, the waveform is reconstructed and compared to the raw waveform. If significant residuals are found, we assume this corresponds to undetected echoes, and the optimization process restart for all the echoes. New comparisons and new peak additions are performed until the residuals are lower than a given threshold ϵ. The full workflow is described in Figure 3.3.

A waveform is composed of a sequence of N points uniformly-spaced $\{(x_j, y_j)\}_{j=1,..,N}$ sampled traditionally at 1 GHz. We aim to decompose each sequence into a sum of components representing the targets located within the travel path of the laser beam as

$$y = \sum_{i=1}^{n} f_i(x) + b \tag{3.2.1}$$

where b is the noise of the waveform, n is the number of components, and f_i the echo model. Here, we have the same model for all the echoes.

3.2.2 Model selection

Based on Equation 2.1.8, the received power as a function of time can be expressed as follows:

$$P_r(t) = \sum_{i=1}^{n} k_i \; S(t) * \sigma_i(t) \tag{3.2.2}$$

where n is the number of echoes, k_i is a target-dependent value varying with range between sensor and target, $S(t) = P_t(t) * \eta_{\text{sys}}(t)$ is the system waveform of the laser scanner, and $\sigma_i(t)$ the apparent cross-section of the i^{th} target. [WUD+06] describe that S has a Gaussian behaviour for RIEGL LMS-Q560 system and we consider that this hypothesis is still valid for our Optech 3100 device. Several other models exist and are tailored to specific sensors (uniform for [Ste00] or with exponential decrease for [JS06]). Furthermore, by assuming that the scattering properties of the n targets can also be described by Gaussian functions, [WUD+06] have derived the following formulation of the received power, that, due to convolution properties, allows to represent waveforms as a series of Gaussian functions.

$$P_r(t) = \sqrt{2\pi} A_0 w_0 \sum_{i=1}^{n} k_i \frac{A_i w_i}{\sqrt{w_0^2 + w_i^2}} \; \exp - \frac{(t - t_i)^2}{2(w_0^2 + w_i^2)} \tag{3.2.3}$$

where $A_{0/i}$ and $w_{0/i}$ are the amplitude and the width of the emitted pulse and the target i respectively.

The Gaussian function has been widely adopted in the literature, and has been shown to be an efficient model for small-footprint waveforms, however incorrect in theory. It is known that the transmitted pulse is slightly dissymmetric and thus not really Gaussian. Moreover, we have no a priori knowledge on the true parametric form of the target cross-sections. The simulation of waveforms over different kinds of shape, vegetation, or urban items has been barely investigated and no analysis by synthesis reasoning has been derived

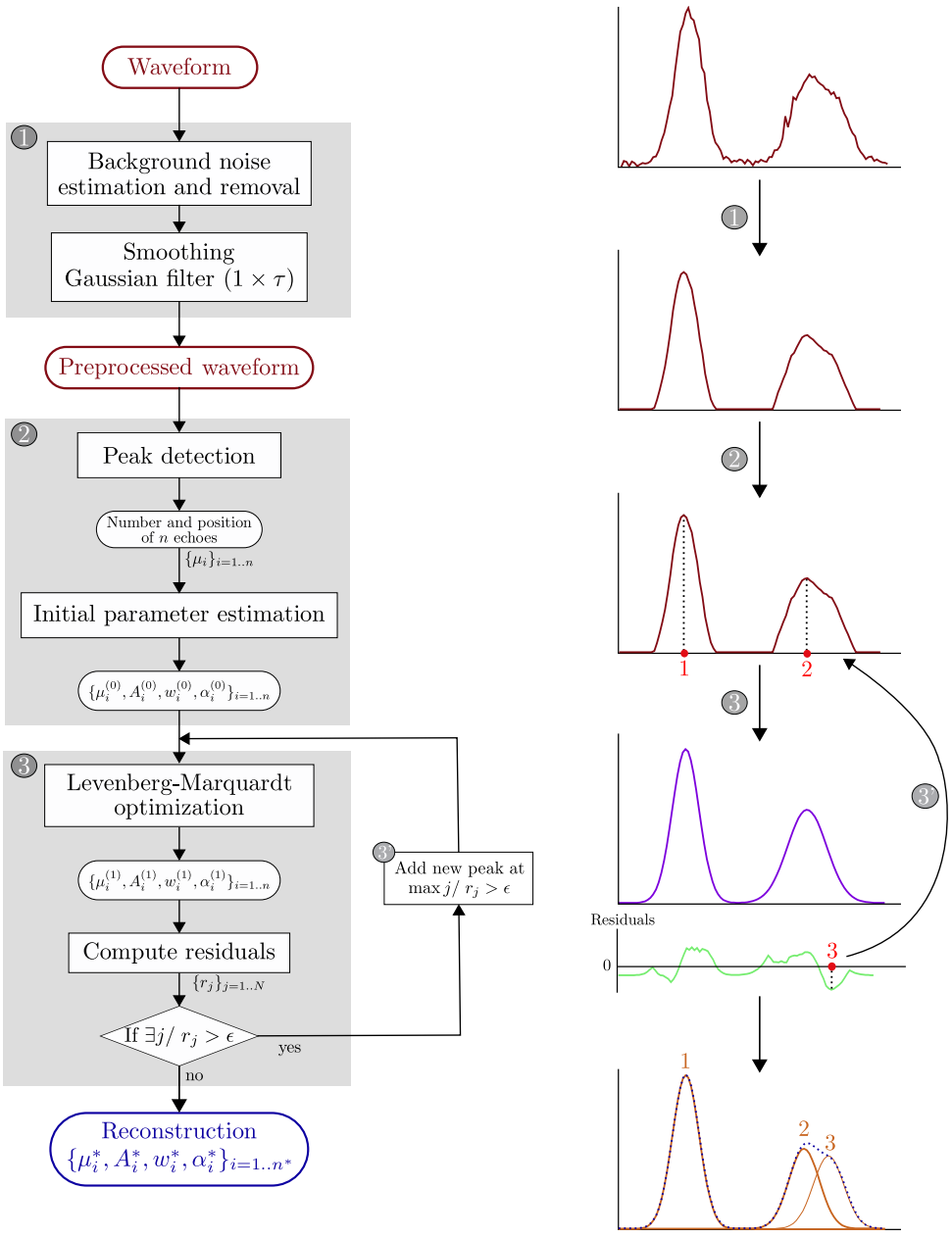


Figure 3.3: Flowchart of Generalized Gaussian waveform processing. The iterative procedure allows to retrieve three echoes instead of two.

yet. Thus, the Gaussian model has appeared to be a simple and efficient solution: 1/ many authors have exhibited very good fitting quality results based on this assumption; 2/ amplitude and width parameters are discriminant features for classification (*cf.* Section 2.6.3) It would be therefore logical to adopt such solution. However, a close look at the data shows that many echoes are indeed symmetric with the well-known bell curve, but are slightly distorted on the top. It is therefore of interest to extend waveform processing capabilities by using a more complex parametric model that is able to deal with such behaviour.

The choice of the new modelling function relies on several criteria. First, the function has to be unimodal, and its parameters have to be related directly to the shape of the target. Second, it should be capable of modelling a larger variety of statistical behaviours (flexibility). Third, for numerical stability and tractability, it should not require the estimation of a large number of parameters and the derivative of the function should have an explicit formulation, in order to use classical optimization algorithms. Among the models available in the literature, the **Generalized Gaussian** (**GG**) distribution is a particularly attractive candidate as it requires only one additional parameter to be estimated compared to the Gaussian model, and it can approximate a large class of statistical distributions thanks to its shape parameter α (*e.g.*, impulsive, Laplacian, Gaussian, and uniform distributions).

$$f_{\mathrm{GG}}(x) = A \exp\left(-\frac{(x-\mu)^{\alpha^2}}{2w^2}\right) \tag{3.2.4}$$

The parameter set is $\boldsymbol{\Theta} = \{\mu, A, w, \alpha\}$, where:

- $\boldsymbol{A}$ and $\boldsymbol{w}$ give the amplitude and the width of the Gaussian model.
- Shift parameter $\boldsymbol{\mu}$ still indicates the position of the mode of the function.

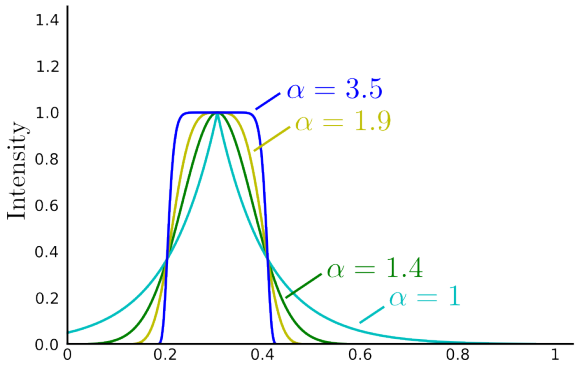


Figure 3.4: Behaviour of the Generalized Gaussian function.

The behaviour of the GG model is displayed in Table 4.1. The main limitation of its introduction is that with a simple non-linear least-squares algorithm, it will increase the number of fits that do not converge. This is due to the increasing number of degrees of freedom of the function and also to the more complex expression of the gradient [AAB99]. The Generalized Gaussian is a popular solution to model SAR amplitude [MZS06], image texture [dWSD99; DV02] or even outliers in image matching [HSSV03]. One can notice that we have not adopted the traditional formulation, involving the Beta function. On the one hand, the Beta function allows higher flexibility as our simple GG formulation. On the other hand, when processing millions of waveforms, its evaluation and the evaluation of its derivation become computationally very expensive in time. This is why it has been removed.

3.2.3 Peak detection and initial parameters estimation

The two first steps of the algorithm are a coarse peak detection and an estimation of the model parameters in order to initialize the Non-Linear Least-Squares (NLS) fitting (Step 2 in Figure 3.3). Peak detection can be carried out by traditional pulse detection methods: leading-edge, constant fraction or center-of-gravity methods are valid and often used in the literature (see [WUM+04] and Chapter 7 of [ST09] for further explanations).

The basic detection method is based on the zero crossings of the first derivative on the thresholded version of the waveform. We estimate the number of components n from the number of inflection points of the signal and obtain initial estimates of the locations from the positions of its consecutive inflection points.
As such a method is affected by noise present in the waveform, the background noise is first thresholded. We first set the background noise to a rough value b equal to the digital current offset value (*i.e.*, an offset value of a signal from zero is included when the signal is converted from analogue to digital – *e.g.*, 5 for the Riegl LMS-Q560 datasets, and 9 for the Optech 3100 device). The real noise value is computed with the mean amplitude value of all the bins which amplitude is inferior or egal to this first value.
Then, smoothing is performed by convolving a Gaussian of width τ with the observed waveform.
The detection algorithm takes into account a minimal number of samples separating two detected peaks. This bound is set to the sensor range resolution which is $c\tau/2$, where τ is the half width of the impulse response. Our ability to resolve these peaks is limited by the vertical resolution of the device and determined by factors such as the laser pulse width, the emitted pulse, and the digitizer sampling rate.
Difficulties arise when two neighbouring peaks are close enough together that only two inflection points (instead of four) are detected making it impossible to isolate the pair. One solution is to perform a second pulse detection on the residuals between the observed waveform and the previous fit. If for a bin, the residual is superior to a given threshold ϵ, a new peak is located here and a new fit is run with the new component. The iterative procedure is similar, for instance, to [HMB00; LMSV10]. [HMB00] have adopted a relevant ranking strategy. A high number of peaks is detected. Echoes with significant amplitude and width are used for the initialization step. Then, peaks are successively added following a given order based on the proximity of the peak to previously accepted peaks. Echoes are introduced only if the quality parameter is not satisfactory. [LMSV10] focus their efforts on asymmetric peaks arguing such behaviours result from the superposition of the contribution of two targets (see Section 2.4.2 for more details).

Once peak locations $\{\mu_i\}_{i=1..n}$ have been estimated,

- their amplitude are set to the mean value of the two adjacent bin values:

$$\forall i, A_i = \frac{y_{\lceil \mu_i \rceil} - y_{\lfloor \mu_i \rfloor}}{2}$$

- w and α are initialized with constant values for all the waveforms, respectively $w = \tau$ ns and $\alpha = \sqrt{2}$.

Due to the low SNR, rough noise reduction is performed and prevent the detection of weak pulses (*i.e.*, with amplitude slightly over the noise level). One solution should be a cross-waveform pulse detection, based, for instance, on stacking (Section 2.4.2), performing a round trip between the waveform of interest, the previous and following waveforms (temporally speaking), and an accumulation of spatially close waveforms.

3.2.4 Fitting the waveforms

We aim to find the optimal values for n GG functions, knowing N bins of a waveform. The problem is thus described by a system of N observations with $4n$ unknown parameters. Such an optimization problem can be efficiency solved using non linear least-squares methods (since the modelling function f is not linear).
As presented in Section 2.5.2, a variety of such techniques exist for the GMM problem (simple linear superposition of Gaussian components), which find the solution by iteratively trying a series of combinations of the parameters until a solution is found: [JS05] use the Gauss-Newton, many authors such as [HMB00] prefer the Levenberg-Marquardt technique, whereas [PSTA05] estimate the Gaussian parameters with the EM algorithm. One can notice that in case of Gaussian mixtures, the least squares and maximum likelihood problems are exactly the same [Bis06]. In the context of Generalized Gaussian distributions, this is still true and such techniques are still valid.

We choose the Levenberg-Marquardt algorithm to compute the fit. The Levenberg-Marquardt method is a simple variation on Gauss-Newton iteration, designed to provide faster convergence and regularization in the case of overparametrized problem [Lev44; Mar63]. It requires that the first derivatives of the minimizing function for each parameter to be known. In a linear least-squares problem, knowing $\{x_i, y_i\}_{i=1,..,N}$, we aim to optimize the set of parameters $\mathbf{\Theta}$ so that the sum of the squares of the deviations

$$E_{\mathrm{LS}}(\mathbf{\Theta}) = \sum_{i=1}^{N} \|r_i\|^2 = \sum_{i=1}^{N} \|y_i - f(x_i|\mathbf{\Theta})\|^2 \tag{3.2.5}$$

becomes minimal. r_i is the residual between the observed waveform and the predicted one (adjustment). Thus, we look after:

$$\mathbf{\Theta}^* = \underset{\mathbf{\Theta}}{\operatorname{argmin}} E_{\mathrm{LS}}(\mathbf{\Theta}) \tag{3.2.6}$$

For functions f that are not linear in the unknown parameters $\mathbf{\Theta}$, the problem is solved by iteratively re-linearizing Equation 3.2.5 around the current estimate of $\mathbf{\Theta}$ using the gradient derivative $J = \frac{\partial f}{\partial \mathbf{\Theta}}$; and computing an incremental improvement $\Delta\mathbf{\Theta}$. $J \in \mathbb{R}^{4n}$ is the Jacobian matrix, and f is thus approximated using the Taylor expansion: $f(\mathbf{\Theta}+\Delta\mathbf{\Theta}) = f(\mathbf{\Theta}) + J\Delta\mathbf{\Theta} + O(\|\Delta\mathbf{\Theta}\|^2)$. To minimize the problem, we iteratively find an update $\Delta\mathbf{\Theta}$ to the current parameter estimate $\mathbf{\Theta}$ by minimizing:

$$E_{\mathrm{NLS}}(\Theta) = \sum_{i=1}^{N} \|y_i - f(x_i|\mathbf{\Theta})\|^2 \tag{3.2.7}$$

$$\approx \sum_{i=1}^{N} \|J(x_i|\mathbf{\Theta})\Delta\mathbf{\Theta} - r_i\|^2 \tag{3.2.8}$$

One can notice that two issues appear: 1/ this approximation only holds near a local minimum; 2/ for small values of $\Delta\mathbf{\Theta}$, the update $\mathbf{\Theta} \leftarrow \mathbf{\Theta} + \Delta\mathbf{\Theta}$ may not always decrease the overall residual error. In order to mitigate this problem, a solution is to take a smaller step $\mathbf{\Theta} \leftarrow \mathbf{\Theta} + \alpha\Delta\mathbf{\Theta}$ with $\alpha \in [0, 1]$. α can be found by linear search applied to the function $\phi(\alpha) = f(\mathbf{\Theta} + \alpha\Delta\mathbf{\Theta})$ [MNT04; NW06]. Another solution to ensure a downhill step in error is to add a diagonal damping term to be approximate Hessian matrix $H = \sum_{i=1} J^{\mathrm{T}}(x_i)J(x_i)$. We therefore have to solve:

$$[H + \lambda \operatorname{diag}(H)]\Delta\mathbf{\Theta} = \sum_{i=1}^{N} J^{\mathrm{T}}(x_i)\, r_i \tag{3.2.9}$$

The damping parameter λ is increased (multiplied by 10) if the squared residual is not decreasing as fast as predicted in Equation 3.2.8, and is decreased (divided by 10) if the expected decrease is obtained. λ is initially set to typically 10^{-3}. The algorithm moves between Newton iteration when λ is very small (rapid convergence in the neighbourhood of the solution), and a descent approach when *lambda* is large (substantially larger than 1), which still enables a decrease of the squared residuals. Since we minimize with respect to a small number of parameters, there is no need to adopt enhanced versions of the algorithm, such as the sparse Levenberg-Marquardt algorithm [HZ00].

Problems occur when the fit ends up in a local minimum, which may not be the best solution possible: this is mainly due to initial estimates significantly far from the true solution. However, waveforms have "known" shapes. Thus, the initialization step, constrained by our lidar physical knowledge, is very rarely wrong. Nevertheless, Levenberg-Marquardt can fail to converge. Contrary to the Newton algorithm, the Levenberg-Marquardt algorithm requires the inversion of the Hessian H. It can lead to inconsistent numerical values for $\mathbf{\Theta}$ parameters: such cases happen very rarely since in practice we have to find a maximum of 28 parameters, considering modelling waveforms composed of a maximum of seven echoes with the GG function.

As indicated in the previous section, if the fit does not approximate the return waveform to a prescribed accuracy, the full optimization process is reiterated. If for one bin of the waveform, the residual between the estimate and the raw waveform is superior to a threshold ϵ, a new peak is found at that location and $4(n + 1)$ parameters are optimized. In practice, ϵ has to set to $\frac{3}{2}b$.

3.2.5 Georeferencing process

The output of the waveform processing step is a 3D point cloud with p additional features. In our case, $p = 3$, and the features are the parameters of the Generalized Gaussian model A, w, and α.

As described in Figure 3.1.a, the peak location μ_i gives the elapsed time to the beginning of the laser pulse transmission. This value must be subtracted by μ_0 which is the location of the transmitted pulse. Since this pulse slightly fluctuates in shape and its amplitude

and width are required for subsequent feature normalization (see Section 3.4), a Generalized Gaussian fit is also performed on it. The main difference with waveform fitting is that we know that only one easy-to-find peak is present in the T0 pulse. The real range value between the sensor and the target is $\forall i \in [1, n], d_i = c\,(\mu_i - \mu_0)/2$. Each target coordinate is expressed in the scanner coordinate frame, in polar coordinates (d_i, γ_i). The so-called georeferencing process is now performed to turn these coordinates to a (x, y, z) triplet within a given geographic datum. The GPS position of the aircraft, and the sensor attitude values for each laser shot are recovered from the full-waveform data file for each laser shot to compute the {x, y, z} in the geocentric datum. The final step consists in transforming these positions in some cartographic projection (planar coordinates). The transformation formulas cannot be expressed in details because they differ from a sensor to another. Offset values are different, depending on the configuration of the laser system, GPS and Inertial Measurement Unit (IMU) devices. However, some basic transformation formulas are provided below.

Direct georeferencing is illustrated in Figure 3.5. The following steps are necessary to transform the polar coordinates expressed in the laser local frame:

1. Compute the cartesian position of the point in the sensor frame from the polar coordinates;
2. Compute the position in the IMU/GPS frame using sensor and GPS calibration data;
3. Compute the position in the geocentric datum using plane position (provided by the GPS device) and attitude (roll, pitch, heading – provided by the IMU).

The global transform can be expressed as follows:

$$c_{\mathrm{P}}^{\mathrm{gc}} = c_{\mathrm{gps}}^{\mathrm{gc}} + \mathcal{R}_{\mathrm{gps}\to\mathrm{gc}} \left(\mathcal{T}_{\mathrm{las}\to\mathrm{gps}} + \mathcal{R}_{\mathrm{las}\to\mathrm{gps}}\, c_{\mathrm{P}}^{\mathrm{las}} \right) \tag{3.2.10}$$

where:

- $c_{\mathrm{P}}^{\mathrm{gc}}$ are the coordinate of the point P in the geocentric datum;
- $c_{\mathrm{gps}}^{\mathrm{gc}}$ are the coordinate of the GPS in the geocentric datum;
- $\mathcal{R}_{\mathrm{gps}\to\mathrm{gc}}$ is the rotation from the GPS frame to the geocentric datum;
- $\mathcal{T}_{\mathrm{las}\to\mathrm{gps}}$ is the translation of the laser frame to the GPS frame;
- $c_{\mathrm{P}}^{\mathrm{las}}$ are the coordinates of the laser point in the laser frame;
- $\mathcal{R}_{\mathrm{las}\to\mathrm{gps}}$ is the rotation from the laser frame to the GPS frame.

$\mathcal{T}_{\mathrm{las}\to\mathrm{gps}}$ and $\mathcal{R}_{\mathrm{las}\to\mathrm{gps}}$ are constants that are calibrated when the lidar device is mounted. Rotation $\mathcal{R}_{\mathrm{gps}\to\mathrm{gc}}$ is given by the plane attitude and position $c_{\mathrm{gps}}^{\mathrm{gc}}$ is provided by the GPS. Since laser PRF is superior to 100 kHz and GPS and IMU frequencies at around 200 Hz and 2 Hz respectively, the plane position and attitude are simply linearly interpolated between two values.

The geometric accuracy of the estimate of the peak position has an impact on the final planimetric and altimetric qualities of the georeferenced 3D point. If the effect of beam divergence caused by diffraction is neglected and considering that the off nadir angle of our data is usually small, therefore the laser profiles diverge only little from nadir view and the peak location misestimation effects mostly the vertical direction (see Figure 3.6). The altimetric and planimetric shifts are respectively equal to $\varepsilon \cos\theta$ and $\varepsilon \sin\theta$, where ε is

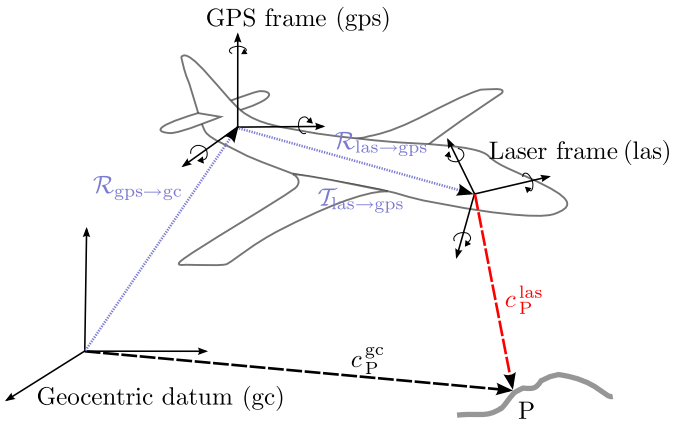


Figure 3.5: Georeferencing process and notations.

the line-of-sight distance between the real maximum position and the estimated one, and θ the emission angle of the laser impulse. If we assume an off-nadir angle of 8 degrees, a line-of-sight distance d=7.5 cm (half of range sampling grid of 15 cm or 1 ns) is projected to a line of length of 7.4 cm on the vertical axis and 1.2 cm on the horizontal axis.

Such values are mitigated by the transform of the point coordinates in the laser frame to geographic coordinates. This impact is difficult to evaluate theoretically since the georeferencing process involves metadata that both depend on the sensor (*e.g.*, attitude offsets) and the current survey (flight height and plane attitude).

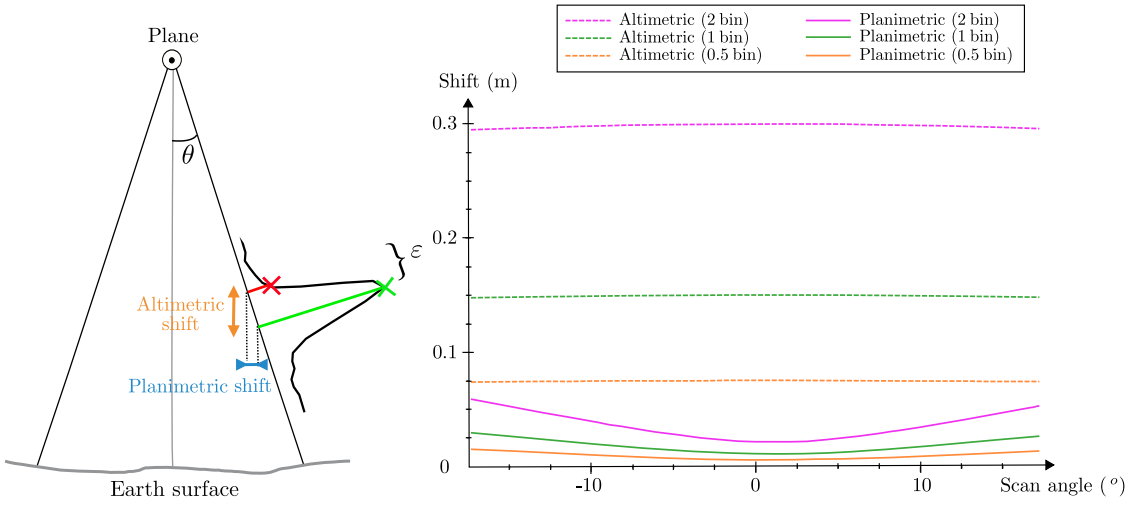
3.3 Results and evaluation

Amiens and Biberach datasets have been entirely processed. The number of points that are retrieved as well as the associated fitting quality are provided in the two following sections. The proposed approach can be quantitatively evaluated comparing these results firstly with the multiple-pulse point cloud, and secondly with the standard Gaussian decomposition.

3.3.1 Decomposition results

Waveform decomposition (or peak detection) is qualified by the number of echoes that are retrieved. In such a case, "ground truth" is almost impossible to obtain since a manual labelling of a sufficiently significant number of waveform is too tedious to be carried out. Consequently, we cannot provide detection results in terms of "*true positive*", "*false positive*", and "*false negative*". A close look at the decomposition results, waveform per waveform has been carried out in order to assess the reliability of the proposed approach, and to set up the underlying parameters (width of the Gaussian filter τ, and residual threshold ϵ). However, it only provides very good hints about the general behaviour of our fitting algorithm on different kinds of shapes, without comprehensive evaluation.

One solution for evaluation is to compare the decomposition results with the number of points that are delivered to an end-user (*i.e.*, who has no access to the lidar waveforms).



(a) Illustration of the planimetric and altimetric shifts caused by a peak misestimation of ε bins.

(b) Plots of the planimetric and altimetric shifts, in the geocentric datum, between a 3D point with correct peak estimation and a 3D point with shifted peak estimation. Three simulations are carried out with the Optech sensor and with three different shifts: 0.5, 1, and 2 bins (1 bin=1 ns). Scan angle also varies.

Figure 3.6: Influence of the peak misestimation on the planimetric and altimetric accuracies of the resulting 3D point.

Tables 3.1 and 3.2 provide the full decomposition results for Amiens and Biberach areas, respectively. MP stands for "Multiple Pulse", in contrast to FW ("Full-waveform"). It refers to the end-user point cloud. Such a point cloud is not generated identically for the two datasets. For Amiens, a maximum number of four range values (first-second-third-last) is provided on-the-flight by the Optech sensor, using the constant fraction method (Table 3.1). For Biberach, the MP point cloud has been generated by Riegl's software RiAnalyze. First, a simple threshold-based peak detection is performed. Afterwards, a Gaussian decomposition is carried out (Table 3.2).

From Table 3.1, one can see that full-waveform processing allows to retrieve more points that multiple pulse sensors. This is a already widely known assessment. Approximately, 4.1% additional points are found over Amiens. This mainly correspond to building edges and vegetated areas, mainly trees (see Figure 3.11). Considering a survey of more than 20 km^2, this leads to around 2.7M additional points. Compared to additional percentages that are traditionally reported in the literature, especially in forested areas, such a figure can be considered as relatively low. The reasons are twofold. Firstly, Amiens in a urban area, surrounded by crops and fields, acquired under leaf-off condition (February). The probability of multiple reflections (and subsequent complex waveforms) are therefore much lower than in a wooded area under leaf-on conditions. Secondly, the laser beam divergence was 0.8 mrad, leading to a footprint size of 0.8 m. Such a specification favors waveforms with multiple peaks (Figure 2.1), but does not make the waveforms more complex. Since the laser energy is spread in a larger patch, the targets should exhibit strong backscattering properties to generate a signal with sufficient amplitude, especially under low SNR con-

ditions. Table 3.1 gives the results for Amiens, strip by strip. Strip 8 exhibits the higher additional percentage of all the strips (6.2%). Such percentage locally reaches more than 10% in the Southern part of Amiens.

Strip	# waveforms	# FW points	# MP points	# additional	% additional
1	3,713,105	3,906,624	3,762,797	143,827	3.8
2	10,552,552	11,147,831	10,733,378	414,453	3.86
3	7,0603,46	7,394,908	7,162,135	232,773	3.25
4	4,685,378	4,892,859	4,725,393	167,466	3.54
5	5,958,867	6,328,962	6,056,634	272,328	4.49
6	5,426,282	5,672,696	5,454,878	217,818	3.99
7	6,502,702	6,977,330	6,627,451	349,879	5.3
8	2,415,712	2,574,170	2,423,806	150,364	6.2
9	6,030,496	6,444,359	6,146,551	297,808	4.8
10	6,204,585	6,471,038	6,261,637	209,401	3.3
11	5,254,211	5,532,875	5,318,653	214,222	4.03
Total	63,804,236	67,343,652	64,673,313	2,670,339	4.1

Table 3.1: Decomposition statistics for the Amiens data set. **FW**: 3D points extracted with the proposed approach. **MP**: 3D points extracted using the range values provided directly by the Optech sensor.

For Biberach area, the data have been acquired with a very small footprint sensor, and under leaf-on conditions. Moreover, the area has a higher proportion of vegetated areas, leading to more complex waveforms. The ratios of the number of detected peaks over the number of waveforms are approximately 1.05 and 1.43 for Amiens and Biberach, respectively. Table 3.2 shows that the additional percentage of retrieved points is much higher, and reaches 24%. This is also a consequence of the peak detection algorithm that is employed. The echoes have been estimated with a simple threshold-based method for Biberach, which leads to coarser results, compared to the constant fraction method.

Strip	#waveforms	#FW points	#MP points	#additional	% additional
Biberach	2,027,547	2,903,976	2,341,916	562,060	24

Table 3.2: Decomposition statistics for the Biberach data set. **FW**: 3D points extracted with the proposed approach. **MP**: 3D points extracted using the range values provided by Riegl's sotfware RiAnalyze (simple waveform thresholding and subsequent Gaussian decomposition).

The specific interest of our proposed algorithm can also be assessed in terms of decomposition performance. First, the choice of the Generalized Gaussian model instead of the Gaussian function only slightly changes the number of detected echoes. For Amiens, around 10 points per strip are additionally found with the GG model, which is negligible. These points are located inside the tree canopy, and this is difficult to ensure whether they correspond to genuine objects. This may due to the fact the GG function allows to deal with more rectangular echoes, leading to potential higher residuals in the ascending and descending parts of the peaks, compared to a smoother Gaussian curve.

Furthermore, the impact of the iterative procedure can be seen on the two first columns of

Tables 3.3 and 3.5. For Amiens dataset, 0.68% of the waveforms exhibited residuals above the threshold and were processed with one additional echo. This second step has lead to approximately 267,000 more points, which correspond to 0.69% more points (among the total of 4.1%, presented in Table 3.1). For Biberach, 84,000 additional points are found (3.6% over 24%). Figure 3.7 shows on a small area that such waveforms mainly correspond to waveforms that have reached vegetated areas.

Finally, there are some MP peaks that have no FW homologue (less than 100 for the whole Amiens dataset). This is due to two main reasons. Firstly, some of them exhibit low amplitude, very close to the noise level, which is not possible with our algorithm. Secondly, the Levenberg-Marquardt optimization procedure may fail, and leads to inconsistent parameters values. Such echoes are discarded in the georeferencing process, and do not lead to a 3D point. Moreover, the two-step peak detection method enhances the stability of the optimization method. Indeed, when providing more relevant estimates of echo positions as input data, the fitting procedure finds more easily, and less inconsistent results are finally obtained.

Strip	**% 2 iterations**	**% additional**	ξ_1	ξ_2	ξ_{gauss}
1	0.63	0.62	2.33	2.09	2.93
2	0.65	0.64	2.15	2.01	2.51
3	0.6	0.61	1.69	1.55	2.41
4	0.61	0.61	2.20	1.97	2.78
5	0.62	0.61	2.12	1.98	2.56
6	0.61	0.62	1.96	1.67	2.44
7	0.65	0.64	1.87	1.70	2.71
8	0.72	0.73	2.03	1.88	2.19
9	0.68	0.69	1.88	1.64	2.25
10	0.79	0.78	1.97	1.74	2.42
11	0.64	0.65	1.44	1.17	1.79
Total	0.68	0.69	1.99	1.55	2.43

Table 3.3: Comparison of the waveform fitting quality between the proposed Generalized Gaussian decomposition (ξ_1 and ξ_2, for one and two iterations, respectively), and the standard Gaussian approach (ξ_{gauss}), for Amiens data set.

3.3.2 Quality analysis

Two main quality measures are considered of interest for our purpose:

- the range accuracy: does waveform processing allow to better locate the peaks?
- the fitting quality: does an iterative GG decomposition improve signal fitting?

The range accuracy *i.e.*, the ability of our approach to correctly detect the echo maxima within the waveform cannot be evaluated with a self diagnostic. The required detection precision cannot be lower than one bin which is the sampling rate. Some simulations have been carried out. They have shown that the absolute differences between the real peak

Strip	$\overline{\epsilon}$ (m)	$\overline{\sigma_\epsilon}$
1	0.058	0.38
2	0.061	0.385
3	0.056	0.37
4	0.058	0.37
5	0.060	0.37
6	0.056	0.37
7	0.062	0.39
8	0.072	0.41
9	0.059	0.38
10	0.057	0.36
11	0.059	0.37

Table 3.4: Statistics on range measurement differences between the sensor values and the waveform decomposition approach (Amiens dataset).

Strip	% 2 iterations	% additional	ξ_1	ξ_2	ξ_{gauss}
Biberach	4.18	3.6	1.51	1.12	1.87

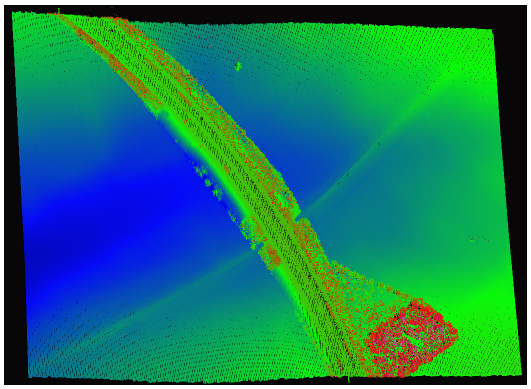
Table 3.5: Comparison of the waveform fitting quality between the proposed Generalized Gaussian decomposition (ξ_1 and ξ_2, for one and two iterations, respectively) and the standard Gaussian approach (ξ_{gauss}) (Biberach data set).

and the estimated one were below half-a-bin. This is totally acceptable but, therefore, this is impossible to give a quantitative evaluation of such differences. Consequently, we have privileged the measure of the range difference between homologous MP and FW points (ϵ, see Figure 3.6.a). Results are shown in Table 3.4, strip by strip for the Amiens dataset. We consider that waveform processing leads to better results that on-line peak detection. $\overline{\epsilon}$ is the mean value of such difference for each strip, and $\overline{\sigma_\epsilon}$ is the associated standard deviation.
Two echoes are considered as "homologue" whether the difference of range values for a given waveform is inferior to 0.75 m. Such an upper bound is necessary to prevent mismatching and set up accordingly to the sensor range resolution. The mean ϵ value is around 0.07 m, with a standard deviation of 0.38. This means, in a Gaussian sense, that 95% of the matched points show an improvement better than 0.83 m $\simeq$ 1 bin, since the sampling rate of the Optech sensor is equal to 1 ns. Finally, considering Figure 3.6.b, we can state that waveform processing allows to generate 3D points with a higher altimetric accuracy of 0.15 m maximum, and with a higher planimetric accuracy of 0.025 m maximum. The latter value can be considered as negligible with respect to the georeferencing process accuracy values and other errors (see [VM10], Chapter 3).

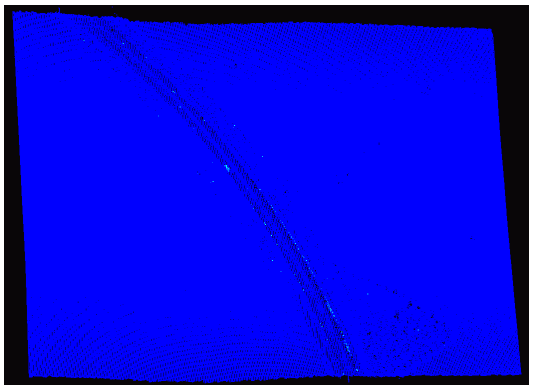
The fitting quality is the global measure of closeness of the reconstructed waveform to the raw signal. Its is evaluated by:

$$\xi = \frac{1}{N}\sum_{i=1}^{N}(y_i - f(x_i|\mathbf{\Theta}^*))^2 \tag{3.3.1}$$

(a) Orthoimage © IGN.



(b) 3D point cloud coloured with the point altitude.

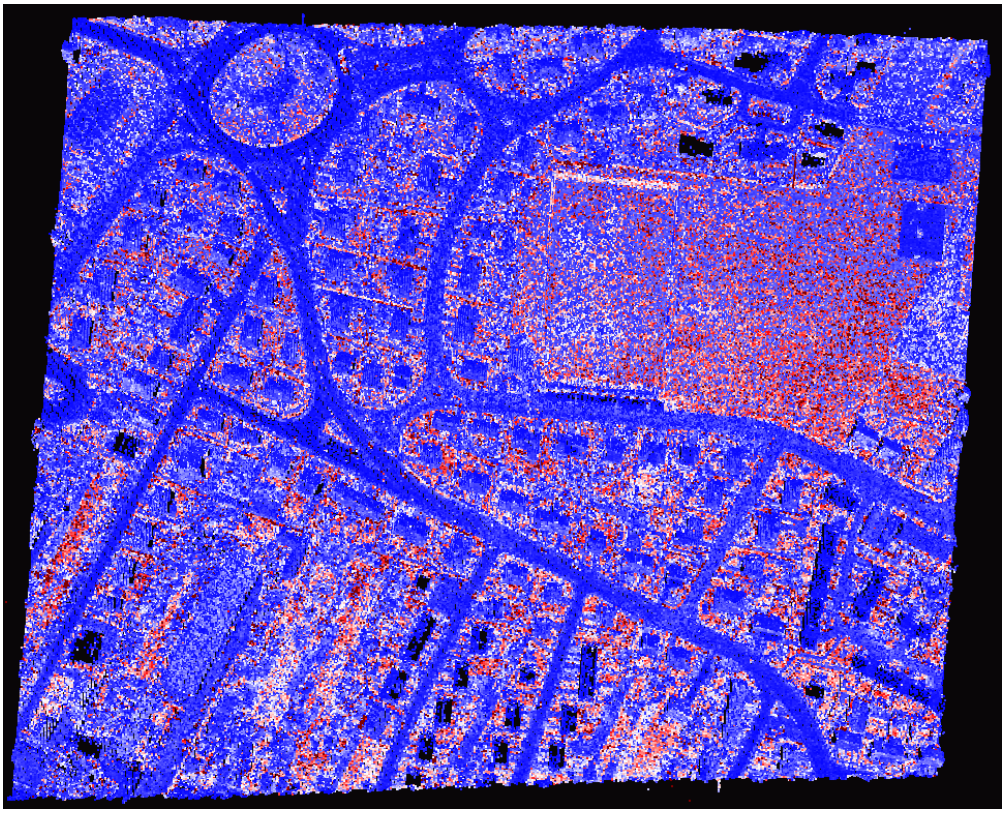


(c) 3D point cloud coloured with the number of fitting iterations (Blue: one – Green: two).

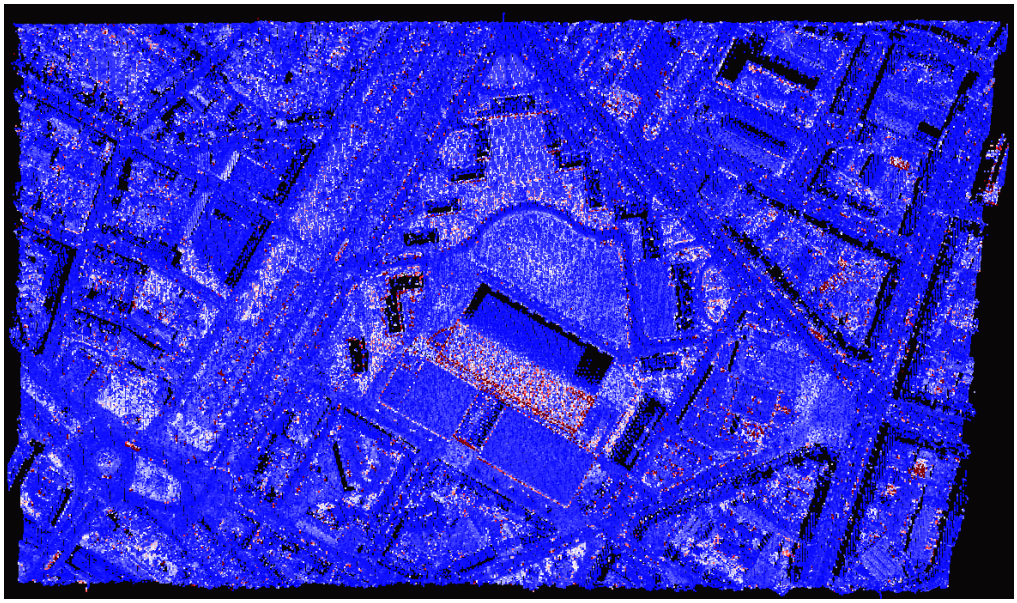
Figure 3.7: Relevance of an iterative waveform decomposition and fitting procedure (Amiens).

where $\mathbf{\Theta}^*$ is the optimal parameter set, estimated with the Levenberg-Marquardt algorithm. Results are shown in the right part of Tables 3.3 and 3.5 for Amiens and Biberach, respectively. Three different figures are provided. ξ_1 and ξ_2 give the mean error per strip with the Generalized Gaussian decomposition method with one and two iterations, respectively, whereas ξ_{gauss} provides quality measures for a one-step Gaussian decomposition (the initial peak detection step is the same as described in Section 3.2.3). The results show that both the introduction of the GG model, and of a second optimization step improve the fitting accuracy (Figure 3.9). The α parameter, introduced to cope with sharp or flat peaks, permits a higher flexibility of the modelling function, and tends to reduce the ξ values *i.e.*, the fitting error. For the purpose of data compression, the Generalized Gaussian model is relevant.

Finally, two 3D point clouds over a residential area and the downtown of Amiens, coloured with respect to the ξ value, are provided in Figure 3.8. Since the quality measure is given for one waveform, two 3D points that belong to the same waveforms therefore have the same value (and colour). One can see the quality criterion exhibits the lowest values for the streets (asphalt ground) and most of the building roofs. Conversely, highest values (*i.e.*, $\xi \geq= 2$) are found on field areas, trees, low above ground objects, and some building roofs depending on the material. This corresponds to more complex targets or targets



(a) Area #1.



(b) Area #2.

Figure 3.8: Waveform fitting quality ξ, displayed for the two areas presented in Figure 3.10.

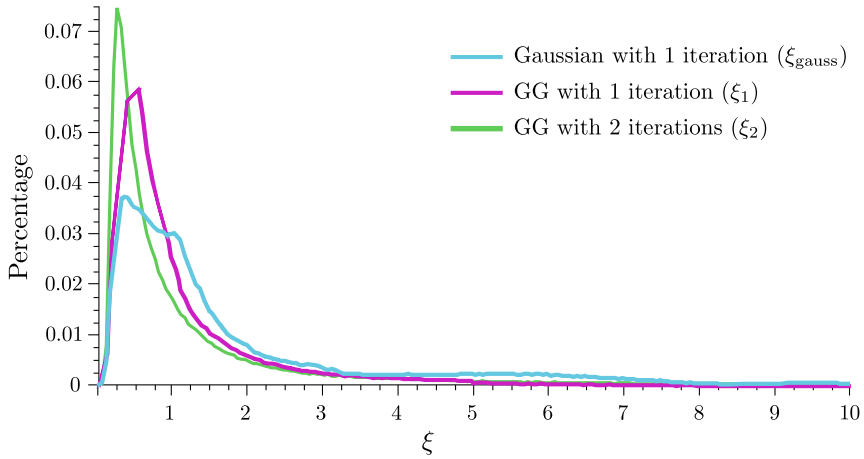


Figure 3.9: Histograms of ξ values for three different scenarii. The values have been computed for Strip 7 of Amiens area.

with highest reflectance (*e.g.*, red flat roof in the middle of Figure 3.8.b), *i.e.*, waveforms with high and narrow pulses. Both cases thus reveal some local misadjustments. Besides, both medium and high values are visible on grass areas depending on the area: lower values correspond to Amiens downtown *i.e.*, public areas in contrast to private gardens in the residential areas. Consequently, the roughness of the terrain has an impact on the fitting accuracy. A close look to the waveforms shows that such waveforms exhibit a slight asymmetry that cannot be perfectly handled by the Gaussian model and its extensions.

3.4 Analysis of the extracted parameters

This section deals with the analysis of the behaviour of the extracted full-waveform features. The georeferencing process after the Generalized Gaussian waveform decomposition has released a 3D point cloud $\{x, y, z\}$, and the so-called FW attributes.

3.4.1 Full-waveform parameters

The Generalized Gaussian waveform decomposition allows to retrieve two kinds of features:

- the model parameters, namely the amplitude A, the width w, and the shape α (see Equation 3.2.4);
- the features that can be computed using the above-mentioned parameters: the Full-Width-at-Half-Maximum $FWHM$, and the backscattered cross-section σ (see Table 3.6).

Parameter μ of the GG function is not a morphological parameter, but indicates the position of the peak maximum. This is therefore the key feature for the georeferencing process. The formulas for deriving $FWHM$ and σ from the Gaussian model parameters are known (see, for instance, Equation 2.6.2). Since the Generalized Gaussian model has been adopted,

Figure 3.10: Four areas of interest over Amiens: (a) Area #1 – (b) Area #2 – (c) Area #3 – (d) Area #4 © IGN.

these formulas are therefore modified (see Table 3.6). In practice, α is close to $\sqrt{2}$ and the $FWHM$ and σ values do not much vary.

Feature	Gaussian	Generalized Gaussian
FWHM	$2w\sqrt{2\ln 2}$	$2\sqrt[\alpha^2]{2\ln 2\, w^2}$
Pulse energy	$A\,w\sqrt{2\pi}$	$\frac{A\,w}{\alpha^2}\Gamma(\frac{1}{\alpha^2})$
Backscattered cross-section	$C_{\text{cal}}\,R^4\,A\,w$	$C_{\text{cal}}\,R^4\,\frac{A\,w}{\alpha^2}\Gamma(\frac{1}{\alpha^2})$

Table 3.6: Full-waveform derived features from the Gaussian and Generalized Gaussian modelling functions.

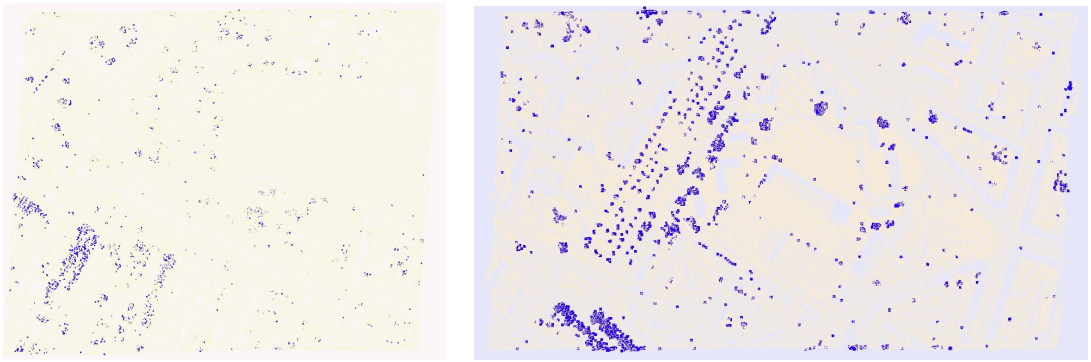


Figure 3.11: Additional points retrieved with our approach (yellow points), compared to those provided by the sensor (dark green points). The areas correspond to those of Figures 3.10a and b.

Figures 3.12 and 3.13 provide several examples of these FW features for both residential, urban, and dense urban areas. It can be clearly noticed that their behaviours both depend on the geometry and the backscattering properties of the targets.

3.4.2 Calibration and correction

As already mentioned and explained in details in Section 2.6.2, both radiometric calibration and feature correction steps are required.

The radiometric calibration is necessary to compute the backscattered cross-section σ for each point. The calibration constant C_{cal} has to be derived. For that purpose, for each lidar strip, small asphalt areas close to the nadir view of the plane are manually delineated. According to the Aster Spectral Library [BHGR09], the reflectance value of such targets is set to 0.2. C_{cal} is calculated for each 3D point lying on the selected areas and labelled as "single" echo of a waveform (both first and last). We follow Equation 2.6.1, derived for the GG model. Finally, the mean value of all the computed constant is taken as reference value. The number of samples points ranges, depending on the strip, from 950 to 3200.

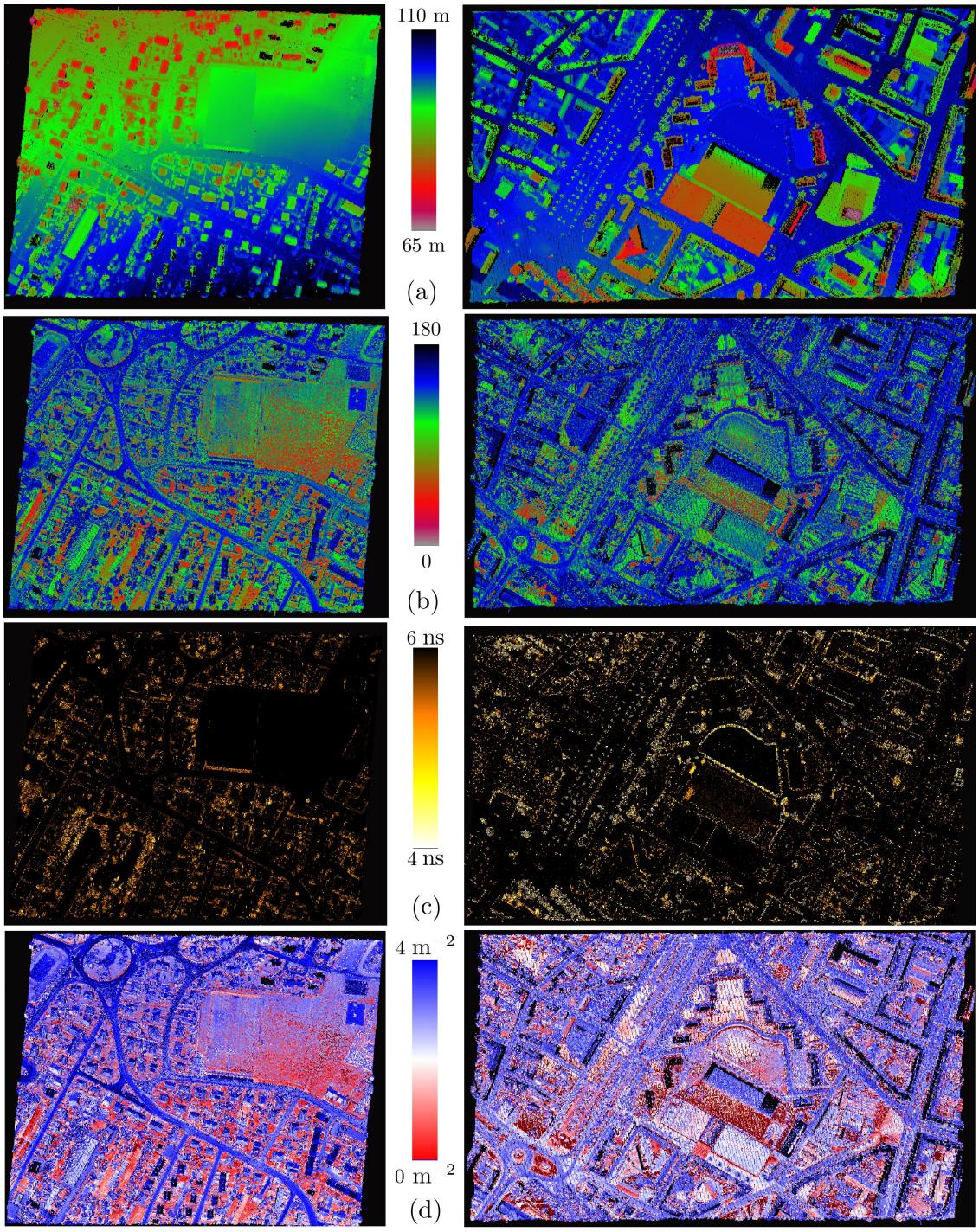


Figure 3.12: Full-waveform features for two areas of Amiens (namely Area #1 on the left column, and Area #2 on the right one). 3D point cloud coloured with respect to the: (a) altitude, (b) amplitude, (c) FWHM, (d) cross-section.

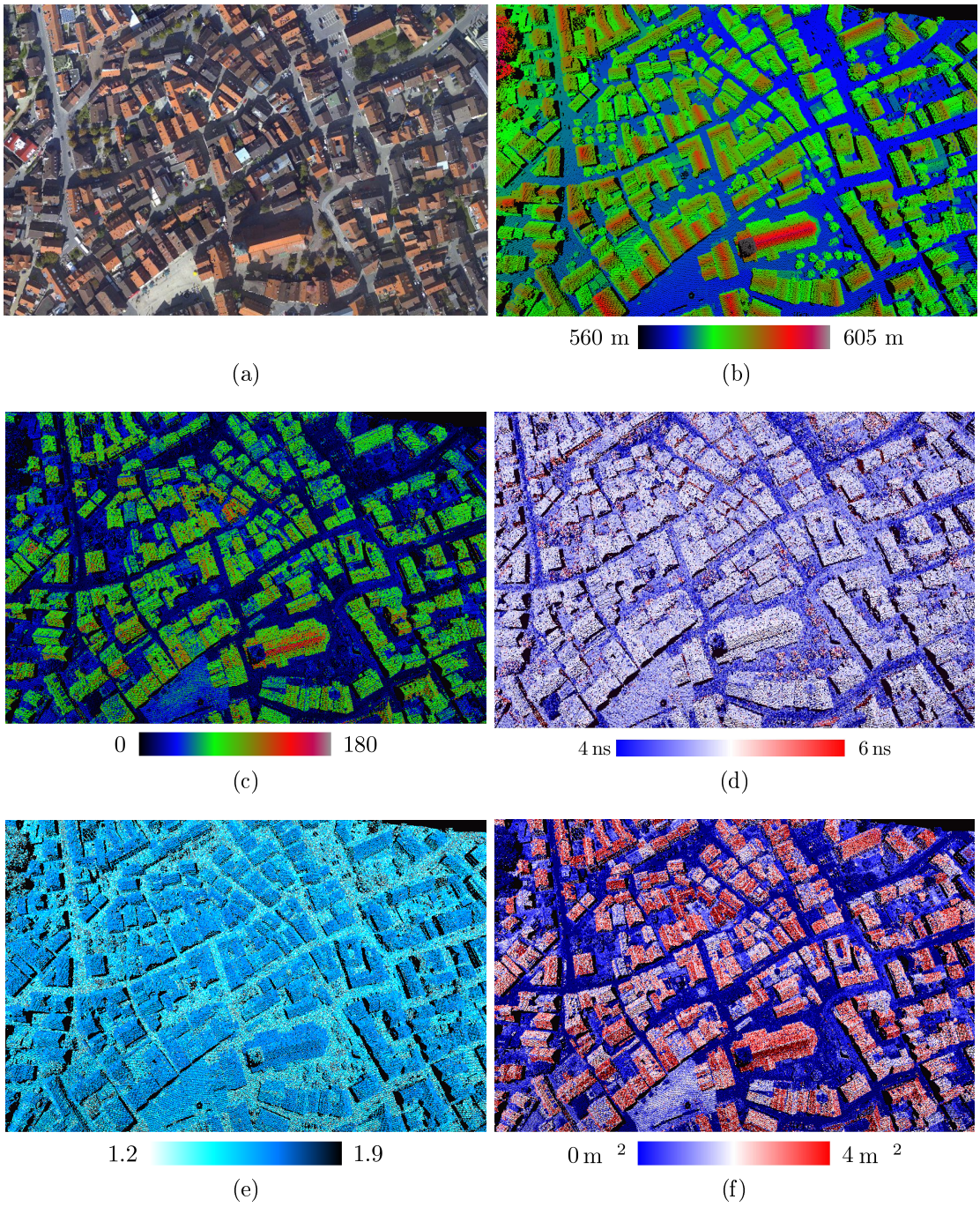


Figure 3.13: Full-waveform features for a small areas of Biberach. (a) Ortho-image. 3D point cloud coloured with respect to the: (b) altitude, (c) amplitude, (d) FWHM, (e) shape (α), and (f) cross-section.

Echo amplitude, width, and shape have to be corrected from the emitted pulse shape fluctuations. This is no longer needed to assume it constant. The T0 pulse is recorded, which permits to take into account its specific shape for each waveform. For that purpose, it is also modelled by a Generalized Gaussian function, and the same optimization procedure is applied than for return signals. The {A,w,α} triplet can be corrected by applying a normalizing factor which is the ratio between the mean feature value of all the emitted pulses of a given strip ($\{\overline{A_e^{\text{strip}}}, \overline{w_e^{\text{strip}}}, \overline{\alpha_e^{\text{strip}}}\}$), and the estimated amplitude of the current emitted pulse ($\{A_e, w_e, \alpha_e\}$).

Figure 3.14 shows the variation in amplitude and length of the T0 pulse for four strips of Amiens. One can notice that the emitted pulse tends to spread (wider with lower amplitude) with the acquisition time. Such a normalization process is therefore necessary to be sensor independent but also not be strip specific. Finally, following Equation 2.6.3, the 3D point amplitude can be additionally corrected from the run length dependency of the laser beam, and the influence of the target geometry, in order to minimize the effect of the slope. Such a process has been carried out for both datasets, and an example of improvement over Biberach is provided in Figure 2.8.

Finally, the improvement can also be assessed with Figure 3.16.a that displays for the whole Biberach area the amplitude histograms before and after correction. One can notice that the small fluctuations on the raw amplitude values are smoothed with the geometrical correction. Such effect comes from the building roofs that compose most of the area of Biberach. Many buildings are composed of gable or hipped roofs and are located on each side of the plane trajectory. Therefore, the laser beam will hit one face with an orthogonal incidence angle (no geometrical influence), and the other side with a significant angle, leading to strong attenuation and lower amplitude values (see Figure 2.8). Each type of building (*i.e.*, with almost the same geometry and radiometry) will therefore generate locally one high value and one low value in the amplitude histogram.

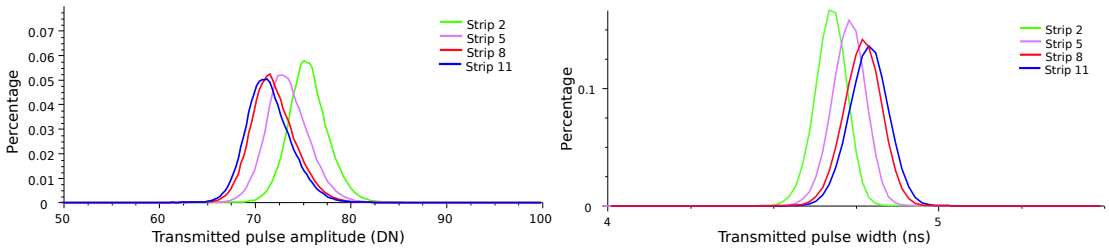


Figure 3.14: Illustration of the transmitted pulse variation with the time. 500,000 points have been randomly picked in four distinct strips of the Amiens dataset to compute the histograms.

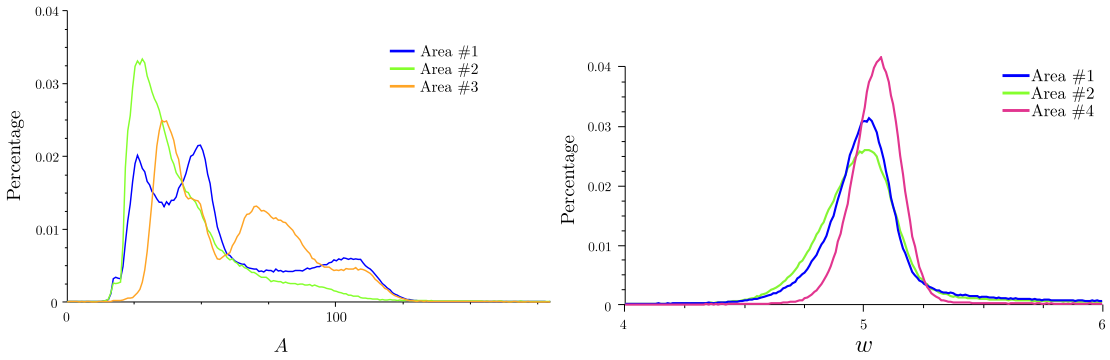
3.4.3 Point-based analysis

The behaviour of four full-waveform features is examined in this section: the amplitude, the width, the shape, and the cross-section. The analysis is carried out with an echo-based approach and not with a segment-based approach. Histograms are computed per feature

first for whole lidar strips. Besides, the point cloud is segmented per echo type (single - first - intermediate - last). Finally, the analysis is performed per various classes of interest. Such a labelling has been carried out manually, and will serve as basis for the ground truth of the classification process (see Chapter 5). On the one hand, we select the three main classes of interest that will be used for subsequent classification: *ground*, *vegetation*, and *building*. Terrain echoes are only composed of single echoes. On the other hand, ground and building classes are splitted for particular in-depth analysis. *Asphalt*, *clay court*, and *grass* areas are delineated for ground points, and several *metal*, *slate*, and *tiled* roofs are labelled for the *building* class. For further analysis, we have also decided for the Biberach area to split two kinds of tiled roofs according to their colour on the associated orthoimage: such classes are called "*orange*" and "*brown*" tile classes.

Amplitude A and width w

Figure 3.15 first shows the general behaviour of these two parameters of four various areas of Amiens. The four areas are visible on Figure 3.10. Area #1 is a residential area mixing vegetated areas and anthropic structures, that is why the associated amplitude histogram is composed of several modes. The amplitude is higher for grasses than woody vegetation due to the fact that the laser energy is scattered more through the multiple layers and less energy is returned to the sensor. In case of dense urban areas, less vegetation is present and targets with poor reflectance properties (such as the asphalt, the cement or slate roofs) prevail: the amplitude histogram thus exhibit lower values (Figures 3.12b 3.15.a). The echo width histograms of these two areas are almost identical, and no conclusion can be made at the moment on the target influence. Area #3 is a region with two distinct crop areas. They have not the same roughness and radiometric properties which explains the two modes in the associated histogram. In such simple conditions, an amplitude-based discrimination is possible. Finally, area #4 correspond to a crop with small areas of vegetation (bushes and trees). Figure 3.15.b shows that the echo width values are higher than for urban areas. Vegetation therefore spreads the laser pulse.

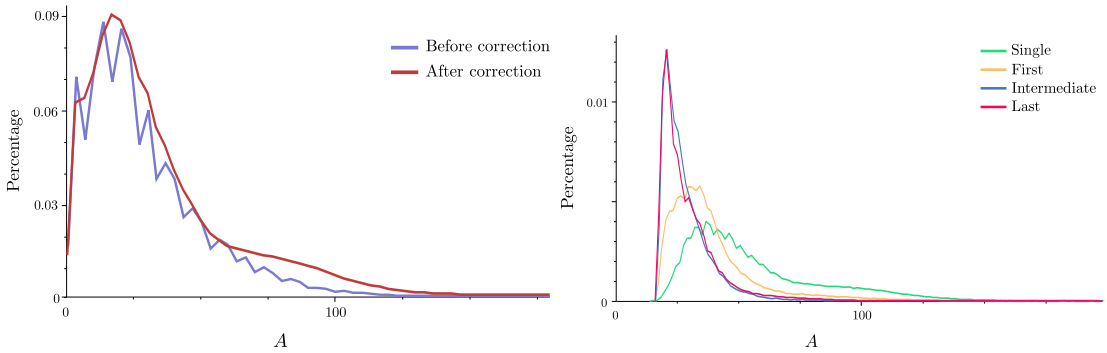


(a) Amplitude. (b) Width.

Figure 3.15: Amplitude and width histograms for the full points of various areas of Amiens (see Figure 3.10).

Figure 3.16.b shows the amplitude behaviour depending on the position of the echo within

the waveform. Since single echoes mainly correspond to building and ground areas, whereas first/intermediate/last pulses stem predominantly from vegetated areas, such histograms are slightly similar to those displayed with respect to the echo label (Figure 3.17.a). Single points exhibit the higher values, and the amplitude decreases with the increasing echo number. First points correspond to the top of the tree canopy and sometimes to building edges. Intermediate and last points have almost the same behaviour. Some last points may be ground points but they are at least second echoes and their amplitude values reflect the loss of laser energy due to multiple reflections.



(a) Before and after correction.

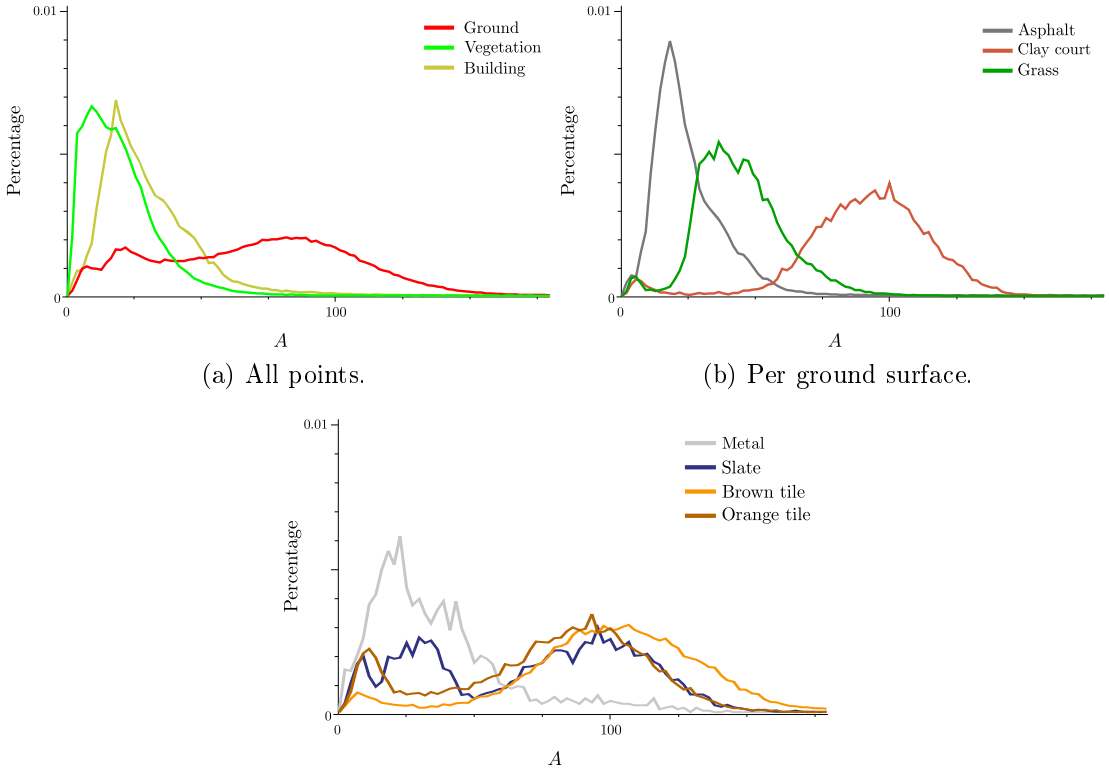
(b) Per type.

Figure 3.16: Echo amplitude (Biberach).

Figure 3.17.a shows that vegetation areas, which are mainly composed of non-single points, exhibit low amplitude values. Indeed the laser pulse is highly absorbed and such behaviour is increased by the multiple reflections. The more echoes occur, the less energy is scattered back with the last ones. Slightly higher values can be noticed for building areas. For the *ground* class, the amplitude values are significantly dispersed. Firstly, this can come from the manual labelling of the point cloud. Delineated ground areas may gather various low above ground items such as cars, street items, etc. with very distinct and unforeseeable behaviours. This is mainly due to the very distinct behaviours of ground surfaces (see Figure 3.17.b). Asphalt areas feature reflectance values (0.1-0.3) leading to low amplitude values, contrary to clay tennis courts and tiled roofs (Figure 3.13.c). Grass points exhibit medium values, which depend on the grass roughness and composition.

Moreover, Figure 3.17.c features amplitude histograms for four different kinds of roofs. Tiled roofs have almost the same behaviour, whereas the metal roofs specifically exhibit low values due to their low reflectance (*e.g.*, large building in the middle of the Area #2 of Figure 3.12b). Finally, vegetation/building/ground classification is not straightforward with the single amplitude feature. Nevertheless, it can be, for instance, a relevant feature for ground surface discrimination (asphalt versus clay courts/grass). It can be noticed on Figure 3.12b both for Area #1 and for Area#2.

Finally, all amplitude histograms feature a low dynamic. The values range from 10/15 to 150, in spite of a digitization process over 8 bits (*i.e.*, until 255). This comes from the calibration process and the survey parameter setting. The record sensitivity is set up such as the emitted pulse and the backscattered signal are fully recorded in amplitude (no



(a) All points.

(b) Per ground surface.

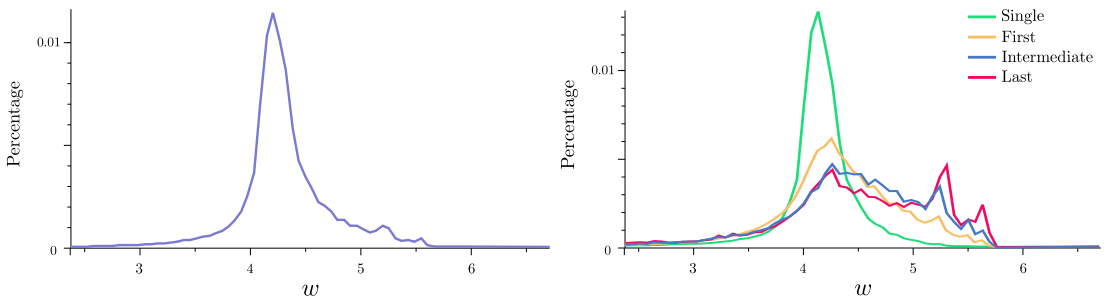
(c) Per roof material.

Figure 3.17: Echo amplitude for various classes of interest (Biberach).

saturation). When such process is not perfectly tunes, this can lead to poor amplitude values that can have a significant impact on the subsequent classification process.

Echo width behaviour is displayed in Figures 3.18 and 3.19. The histogram for all the points over Biberach (Figure 3.18.a) is right skewed and shows little dispersion around the mean value of the emitted pulse ($\overline{w_e}$=4.3 ns). Figures 3.18.b and 3.19.a show that the skewness comes from the intermediate and last echoes *i.e.*, vegetation points. Such histograms are far more scattered than for the building and ground classes. Nevertheless, the two corresponding plots are slightly right-skewed, enhancing the fact that the lidar pulse can be stretched by some kinds of objects. This is particularly the case of building roofs. Figures 3.19.b and 3.19.c, we can see that asphalt surfaces and metal roofs (which are the most flat surfaces of their category) tend to stretch the laser pulse, whereas the clay courts and tiled roofs barely influence the echo width. Finally, grass areas, rougher surfaces, broaden the emitted pulse.

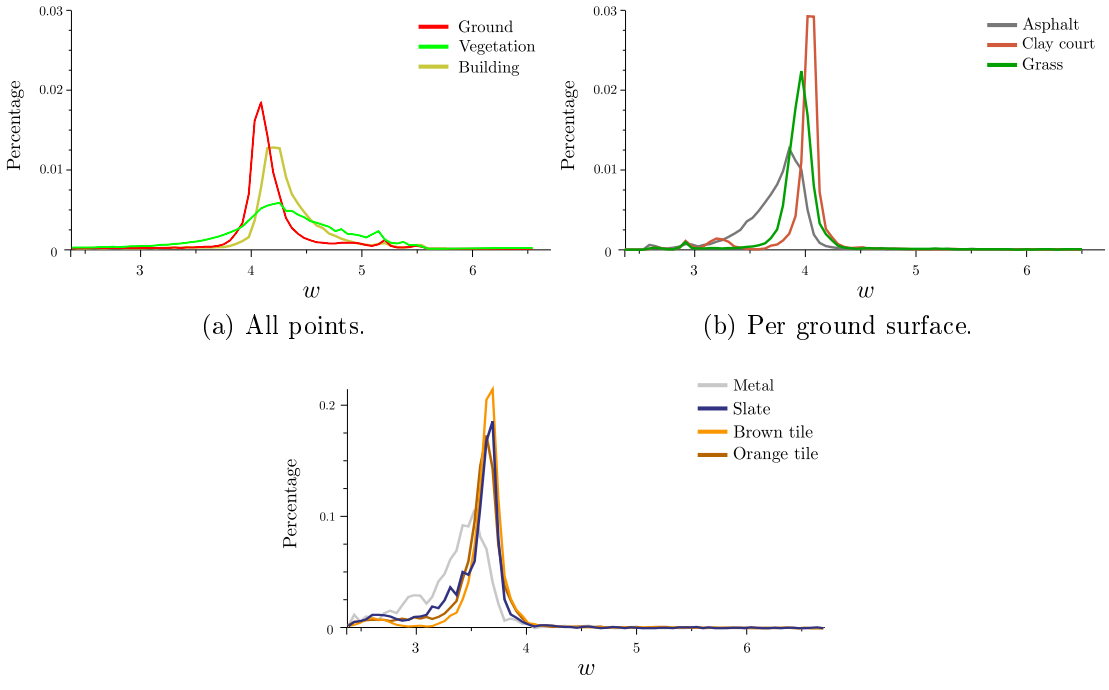
In urban areas, and as already shown in forested and natural areas, echo width may contribute to the classification of vegetated areas. The pulse is wider in such areas (as in Figure 3.13.d). Building edges have also the same effect (see Figure 3.12c). Even if building roofs stretch the pulses more than terrain surfaces, the two histograms are very similar. Finally intra-class discrimination does not seem possible for the building and ground classes, since the histograms of the selected surfaces and roof materials of interest



(a) All points. (b) Per type.

Figure 3.18: Echo width (Biberach).

largely overlap.



(a) All points. (b) Per ground surface.

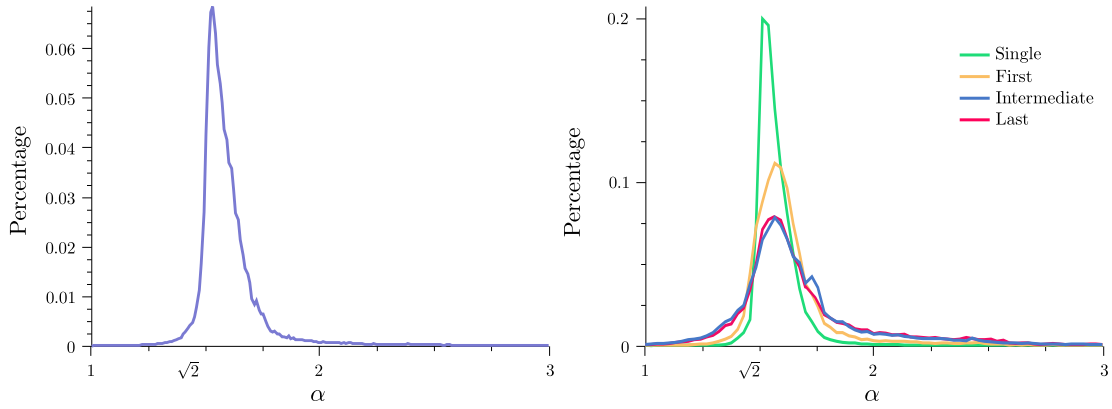
(c) Per roof material.

Figure 3.19: Echo width for various classes of interest (Biberach).

Echo shape α

The echo shape indicates how peaked or how flat an echo is. $\alpha = \sqrt{2}$ means this is a Gaussian curve. Histogram of Figure 3.20.a shows the general non Gaussian behaviour of the lidar echoes. Values range between 1.3 and 2, with a mean value of 1.58. The distribution of the values show a slight asymmetry and a shift to the right. However, most of the values lie between 1.5 and 1.7. We can therefore conclude that such behaviour does not

remote far from the Gaussian assumption. Lower and higher shape values are characteristic of ground points and vegetation, respectively (see Figure 3.13.e). However, the plots of the vegetation and building roof points do not differ. Metal roofs and asphalt surfaces feature the highest values (*i.e.*, flatter peaks) but one can notice that these two classes of interest also feature low echo width and amplitude. Such echoes look consequently flat even if there is no real distortion, conversely to narrow high pulses that look more peaked. This is an effect of the (relatively low) digitization rate of the waveforms. Figures 3.13.e, 3.21.b and 3.21.c eventually enhance the difficulty to use α as a single discriminative feature for classifying terrain surfaces and building roofs in urban areas.

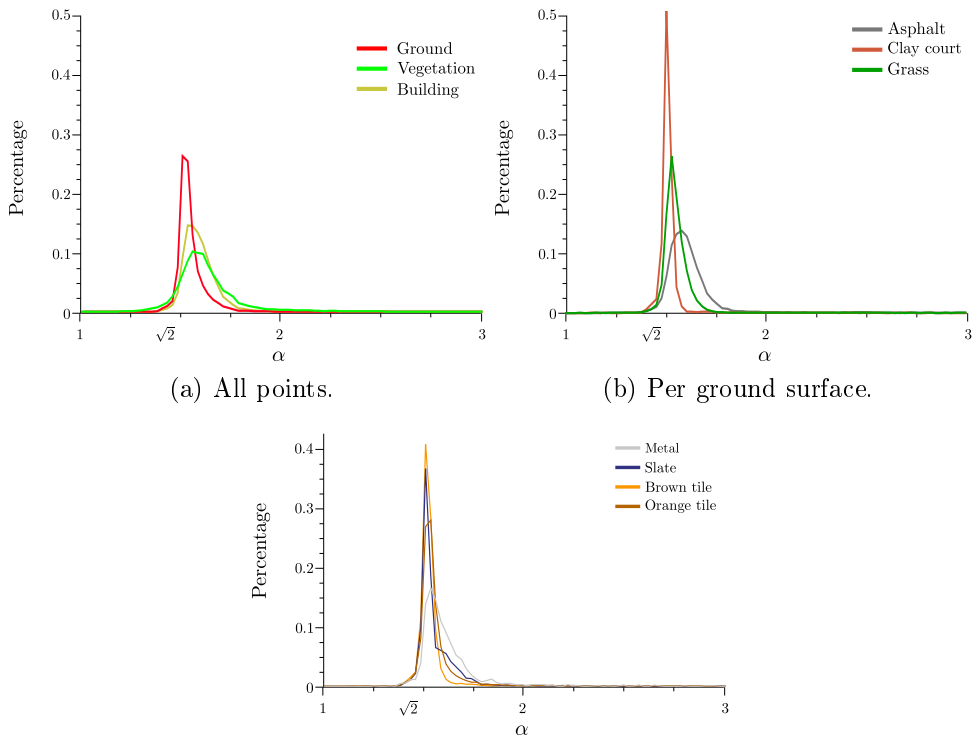


(a) All points. (b) Per type.

Figure 3.20: Echo shape (Biberach).

Cross-section σ

The histogram of σ values computed for the whole Biberach area feature two main modes with a long descending tail (Figure 3.22.a). The cross-section decreases with the increasing position of the echoes within the waveforms. Single and first pulses exhibit higher values than intermediate and last echoes that have narrower histograms and closer to 0 (Figure 3.22.b). Figure 3.23 first confirm the vegetated areas, with smallest targets (with respect to the laser footprint), have the lowest cross-sections. The *ground* class histogram is more scattered than the two other main classes. One can notice in Figure 3.13.f that ground regions can be easily discriminated from building points. However, some roofs feature low σ values since the cross-section is highly influenced by the amplitude that is also influenced by the object reflectance. Furthermore, a close look to the terrain surfaces shows that the asphalt and clay court plots barely overlap. The grass plot lies in between the two latter ones. The σ feature therefore shows a better discrimination between these land-cover classes than the features previously presented. For roof materials, the classes of *tile* and *metal* are distinguished better, having a small overlap. σ is therefore a relevant feature for their discrimination. A visual inspection of the 3D point clouds (Figure 3.12d) assesses such relevance since several land-cover classes are easily distinguishable. Cross-section values of slate roof points are located between these two classes but are not totally overlapping them.

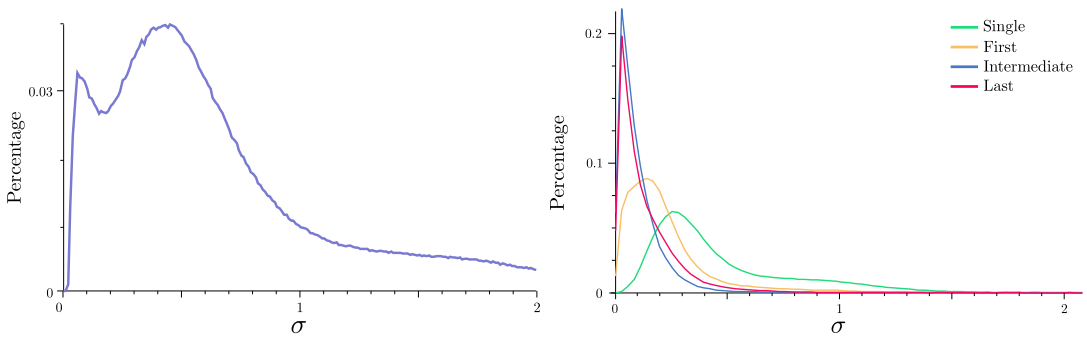


(a) All points.

(b) Per ground surface.

(c) Per roof material.

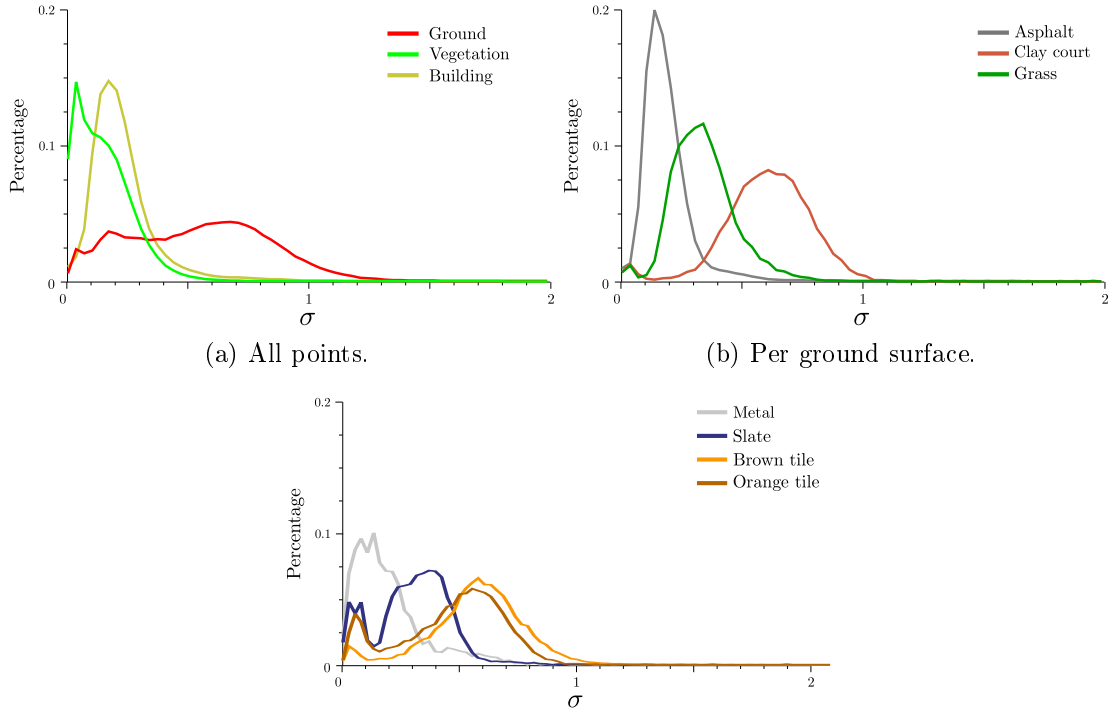
Figure 3.21: Echo shape for various classes of interest (Biberach).



(a) All points.

(b) Per type.

Figure 3.22: Echo cross-section (Biberach).



(a) All points.

(b) Per ground surface.

(c) Per roof material.

Figure 3.23: Echo cross-section for various classes of interest (Biberach).

3.5 Conclusions

We have presented in this chapter a methodology to process airborne lidar waveforms by adopting the standard Gaussian Mixture Model approach. This has already been extensively tackled in the literature. All the echoes of the waveforms are described in the same symmetric analytical function. The number of echoes is first retrieved, and afterwards parameter estimation is performed traditionally using a Non-Linear Least Squares optimization algorithm. We have adopted this framework but have proposed at the same time two main modifications.

Firstly, the performance of such methodology mainly relies on the estimation of the number of modes within the waveforms. Such modes therefore need to be carefully retrieved. This is all the more crucial that incorrect signal processing with a high number of missing points would turn the full-waveform laser scanning to an inefficient technology. Due to the relatively low digitization rate of the waveform, finding the correct number of inflexion points within the lidar signals is not straightforward, especially when dealing with overlapping peaks. We have therefore proposed to smooth the waveforms in order to primarily find a "coarse" number of modes. A first Levenberg-Marquardt optimization step provides us a first reconstruction, which compared to the raw waveform, helps us to assess whether strong differences exist. In this case, a new echo is created where the most important shift exist, and the parameter optimization restarts. Such slight improvement of the standard approach has been proposed by several authors in parallel to this thesis. Results have

shown that two iterations are enough, since we mainly deal with single and double-mode signals. In urban areas, additional points are retrieved on building edges and trees, allowing to fully benefit from the potential of the FW technology, in comparison to multiple pulse sensors. Furthermore, detecting echo location with higher accuracy that on-the-flight signal processing algorithms leads to 3D points with higher geometrical accuracy. In best cases, a higher altimetric accuracy of 0.15 m can be retrieved for the subsequent georeferenced 3D points.

Secondly, concerning the modelling function, we have decided to extend the Gaussian standard by adopting the Generalized Gaussian (GG) model. This is still a symmetric model but it allows to cope to more peaked or flat echoes with one additional parameter α. In the perspective of land-cover classification, this is all the more relevant that such feature may prove to be discriminant for the classes of interest. The waveform processing results show that this improves the fitting accuracy, while still using a simple parametric model. Well-known full-waveforms features, namely the amplitude, the width, and the derived cross-section are consequently better estimated. Radiometric calibration, geometrical correction, and emitted pulse fluctuation normalization also stem from the modelling step of both the waveforms and the emitted laser pulses. These three features show high potential for discriminating our classes of interest (*ground-building-vegetation*) for our landscapes of interest. The newly introduced shape feature is not clearly discriminative since the histograms of α values are almost the same. Nevertheless, the feature will be conserved in Chapter 5 in order to assess its genuine relevance. Further intra-class analysis has eventually been performed for the *building* class and the *ground* class. We can say that the echo cross-section has revealed to be a relevant feature, but more in-depth analysis and subsequent classification must be carried out for full conclusions.
The main limitation of the GG model compared to the Gaussian one may come from the fact that adding one parameter in the optimization process may lead to a higher number of inconsistent results. However, statistics have shown that the same order of magnitude is reached for both the Gaussian and the GG models (approximately 1:700,000).

The proposed workflow has not been specifically designed for urban areas, and is not limited to small-footprint airborne lidar waveforms. The method can be tailored for other applications and for a required fitting accuracy. Changing the modelling function allows to retrieve new parameters that may be discriminant for classifying the resulting 3D point clouds. However, it should be first verified that its parameters can be computed in the optimization process. Least-squares and EM approaches require the knowledge of the first derivatives of the modelling function for each parameter. The Generalized Gaussian model can still be handled by such techniques since this is a simple extension of the Gaussian model. For more complex functions, other algorithms have to be developed.
Moreover, the issue of decomposing terrestrial or spatial lidar waveforms can also be tackled with our approach. For other kinds of landscapes and altimetric profiles, the key point is the initial number of echoes as well as the rule for aggregating new echoes in the second optimization step. For instance, for spatial waveforms over forested areas, as illustrated in Figure 2.3, two options are conceivable. If we aim to retrieve each layer of the tree canopy, a slight smooth of the signal coupled with multiple echo addition based on high residuals should be effective. For a coarse strategy, *i.e.*, in an area-based approach, favoring the main trends of the signal, the raw waveform should be smoothed with a larger kernel and

a single optimization step should be enough.

To top it all, on the one hand, we can say the Generalized Gaussian model, as an extension to the standard Gaussian function, still provide a suitable trade-off between a simple solution with some relevant features and that can be easily handled, and a correct reconstruction accuracy for a whole area of interest. One the other hand, such simple solution can also be considered as very restrictive. First, the symmetric assumption does not always hold: the fitting quality criterion can be relatively superior to the expected even for single-mode waveforms. This enhances the asymmetric behaviour of the waveforms, and more precisely of backscattered signals within the waveforms. Better adjusting the waveforms would lead to better signal reconstruction, and to top it all, to better data compression. Such issue is clearly at stake due to the large amount of data provided by FW sensors. Furthermore, in order to foster information extraction, it may be of interest to deal with full-waveform features, requiring new modelling functions.
The following part introduce a strategy to cope with the two above-mentioned issues at the same time using a stochastic approach based on a marked point process.

Chapter 4

A marked point process approach for waveform modelling

4.1 Motivation

As discussed in Chapter 3, the Gaussian assumption does not always hold, and approximating the waveforms by a sum of Gaussians may be inadequate, depending on the required accuracy, the application and the landscape. In the lidar equation (see Equation 2.1.8), the emitted power P_t and the echoes are usually described by Gaussian functions, but this is not always correct, and waveforms can therefore be composed of modes with non-similar shapes (symmetric and asymmetric). To remove both, the broadening and the asymmetric effects caused by a varying P_t on the received waveforms, a deconvolution step is usually carried out, using, for instance, matched filtering, Wiener filtering [JS06], or B-splines [RBP10]. Indeed, target cross-sections are physical parameters which are independent of the emitted pulses. However, asymmetric peaks are also reported after deconvolution, that is why we have to depart from the symmetric assumption. Such asymmetry has been pointed out when quantifying the fitting accuracy of the Generalized Gaussian decomposition approach. It can stem from various conditions (see Figure 4.1):

- Two overlapping Gaussian echoes of distinct amplitude can lead to a single left or right-skewed pulse. A simulated case illustrates in Figure 4.1.a what is usually found in tree canopies, with right-skewed echoes.
- Waveforms acquired with small-footprint sensors are highly influenced by the local geometry of the intercepted surfaces. They can be positively or negatively skewed by rough surfaces like vegetated areas (upper part of the tree canopy, hedges) or ploughed fields (Figures 4.1.b and 4.1.c).
- Waveforms received from medium and large-footprint sensors are bound to hit a sequence of targets both dispersed in planimetry and in altimetry. They will exhibit non-symmetric altimetric distribution leading to complex pulse shapes (Figures 4.1.d-4.1.e- 4.1.f).

More advanced lidar waveform processing can be consequently motivated by two main reasons. First, the fitting quality measure has shown that, for small-footprint sensors,

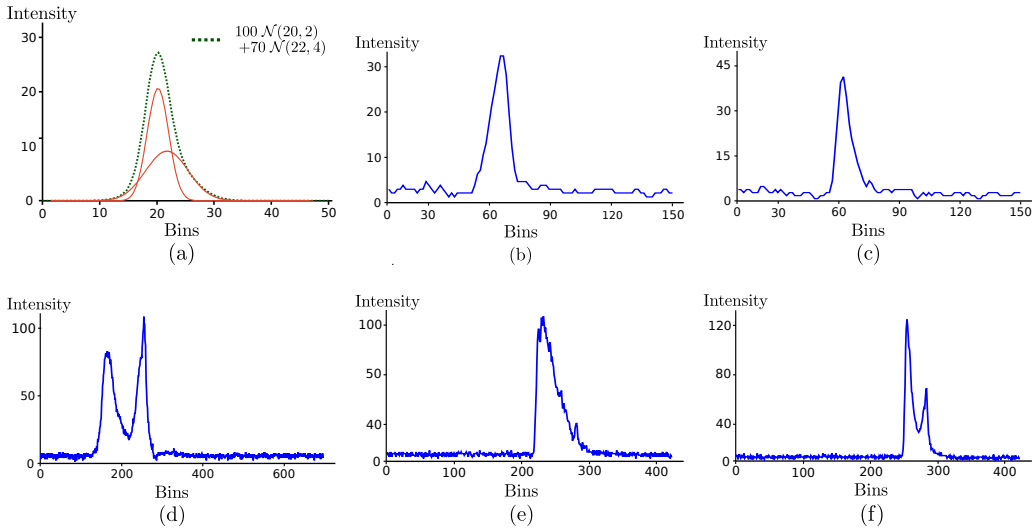


Figure 4.1: Types of waveform signals (a) Right-skewed waveform (dark green dashed curve) simulated with two Gaussian curves (red continuous curves). (b) and (c) Waveforms resulting from the small-footprint laser pulse backscattered from a hedge and a tilted building roof (Biberach dataset). (d), (e), and (f) Several illustrations of complex asymmetric waveforms acquired on tropical vegetated areas with LVIS sensor.

even for simple single mode waveforms, the reconstructions are not always perfect, and the Generalized Gaussian model does not solve the problem. A new modelling function with the adequate behaviour has to be introduced. However, in a large majority of cases, the "symmetric" assumption is valid, and may be conserved for the corresponding echoes. Secondly, we must remind that waveform processing is performed in order to generate a 3D point cloud with additional features. The application of interest is 3D land-cover classification. The features, derived from the modelling function, should be discriminant for the classes of interest. Consequently, in order to foster information extraction, and in addition to the standard echo amplitude, width, and derived cross-section, other morphological parameters may be relevant. The relevance of these features may only be correctly assessed embedded with the selected classifier. Thus, our approach needs to be flexible and be designed to be able to deal with several kinds of functions without modifying the algorithm.

No solution has yet been proposed to transform the well-known parameter estimation problem into an optimal model selection problem for each mixture component where (i) the number of components is unknown, (ii) the parametric models come from a predefined library, and (iii) the most suitable model for each echo is not known beforehand. The coarsest solution would be to adapt the approach proposed in Chapter 3. After initial peak detection, a greedy algorithm may look for the optimal parameter values for each peak for the set of proposed functions *i.e.*, least-squares matching would be performed for each combination. Finally, the minimal value of a well-designed energy would give the reconstruction. However, this suffers from three main drawbacks. Firstly, each echo is segmented separately and no real coherence exists within the waveform. Secondly, such process does not allow to correct peak under-detection and would lead to sparser point clouds. Finally, the least-squares optimization process still limits the introduction of new

functions. As a consequence, many models that we would like to test cannot be well handled since specific techniques are required to estimate their parameters (*e.g.*, the Weibull function, see [Men63]). Besides, modelling functions with a higher number of parameters may be prohibited since a high number of inconsistent results would appear.
Such combinatorial optimization problem may be solved, for instance, using simply simulated annealing, Ant Colony Optimization [DS04], or Genetic Algorithms [YS94]. Stochastic methods based on **marked point processes** [BL93; Lie00], that also use simulated annealing, are very promising for addressing such issue, and have been adopted here. These models, which allow the sampling of parametric primitives, while taking into account complex interactions, have shown very good potential for many applications in remote sensing [HSB03] and especially in image analysis aiming at the extraction of line networks [LDZ05; SSZ07; TP09], vegetation [PDZ05], or 3D urban objects [ODZ07; LDZPD10]. The sampling of the primitives is performed by Markov Chain Monte Carlo (MCMC) techniques [Has70] which exhibit very good signal reconstruction properties [PADF02].

Each waveform is considered as a configuration of objects. Our objects of interest are parametric functions, and we aim to model each mode of the signal. The configurations are assessed using an energy term. The global minimum of this energy is then found by applying a Reversible Jump Markov Chain Monte Carlo (RJMCMC) sampler embedded into a simulated annealing scheme [Gre95]. The Reversible Jump framework is adopted since the number of objects/echoes is not known beforehand.

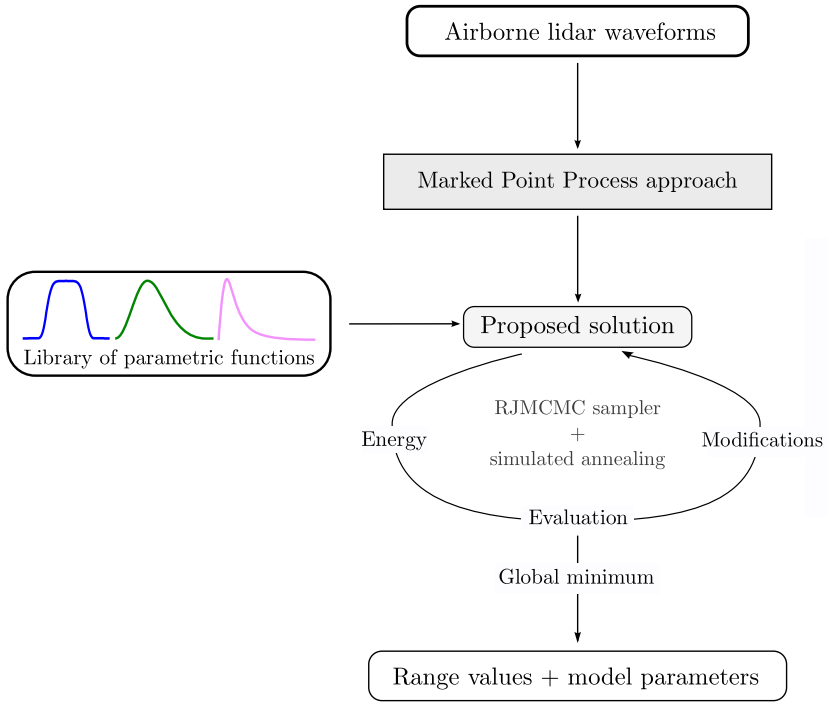


Figure 4.2: Overview of the waveform decomposition and fitting process.

Such techniques have been already adopted immediately before in [HMWG07] where a specific model composed of four exponential parametric functions is fitted to lidar intensity histograms of data affected by significant background noise. The model estimate is

used for counting and locating the reflected returns from surfaces, as well as retrieving their amplitudes. It thus provides an effective algorithm for 3D ranging, all the more since prior knowledge can also be incorporated into the model. However, this approach is not suitable for our airborne lidar waveform: the parameters of the underlying shape model can vary, but this increases dramatically the dimensions of the parameter space and makes the problem much more complex. Thus, the authors of [HMWG07] assume all the peaks of the signals to have a similar underlying shape model, an invalid assumption in our case.

Our model presents several interesting characteristics compared to conventional symmetric (Gaussian) Mixture Model approaches:

- *Multiple function types* - The joint sampling of multiple functions types allows to deal with various parametric functions. First, by using a library of shapes, more accurate estimates are performed compared to classical approaches such as the Gaussian mixture model (see [SGKR09] and Figure 4.1). Secondly, by selecting the most suitable function for each peak, which is unknown beforehand, the estimated parameters are more discriminant for a subsequent classification.
- *Lidar physical knowledge integration* - Prior information on lidar waveform characteristics can be introduced in the energy of the stochastic model formulation (*e.g.*, maximal number of echoes). This permits to get a more realistic model and to achieve better results.
- *Efficient exploration of configuration spaces* - A MCMC sampler associated with relevant proposition kernels allows us to avoid exhaustive explorations of large configuration spaces, which can be both continuous and discrete. It is particularly efficient when the number of functions is unknown.

Eventually, the Reversible Jump Markov Chain Monte Carlo (RJMCMC) [Gre95] algorithm is attractive because in a multi-object framework it can deal with parameter estimation and model selection jointly in a single paradigm.

Section 4.2 first describes the parametric functions that compose the library. Three models have been selected, including the Generalized Gaussian function. Then, the marked point process is briefly presented in Section 4.3. Our goal is not to detail all the mathematical background of point process. Therefore, readers may, for instance, refer instead to [Lie00] for further explanations. Next, the model energy is described in Section 4.4, and Section 4.5 deals with the optimization of the proposed energy. Finally, simulated experiments show the relevance of such an energy, and results for various kinds of signals are presented in Section 4.6. Finally, conclusions and perspectives are drawn in Section 4.7.

4.2 Library of modelling functions

As underlined in Section 4.1, the set of functions has to be chosen carefully since the function parameters shall be used subsequently for classifying the lidar data (presented in Chapter 5). Three options are conceivable:

- Find a single function with a very high flexibility allowing to deal with symmetric, right and left-skewed echoes at the same time. For example, the Levy-skew α stable

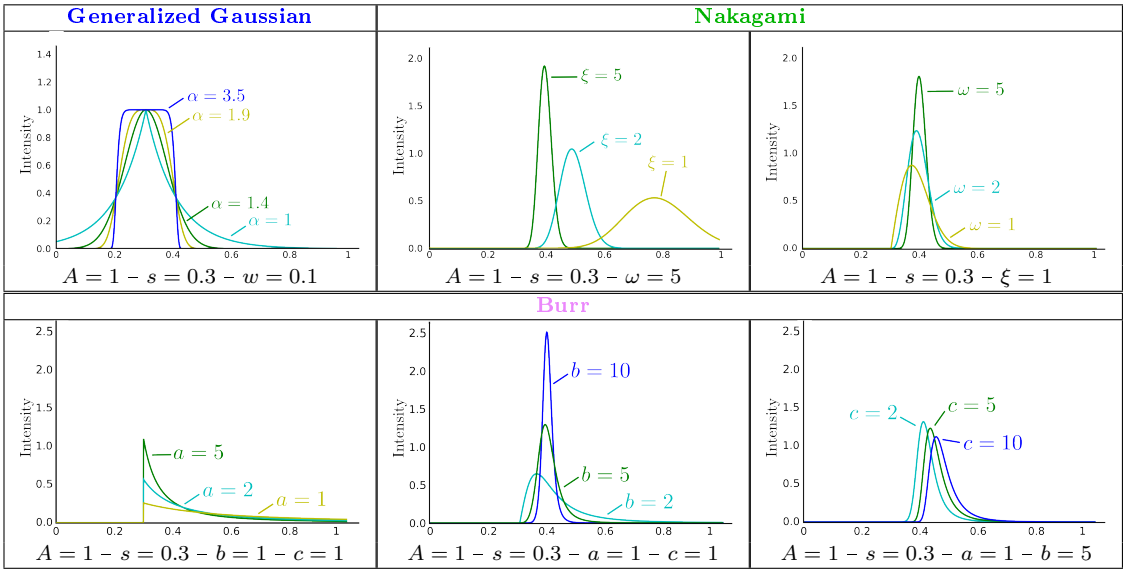


Table 4.1: Overview of various parameter set-ups of the three modelling functions of the library.

function gathers these properties [NS95]. However, it features one main drawback that prohibits its use: its analytical form is bulky, which would result in unfeasible computational load for bin-to-bin comparison.

- Integrate much more models (5-10) to be able to deal with a large variety of lidar waveforms. The process would be more time consuming since the selection of the best-fit function would be a problem of higher dimensionality. A close look to waveforms both from large-sized and small-sized footprints reveals that lidar echoes eventually do not exhibit a large range of distinct behaviours. Therefore, it would be difficult to qualify the process and assess whether each echo is fitted by the "best" model. Introducing models with similar behaviours would lead to random results since mostly noise would govern the choice of model.

- Select a few number of functions (2-4) . The process will be robust if the library is limited to very discriminative modelling functions, and with limited computing time. Each model will fit its own class of echoes.

The strategy with few models has been adopted. Three different distributions are chosen to model the waveforms. They have been selected among a large variety of mathematical functions that arise in physical and engineering problems [AS64].

The Gaussian and **Generalized Gaussian** (GG) models have been shown to fit most of the echoes of small-footprint lidar waveforms in urban areas (see [WUD+06] and Chapter 3). They allow to model symmetric echoes which form the majority of lidar signals. Therefore, the GG model is conserved. The properties of GG function are described in Section 3.2.2.

Nevertheless, the Gaussian assumption does not always hold. Asymmetric echoes are observed within waveforms corresponding to ground surface or tree canopy (Figure 4.1).

Thus, many waveforms exhibit heavier tails and require a more flexible parametric characterization. Moreover, the GG model gives the amplitude, width, and flatness for only symmetric echoes.
Other kinds of functions must therefore be included: functions able to fit asymmetric peaks and those which can cope with both left- and right-skewed curves which therefore deliver other parameters than those provided by the GG model. Several functions have been tested, such as Weibull, Fisher, or Generalized Gamma models, but have not been selected since they were not enough flexible for our purpose. The Nakagami function features all the requested properties, except the ability to fit peaks with heavy tails, that is why the Burr function has been added.

The **Nakagami** distribution is a generalization of the χ distribution and can model right-skewed and left-skewed distributions with a skewness/spread parameter ω:

$$f(x \mid A, s, \xi, \omega) = A \frac{2\,\xi^{\xi}}{\omega\Gamma(\xi)} \left(\frac{x-s}{\omega}\right)^{2\xi-1} \exp -\xi \left(\frac{x-s}{\omega}\right)^{2} \tag{4.2.1}$$

When ω increases, the peak becomes narrower and more symmetric. Scale parameter ξ controls the peak width: large ξ leads to narrow peaks of higher amplitude. The Nakagami function is traditionally used to model Synthetic Aperture Radar (SAR) images to estimate their amplitude probability density functions as well as for subsequent classification [TNTM04]. A large body of literature has presented and studied probability density functions so as to model the dispersion of the received signals produced by different objects, using either theoretical or heuristic models [MZS06; KMSZ08].

Finally, the **Burr** function is especially useful to model asymmetric modes with two shape parameters. It enables to fit right-skewed peaks that the Nakagami model cannot handle. It is a generalization of the Fisk distribution thanks to the parameter c. The scale parameter is a, and b and c are two shapes parameters (b has the same effect as the ω parameter for the Nakagami function). The ratio between peak amplitude and skewness is tuned by c.

$$f(x \mid A, s, c, a, b) = A \frac{bc}{a} \left(\frac{x-s}{a}\right)^{-b-1} \left(1 + \left(\frac{x-s}{a}\right)^{-b}\right)^{-c-1} \tag{4.2.2}$$

For airborne lidar waveforms, there is no physical background justifying their use. Furthermore, contrary to the Gaussian model, their parameters cannot be, at the moment, directly converted into physical parameters: no calibration method has been developed yet, as for the amplitude feature. On the one hand, we admit that there is no physical entity exclusively attached to these curves. On the other hand, they enable us to handle asymmetric peaks and therefore we expect their application will outperform standard approaches. These distributions are defined in continuous domains. Table 4.1 provides some representations of these functions with critical parameter variations.
Finally, one can notice that the library can evolve and be extended to other models to become more generic and be able to deal with a larger scope of signals.

4.3 Marked point processes

Marked point processes are stochastic tools which have been introduced in signal and image processing by Baddeley and Van Lieshout [BL93], and extended further in [RS98; PG98;

Lie00]. These models can be considered as an extension of conventional Markov Random Fields [GG84] such that random variables are not associated with signal values but with parametrical functions describing the signal. An overview of marked point processes is given below. We mainly follow the theory fully presented in [LDG10].

4.3.1 Point processes

A configuration x is a unordered finite set of points $\{x_1, ..., x_n\}$. X is a measurable mapping from an abstract probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to Ω. Ω is the configuration space, $\mathcal{A}$ is a σ-algebra , and $\mathbb{P}$ the associated measure. Ω can be written:

$$\Omega = \bigcup_{n \in \mathbb{N}} \Omega_n \tag{4.3.1}$$

$K = [0, L_{\max}]$ ($L_{\max}$ is the length of the 1D signal) is a continuous bounded set. For every Borel set $B \in K$, $N_X(B)$ defines the number of points of X in B. A *point process* X of points in K is a measurable mapping that associates with an event $\omega \in \Omega$ a configuration of points $X(\omega) = \{x_1, ..., x_{n(\omega)}\}$ with $x_i \in K$. $n(\omega)$ represents the number of points associated with the event ω. Accordingly, a *point process* is a random variable whose realizations are random configurations of points.
The most random point process with respect to the entropy is the *Poisson point process*. This is the reference point process. We consider $\nu(.)$ as a positive measure on K. A Poisson point process X with intensity measure $\nu(.)$ has the two following properties:

- For every Borel set $B \subset K$, $N_X(B)$ follows a discrete Poisson distribution with the mean $\nu(B)$, *i.e.*:
$$P(N_X(B) = n) = \frac{\nu(B)^n}{n!} \, \mathrm{e}^{-\nu(B)}.$$
- If the Borelian sets $B_1, ..., B_l$ (finite sequence) are disjoint, the random variables $N_X(B_1), ..., N_X(B_l)$ are independent.

The Poisson process induces a complete spatial randomness, given by the fact that the positions are uniformly and independently distributed.

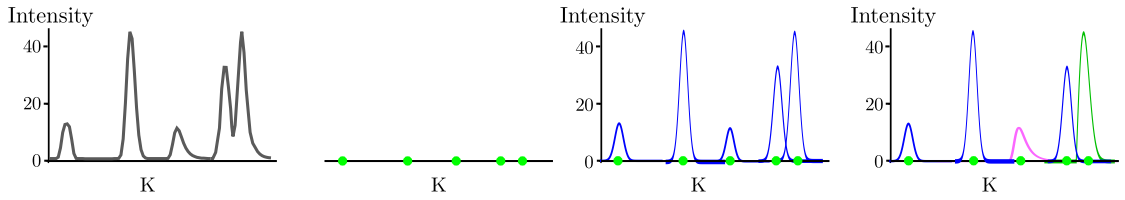


Figure 4.3: Illustration of various point processes - **From left to right**: a 1D signal defined on the support K, realizations of a point process on K, a marked point process of Gaussian functions, and a point process specified by a library of various functions.

4.3.2 Density of a point process

An interesting feature of point processes is the possibility of defining a point process distribution by its probability density function. In most applications, this is not realistic to

assume points are scattered randomly. Poisson processes are useful to build more complex models. Complex point processes introducing both, consistent measurements with data and interactions between points, can therefore be defined by specifying a density with respect to the distribution of a reference Poisson process. $\nu(.)$ is the intensity measure of a Poisson process. $h(.)$ is a non-negative function on the configuration space $\mathcal{C}$ (see Section 4.3.3). Then, the measure $\mu(.)$ having a density $h(.)$ with respect to $\nu(.)$ is defined by:

$$\forall B \in \mathcal{B}(\mathcal{C}),\ \ \mu(B) = \int_B h(\mathrm{x})\nu(d\mathrm{x}) \tag{4.3.2}$$

If $0 < \mu(B) < \infty$, then μ can be normalized to make a probability measure π defined by: $\mu(B)/\mu(\Omega)$. A Gibbs energy $U(x)$ can also be used to specify a point process. The density $h(\mathrm{x})$ of a configuration x is then formulated using the Gibbs equation:

$$h(\mathrm{x}) = \frac{1}{Z}\, \mathrm{e}^{-U\mathrm{x})} \tag{4.3.3}$$

where Z is a normalizing constant such that $Z = \int_{\mathrm{x}\in\mathcal{C}} \mathrm{e}^{-U(\mathrm{x})}$. Defining the Gibbs density of the associated marked point process with respect to the Poisson measure allows to reduce the issue to an energy minimization problem. Generally, a Monte Carlo Markov Chain sampler coupled with a simulated annealing is used to find the maximum density estimator[1] $\widehat{x} = \arg\max h(.)$. This optimization process is particularly interesting since the density $h(.)$ does not need to be normalized. Thus, the complex computation of the normalizing constant Z is avoided.

4.3.3 Marks and object library

The configurations of points described so far only include simple points of $\mathbb{R}$. To describe random configurations of geometrical objects, random marks are added to each point. In order to model signals in terms of parametric functions, it is possible to extend a point process by adding specific *marks* that associate a parametric function (also called an *object*) to each point[2].
A *marked point process* in $S = K \times M$ is a point process in K where each point is associated with a mark from a bounded set M (see Figure 4.3).
The finite configuration space of n objects belonging to S, $\mathcal{C}$, and can be expressed as follows:

$$\mathcal{C} = \bigcup_{n\in\mathbb{N}} (K \times M)^n \tag{4.3.4}$$

For more details on point processes and their applications, the reader should refer to [Lie00].

Usually, the marked point process based models [LDZ05; ODZ07; SSZ07; TP09; GC09] use a single type of object. Some authors [LDG10] have extended the conventional framework in order to sample various kinds of objects extracted from a library. The mark space M

[1] This estimator corresponds to the configuration minimizing the Gibbs energy $U(.)$, *i.e.*, $\widehat{\mathrm{x}} = \arg\min U(.)$.
[2] In many cases, the point corresponds to the mean of the function.

associated with this library is then specified as a finite union of bounded subsets M_q:

$$M = \bigcup_{q=1}^{N_s} M_q \tag{4.3.5}$$

where each mark space M_q corresponds to one of the N_s specific object types. This extension of the marked point processes, which is able to deal with objects having different numbers of control parameters, will be used in the following.
The reference process for building our model is a uniform Poisson process on the space K of the positions within the 1D signal. The objects are the three modelling functions described in the previous section: Generalized Gaussian, Nakagami, and Burr. All the objects have between four and five control parameters. The parameters, detailed in Table 4.2, represent the marks of the object types. They are all defined in continuous domains. The chosen set includes all basic objects used in the conventional marked point process based models.

Object type	Marks	Definition domains
Generalized Gaussian	I, s, w, α	$[I_{\min}, I_{\max}] \times [s_{\min}, s_{\max}] \times [w_{\min}, w_{\max}] \times [\alpha_{\min}, \alpha_{\max}]$
Nakagami	I, s, ξ, ω	$[I_{\min}, I_{\max}] \times [s_{\min}, s_{\max}] \times [\xi_{\min}, \xi_{\max}] \times [\omega_{\min}, \omega_{\max}]$
Burr	I, s, a, b, c	$[I_{\min}, I_{\max}] \times [s_{\min}, s_{\max}] \times [a_{\min}, a_{\max}] \times [b_{\min}, b_{\max}] \times [c_{\min}, c_{\max}]$

Table 4.2: Parameter definition of the marks.

4.4 Energy formulation

Let x be a configuration of parametric functions (or objects) x_i extracted from the above library. The energy $U(\mathrm{x})$ measures the quality of x such as:

$$U : \mathcal{C} \to \mathbb{R} \tag{4.4.1}$$
$$\mathrm{x} \to U(\mathrm{x}) \tag{4.4.2}$$

U is composed of both a data term $U_{\mathrm{d}}(\mathrm{x})$ and a regularization term $U_{\mathrm{p}}(\mathrm{x})$ such that:

$$U(\mathrm{x}) = (1-\beta)\, U_{\mathrm{d}}(\mathrm{x}) + \beta\, U_{\mathrm{p}}(\mathrm{x}) \tag{4.4.3}$$

where $\beta \in [0, 1]$ tunes the trade-off between the data term and the regularization.

4.4.1 Data term

The data energy steers the model to best fit the lidar waveform. It measures the likelihood of the waveform w.r.t. the objects. It can be obtained by computing a distance between the given signal $\mathcal{S}_{\mathrm{data}}$ and the estimated one $\mathcal{S}_{\mathrm{x}}$, which thus depends on the current objects on the configuration x:

$$U_{\mathrm{d}}(\mathrm{x}) = \sqrt{\frac{1}{|K|} \int_K (\mathcal{S}_{\mathrm{x}} - \mathcal{S}_{\mathrm{data}})^2} \tag{4.4.4}$$

The term $U_d(x)$ measures the quadratic error between the signal and its estimation: it allows to be sensitive to high variations *i.e.*, to local strong errors in the signal estimate that correspond to poorly fitted peaks.
The L_2 norm has been chosen among several distances that have been tested to assess which of them would be the most discriminative for measuring the consistence between the estimated signal and the reference one. For that purpose, two reference signals have been used: a real waveform over a building roof consisting of two close echoes embedded into background system noise (Figure 4.4a), and a simulated one composed of three peaks of which two overlap significantly (Figure 4.4b). They have been compared with various correct or erroneous configurations (see Table 4.4), and their distances to the two reference signals have been measured with seven distinct measures: the L_1, L_2, L_∞ norms, the Kullback-Leibler, χ_2, Matusita, and Bhattacharrya distances. These distances are described in detail in [SB91; PHB97; CS02]. They are based on bin-to-bin distances, *i.e.*, they aim to measure the correspondence for each bin between the two signals without taking into account information for neighbouring bins. Such similarity measures are therefore sensitive to the signal length.

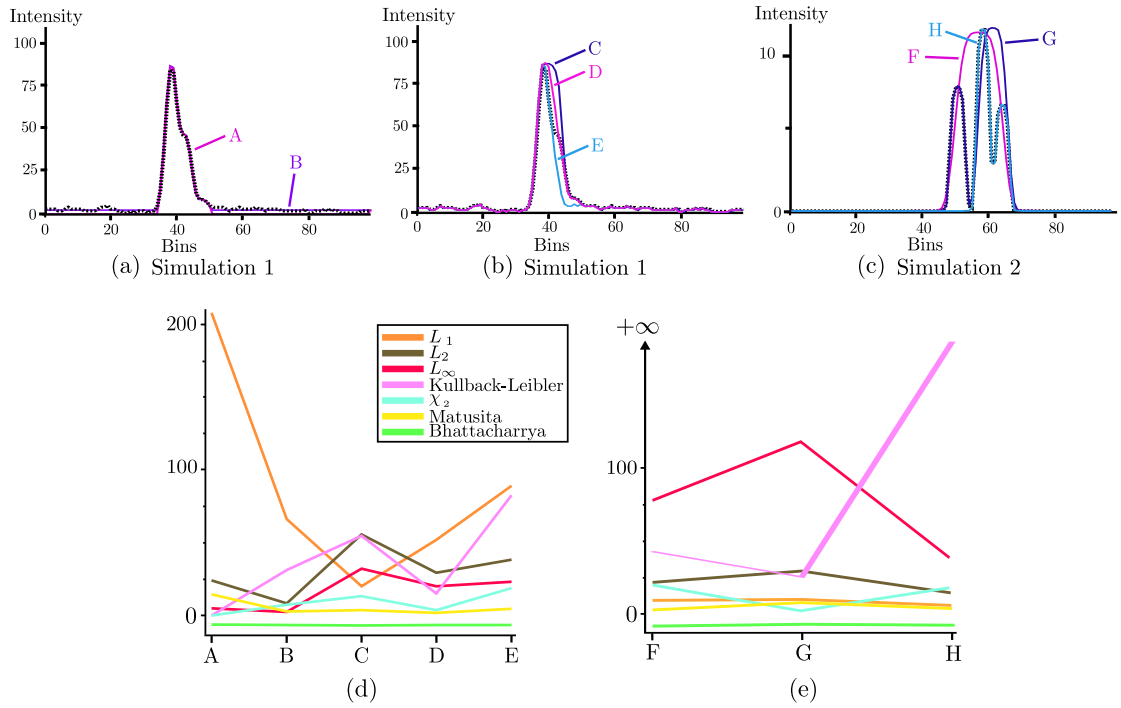


Figure 4.4: Comparison of several metrics measuring the consistence between the reference signal and a good estimated signal. **(a-b-c)** show several examples of fitted signals (coloured curves) with various qualities, compared to the reference signal (black dashed curve). **(d)** gives (respectively **(e)**) the values of the seven tested distances for simulations presented in **(a)** and **(b)** (resp. **(c)**).

Figure 4.4d shows that the L_1 norm is not adapted to real waveforms since it is greatly influenced by the background noise (configurations A and B), however it allows to clearly detect missing peaks (configurations E and H). The L_∞ norm behaves similarly and is able to enhance erroneous results (configuration G), and unfitted single echoes (configuration H), but is less effective when dealing with real waveforms that are poorly adjusted

Distance	Definition	Distance	Definition
L_p norm	$\frac{1}{\lvert K\rvert}\left(\int_K \lvert\mathcal{S}_\mathrm{x}-\mathcal{S}_\mathrm{data}\rvert^p\right)^{\frac{1}{p}}$	$\chi 2$	$\frac{1}{2\lvert K\rvert}\int_K \frac{(\mathcal{S}_\mathrm{x}-\mathcal{S}_\mathrm{data})^2}{\mathcal{S}_\mathrm{x}+\mathcal{S}_\mathrm{data}}$
L_∞ norm	$\sup_K \lvert\mathcal{S}_\mathrm{x}-\mathcal{S}_\mathrm{data}\rvert$	Matusita	$\frac{1}{\lvert K\rvert}\sqrt{\left(\int_K(\sqrt{\mathcal{S}_\mathrm{x}}-\sqrt{\mathcal{S}_\mathrm{data}})^2\right)}$
Kullback-Leibler	$\frac{1}{\lvert K\rvert}\left(\int_K(\mathcal{S}_\mathrm{data}-\mathcal{S}_\mathrm{x})\log\frac{\mathcal{S}_\mathrm{data}}{\mathcal{S}_\mathrm{x}}\right)$	Bhattacharrya	$\frac{1}{\lvert K\rvert}\left(-\log\int_K\sqrt{\mathcal{S}_\mathrm{x}\mathcal{S}_\mathrm{data}}\right)$

Table 4.3: Definitions on various distances tested as data term U_d (Figure 4.4).

(configurations D and E). Besides, the Kullback-Leibler distance has an appropriate behaviour for missing peaks. However, it is less affected by erroneous adjustments than by the background noise (configuration D compared to B). It can be seen in Figure 4.4d that the χ_2 distance has the requested behaviour, but, as each bin-to-bin measure is weighted by the mean of the two values, the final measure is levelled and is not able to penalize configurations composed of a global peak fitting two close echoes (configurations C and G). Finally, the Matusita and the Bhattacharrya distances are almost constant for all the simulated configurations since the use of the square root smoothes the differences between both signals. Thus, they cannot be used for measuring the consistence between them.

The L_2 norm has been chosen for that purpose. In practice, to reduce the effect of the background noise on the distance computation, a threshold-based removal approach is adopted. The amplitude of the receiver noise is known before from a given survey (fluctuating with ±2 DN). Then, the mean noise value is computed using all the waveform bins which amplitude is lower than the noise amplitude + 4 DN. Finally, all the waveform bins are subtracted from this mean value. It allows configurations that perfectly fit the waveforms in the "echo sections" to have a data term very close to 0.

Configuration		Specificity	Expected value
Simulation 1	A	Reference signal without noise	Low
	B	Reference signal and background noise set to a constant value	Low
	C	Signal with a single coarse peak, instead of two	High
	D	Signal with a single peak located between the first and the second one	Moderate
	E	Signal with the second echo missing	Moderate
Simulation 2	F	Signal with a single coarse peak, instead of three	Very high
	G	Signal with two echoes (one correct, and a peak locate) between the second and the third one)	High
	H	Signal with the first echo missing	Very high

Table 4.4: Description of the simulated configurations, presented in Figures 4.4a-b-c.

4.4.2 Regularization term

The term $U_{\mathrm{p}}(\mathrm{x})$ allows the introduction of interactions between objects of x and to favor/penalize some configurations.

$$U_{\mathrm{p}}(\mathrm{x}) = U_n(\mathrm{x}) + U_e(\mathrm{x}) + \sum_{\mathrm{x}_i \sim \mathrm{x}_j} U_m(\mathrm{x}_i, \mathrm{x}_j) \tag{4.4.5}$$

where $\mathrm{x}_i \sim x_j$ constitutes the set of neighbouring objects in the configuration x. This neighbourhood relationship $\sim$ is defined as follow:

$$\mathrm{x}_i \sim \mathrm{x}_j = \{(\mathrm{x}_i, \mathrm{x}_j) \in \mathrm{x} \quad | \quad |\mu_{\mathrm{x}_i} - \mu_{\mathrm{x}_j}| \leq r\} \tag{4.4.6}$$

Parameter μ_{x_i} (resp. μ_{x_j}) represents the mode (*i.e.*, the position of the maximum amplitude of the echo) of the associated function to object x_i (resp. x_j), and r is constrained by the lidar sensor range resolution (*i.e.*, the minimum distance between two objects along the laser line of sight that can be differentiated) as well as the complexity of the reconstruction we aim to achieve.
For airborne lidar waveforms the prior knowledge is set up by physical limitations in the backscatter of lidar pulses. These limitations are modelled by three terms U_n (echo number limitation), U_e (backscatter laser energy limitation), and U_m (reconstruction complexity) that are described below.

Echo number limitation

The two first echoes of a waveform contain in general most of the total reflected signal power. Consequently, even for complex targets like forested areas, a waveform empirically reaches a maximum of seven echoes and it is quite rare to find more than four echoes. In urban areas, most of the targets are rigid, opaque structures like buildings and streets. Thus, more than two echoes are usually only found in open forests. We therefore aim to favor configurations with a very limited number of reflecting objects (*i.e.*, echos) per emit pulse, with an energy given by:

$$U_n(\mathrm{x}) = -\log \mathrm{P}_{\mathrm{card(x)}} \quad \text{with} \quad \sum_{n=0}^{\infty} \mathrm{P}_n = 1 \tag{4.4.7}$$

where P_n is the probability for the waveform to have n echoes. The probabilities were empirically determined by a coarse mode estimate on the Amiens dataset. Here, we have: $P_1 = 0.6$, $P_2 = 0.27$, $P_3 = 0.1$ and $P_{4 \leqslant n \leqslant 7} = 0.01$. For $n > 7$, $U_n(\mathrm{x})$ is set to a very high positive value, which bans such configurations in practice. This defines the upper bound for the acceptable number of objects within a configuration. Since the number of echoes within a waveform is not known beforehand, this energy only slightly favors waveforms with few objects. In case of more complex signals with a significant number of pulses, this term might be discarded. Or, we could choose a probability decreasing with the number of echoes until a predefined upper bound, in order to model the expect loss of energy on the signal path.

Backscatter energy limitation

We take advantage of the law of conservation of energy and define an upper bound for the backscatter energy as well. This upper bound depends on the emitted laser energy

and the target reflectance and scattering properties. This reference energy $\mathcal{E}_{\text{ref}}$ can be set empirically to $\sqrt{2\pi}A_{\max}w_{\max}$, which is the energy of a Gaussian pulse of amplitude $A_{\max}$ and width $w_{\max}$. $A_{\max}$ and $w_{\max}$ are upper bounds for the amplitude and the width of echoes within the waveforms over the area of interest. Waveforms with larger pulse energy are penalized as follows:

$$U_e(\mathrm{x}) = \pi_e \, 1_{\{E(\mathrm{x})>\mathcal{E}_{\text{ref}}\}}(E(\mathrm{x}) - \mathcal{E}_{\text{ref}})^2 \tag{4.4.8}$$

where $1_{\{.\}}$ is the characteristic function, $E(\mathrm{x}) = \int_K \mathcal{S}_{\mathrm{x}}$ is the pulse energy of $\mathcal{S}_{\mathrm{x}}$, compared to a reference energy $\mathcal{E}_{\text{ref}}$ (see Figure 4.5), and $\pi_e \in \mathbb{R}$ is a weighting parameter. In our data sets, transmitted waveform energy is always much higher than received signal energy. Therefore, $U_e = 0$ for most of the configurations, and only penalizes configurations with unrealistic pulse shapes.

Reconstruction complexity

Our aim is twofold:

- to penalize objects spatially closer along the line of sight than the sensor range resolution;
- to favor configurations with a small number of objects, following the Minimum Description Length (MDL) principle [Ris78].

Such energy is given by:

$$U_m(\mathrm{x}_i, \mathrm{x}_j) = \pi_m \exp\left(\frac{r^2 - |\mu_{\mathrm{x}_i} - \mu_{\mathrm{x}_j}|^2}{\delta^2}\right) \tag{4.4.9}$$

where π_m, $\delta \in \mathbb{R}$ are weighting parameters. This means that a mode of a waveform may be either reconstructed by a single peak or by a sequence of peaks whose accepted minimum distance is governed by parameter r (see Figure 4.5). The lower bound of r is given by range resolution $\tau \times c/2$ (where τ is the laser pulse duration, and c the speed of light), while the upper bound of r is thus model based and may be chosen depending on the scene. For example, if we know that the data were acquired in a forested area in the leaf-off period and the trees have preferably few, but strong branches, we would chose a large r.

4.4.3 Parameter estimation

Physical and weight parameters can be distinguished in the energy. Physical parameters are r and δ. r is tuned according to sensor specifications and our knowledge on acquired waveforms. Indeed, for our datasets, τ=4 ns. It leads to r=0.75 m. δ is chosen in order to set how strong close echoes are penalized (see Figure 4.5.b). In our experiments, we set δ to 0.01. Thus, $U_m(\mathrm{x}_i, \mathrm{x}_j)$ exhibits large values when $\mu_{\mathrm{x}_i} \to \mu_{\mathrm{x}_j}$.

Data and regularization terms are weighted with respect to each other using a factor β (see Equation 4.4.3). Regularization energy terms are weighted with the two π_e and π_m parameters. Several methods allowing to estimate such kinds of parameters exist, such as the Expectation-Maximization algorithm [DLR77] or its stochastic version [CD85].

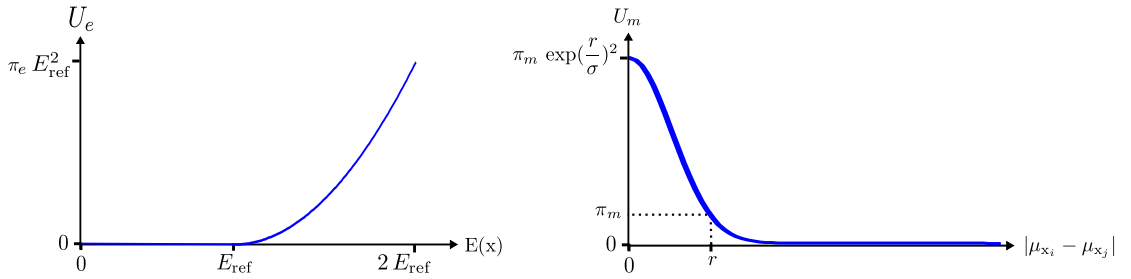


Figure 4.5: **Left**: Backscatter energy limitation term plotted against the energy of the current configuration $E(\mathrm{x})$. **Right**: Reconstruction complexity term plotted against the absolute distance between two neighbouring objects of the current configuration.

However, these algorithms are very general and are not specifically adapted to our problem. These parameters are therefore determined by "trial-and-error" tests. For instance, good results are found with β values included in the interval [0.45 - 0.6].

4.5 Optimization

We aim to find the configuration $\widehat{\mathrm{x}}$ which gives the minimum value of the energy U:

$$\widehat{\mathrm{x}} = \arg\min_{\mathrm{x}\in\mathcal{C}} U(\mathrm{x}) \tag{4.5.1}$$

This is a non-convex optimization problem in a high dimension space $\mathcal{C}$. Moreover, since the functions of the library are defined by different numbers of parameters, $\mathcal{C}$ is the union of variable dimension spaces. Two options are therefore conceivable:

- A stochastic algorithm adapted to our problem but with a significant computing time. Markov Chain Monte Carlo (MCMC) methods are very popular techniques to solve this issue [RC99].
- Other techniques that require to simplify the optimization problem. First, the state space may be narrowed: it necessitates to reduce the definition domain of the marks or adopt discrete values. Both solutions are not adequate since it may prevent the developed algorithm to find the best solution or even a suitable solution. Another possibility would be to simplify the proposed energy in order to be able to use relatively fast optimization techniques such as stochastic diffusion or jump-diffusion based approaches [SSF02; HTZ04].

The first solution has been adopted since our first implementations have shown a satisfactory computing time.

4.5.1 RJMCMC sampler

Markov Chain Monte Carlo (MCMC) sampler techniques consist in simulating a discrete Markov Chain (X_n), $\mathrm{n} \in \mathbb{N}$ on the configuration space, having an invariant measure specified by the energy U. The transitions of this chain correspond to local perturbations of

the current configuration. This means that a single object of the configuration is perturbed. This sampler performs "jumps" between spaces of different dimensions respecting the reversibility assumption of the Markov chain. They consist in two main steps: firstly, a perturbation of the current step is proposed; secondly, this perturbation is accepted or refused. One of the advantage of this iterative algorithm is that it does not depend on the initial state.

The formalism of MCMC samplers have been introduced in [MRTT53] and [Has70] (Metropolis-Hastings algorithm), and extended in [GM94] and [Gre95] (Metropolis-Hastings-Green algorithm), where sampling configuration spaces with variable dimension is now possible. Since it is required to sample from parameter spaces of varying dimensions, the Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm is well adapted to our problem. [Gre95] has proposed this extension for a model selection problem in case of a mixture of k Gaussian, with k unknown. Several papers have shown the efficiency of the RJMCMC sampler for the problem of multiple parametric object recognition [DTC04; Rip08; TP09] in image processing and computer vision.

Metropolis-Hastings-Green algorithm

In the framework of Marked Point Processes, [GM94] have designed a sampler for space state with variable dimensions. The Markov Chain $(X_n)_{n\in\mathbb{N}}$ evolves with object birth or death proposals. These two perturbations allow an efficient exploration of the configuration space, with a variable number of objects. The convergence of the chain is ensured since the two moves allow to reach any variation of the space state.
In a more general framework, [Gre95] has proposed a sampler that gathers the properties of the Metropolis-Hastings sampler (object perturbations) and the Geyer and Møller sampler (object births and deaths). Adding perturbation kernels allows to fasten the convergence process.

The jumps are realized according to these families of moves m called proposition kernels and denoted by $Q_m(\mathrm{x},.)$. Each kernel is associated with a probability p_m. The jump process performs a move from an object configuration x to y according a probability $Q_m(\mathrm{x}\rightarrow\mathrm{y})$. Then, the move is accepted with the probability $\min(1,\ R_m(\mathrm{x},\mathrm{y}))$, where $R_m(\mathrm{x},\mathrm{y})$ is called the *Green ratio*, and defined as follows:

$$R_m(\mathrm{x},\mathrm{y}) = \frac{Q_m(\mathrm{y}\rightarrow\mathrm{x})}{Q_m(\mathrm{x}\rightarrow\mathrm{y})}\ \exp -(U(\mathrm{y}) - U(\mathrm{x})) \tag{4.5.2}$$

Finally, the Metropolis-Hastings-Green algorithm is detailed in Algorithm 4.1, and the proposition kernels are presented in Section 4.5.2.

In practice, if the ratio is larger than 1, the move is forbidden. Otherwise, the acceptance is computed by comparing the Green ratio with a value κ sampled uniformly in the interval $[0,1]$.

Algorithm 4.1: Metropolis-Hastings-Green algorithm

Data: $\mathrm{x}^{(0)}$= initial configuration
$Q_m(\mathrm{x}, .)$= a set of proposition kernels
p_m= the probabilities associated to the proposition kernels.
Result: Configuration associated to the minimal energy $\widehat{\mathrm{x}}$

At **iteration t**, if $\mathrm{x}^{(t)} = x$

1. Choose a perturbation kernel $Q_m(x, .)$ with probability p_m;
2. Propose a new configuration y according to Q_m;
3. Compute Green ratio $R_m(x, y)$ according to Q_m using Equation 4.5.2;
4. Accept the proposition with a probability $\alpha = \min(1,\ R_m(x, y))$: $\mathrm{x}^{(t+1)} = y$. Otherwise, keep $\mathrm{x}^{(t+1)} = x$.

4.5.2 Proposition kernels

Two families of moves are used in order to perform jumps between the subspaces. Birth-and-death and perturbation kernels are designed to uniformly perturb the state space. Another type of move which is more specifically dedicated to the exploration of such subspaces is also proposed: the switching kernel.

- **Birth-and-death kernel** Q_{BD}: an object is added or removed from the current configuration x, following a Poisson distribution. These transformations corresponding to jumps into the spaces of higher (birth) and lower (death) dimension are theoretically sufficient to visit the whole configuration space. However, other kernels can be specified, which fit better to our problem. The aim is to speed up the process convergence by proposing relevant configurations more frequently. Therefore, two other kernels have been introduced.
- **Perturbation kernel** Q_P: the parameters of an object belonging to the current configuration x are modified according to uniform distributions.
- **Switching kernel** Q_S: the type of an object belonging to x is replaced by another type of the library. Contrary to the previous kernel, this move does not change the number of objects in the configuration. However, the number of parameters can be different (*e.g.*, four parameters for the Nakagami model are substituted by five parameters for the Burr one). This kernel creates bijections between the different types of objects [Gre95].

The impact of the four distinct kernels for a given configuration is illustrated in Figure 4.6. If an object is added, its type and its associated parameters are randomly chosen. Because no assumption can be made which move is more relevant at the current state, we choose equiprobability of the kernels in order not to favor one with respect to another. The computation of these kernels for the Green ratio is detailed in [LDG10].
Data-driven perturbations would also be possible to explore more efficiently the configuration spaces [TZ02]. Instead of using a uniform law, a Normal law can be adopted for

parameters which mean and variance have been estimated beforehand. It would be possible, for instance, for the amplitude and the width of the Generalized Gaussian model, but much more difficult for the two other functions since several combinations of their parameters can simulate almost the same realistic shapes. Therefore, data-driven perturbation kernels have not been designed.

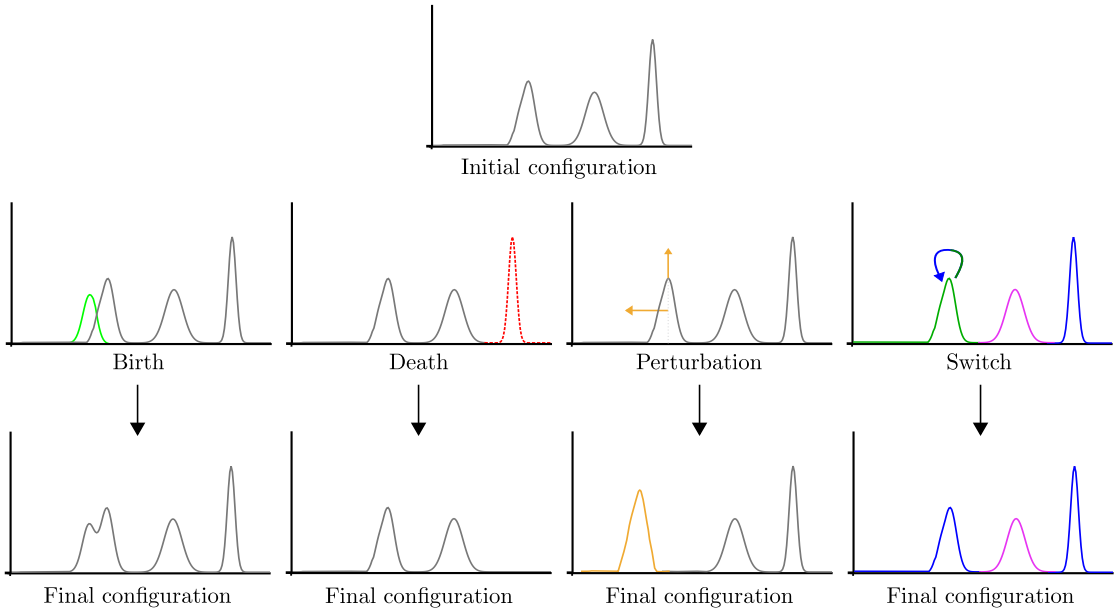


Figure 4.6: Illustration of the effect of the four adopted proposition kernels.

4.5.3 Simulated annealing

In order to converge on the desired solution, the RJMCMC sampler is coupled with a relaxation process. The simulated annealing algorithm, introduced by [MRTT53], is a stochastic global optimization procedure which exploits an analogy between the search of a minima and a statistical thermodynamic problem. It allows to reach the global maximum x^* of the probability density $h(.)$ for all initial configurations x_0. A series of densities $h_t(.)$ is generally built as follows:

$$h_t(\mathrm{x}) \propto h^{\frac{1}{T_t}}(\mathrm{x}) \tag{4.5.3}$$

The relaxation parameter T_t is defined by a sequence of temperatures decreasing to zero when t tends to infinity. $U(.)$ is thus substituted by $\frac{U(.)}{T_t}$ in the Green ratios (see Equation 4.5.2).
There are three main phases in a simulated annealing process (see Figure 4.7):

- First, the temperature is high. The density is close to the reference point process, barely biased by the energy U: most moves are accepted initially. It allows to explore non contiguous modes of the density. However, this phase is not critical as the relaxation is very fast at high temperatures. Contrary to greedy algorithms where the total energy is not globally minimized, it allows not to be stuck in its local minima.

- The second phase is the most important: the process explores the different density modes (the configurations of interest) and becomes more and more selective.
- Finally, the temperature is close to 0. It corresponds to local adjustments of the objects of the configuration. The rate of rejected propositions is very high.

The simulated annealing decrease scheme represents how the temperature T_t evolves with the time t. It should take into account both the decrease type and speed of the series of temperatures, and the estimate of the initial temperature when $t = 0$. According to [SSF02], the decrease schedule should take into account various characteristics of the density $h(.)$ (or of the associated Gibbs energy $U(.)$: the scale, the landscape, the number of local maxima, and their size. However, in practice, such information cannot be known easily. Thus, information about the energetical variations gathered during the sampling process are used instead.
During the process, a large number of proposed configurations are rejected, leading to a high computing time without any evolution. A sampler can be considered as "efficient" if the acceptance rate reaches 1/3. To deal with this problem and other theoretical issues, [DMZ09] have proposed a multiple birth-and-death process, which allows to propose a higher number of configurations for each iteration of the relaxation process.

Temperature decrease

Simulated annealing allows to reach the global maximum of the energy using a logarithmic temperature decrease. In practice, we prefer to use a geometrical cooling scheme (sometimes also referred as exponential schedule), which is faster and gives an approximate solution close to the optimal one:

$$T_t = T_0 \, \alpha^t \tag{4.5.4}$$

where α and T_0 are the decrease coefficient and the initial temperature, respectively. α can vary and be adapted to the energy variations during the process [Var96; SSF02; PDZ05]. However, this may not really fasten the process. Thus we prefer to use a constant decrease coefficient. The cooling factor is usually set to a number close to 1. In our experiments, α is set to 0.99995, which is much faster that using $\alpha = 0.99999$, or $\alpha = 0.999999999$ etc.

Initial temperature T_0

According to [Whi84], T_0 should be chosen such that the system should initially be sufficiently hot to allow large fluctuations in the energy. This initialization is very important. If T_0 is too high, the first phase of the process will be too long and useless, and the computing time will increase. If T_0 is too low, the simulated annealing will directly start in the selection phase (step 2), which will be shortened (and bound to be too short). T_0 can be therefore estimated in function of the energy variations of a sample of random configurations. The scale of these fluctuations is set by the standard deviation σ_U of the energy computed at infinite temperature. In practice, T_0 is chosen twice as large as σ_U:

$$T_0 = 2\,\sigma_{U_{T=\infty}} = 2\,\sqrt{\langle U^2_{T=\infty}\rangle - \langle U_{T=\infty}\rangle^2} \tag{4.5.5}$$

where $\langle U \rangle$ is the empirical mean of the energy computed from the sample of examples. In practice, for a robust estimation, several hundreds or thousands of configurations are

necessary. However, the time required to estimate the initial temperature is negligible, compared to the simulated annealing process time. The sample of random configurations is generated as follows:

Algorithm 4.2: Automatic set up of the initial temperature in simulated annealing.

Data: n_C= number of configurations to be simulated
m= parameter of the Poisson law
Result: Initial temperature T_0

for *i=1* **to** n_C **do**

1. Compute the number of objects n_o, according to a Poisson law with parameter m.
2. Simulate a configuration with n_o objects. The object parameters are set up randomly using a Uniform distribution.
3. Compute the associated energy U_i

$T_0 = 2\,\sigma(U_{T=\infty})$

Stop temperature T_{final}

[Whi84] suggests taking the stop temperature of the order of the smallest energy scale in the system. It corresponds to the smallest energy variation of a single move. Another more common and simpler solution consists in stopping the process if the energy does not evolve in the last N_{stop} iterations. This solution has been adopted here with $N_{\text{stop}} = 1000$.

Practical aspects

Algorithm 4.2 has been applied to find T_0. Parameter m is the expected value of the Poisson law, which is here the expected mean number of objects. We set m=2 and n_C=10000. This leads to T_0 values superior to 20 millions. This is mainly due to U_e and U_m terms in the proposed energy. Configurations that are randomly proposed may contain almost superposed objects with significant backscatter energy. This leads to very high energy values (setting m=2 only prevents configurations to have a high number of peaks, which concerns the last prior term U_n). Such T_0 values are, in practice, too high, and too much time is spent in the highest temperatures, which slows down the process. Consequently, T_0 has been set up manually by an analysis of the energy behaviour during the simulated annealing process. Thus, T_0=10.

In our experiments, the optimization process requires between 100,000 and 150,000 iterations. Moreover, T_{final} is close to 0.00005. The process lasts approximately 0.5s/waveform. Thus, processing millions of waveforms requires a rather significant computing time. With a 8-core 2.93 GHz Mac Pro with 6 GB RAM, approximately 50,000 waveforms can be processed per hour. In comparison, the Generalized Gaussian decomposition presented in Chapter 3 can process 2.5M signals in the same time lapse (1:50 ratio).

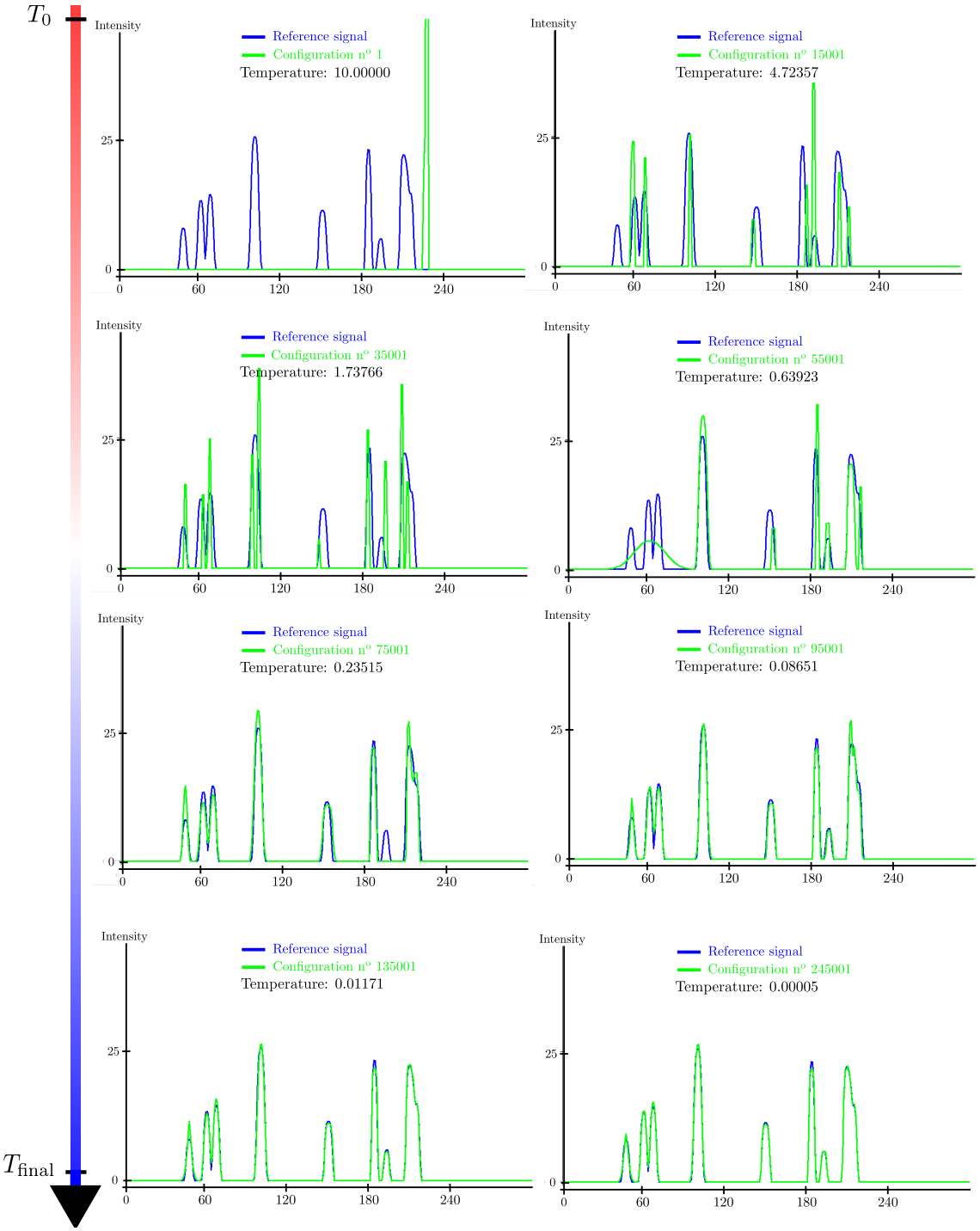


Figure 4.7: Optimization process: evolution of the object configuration from the initial temperature T_0 to the final one T_{final} (left to right, and top to down).

4.6 Results

The proposed approach has been first assessed on simulated data and subsequently applied to different kinds of spatial and airborne lidar signals. The results have been evaluated quantitatively by computing the normalized cross-correlation coefficient ρ and the relative Kolmogorov-Smirnov distance KS between the raw and the estimated signals. We have:

$$\rho = \frac{\sum_{i=1}^{N}(\mathcal{S}_{\text{data}}(i) - \overline{\mathcal{S}_{\text{data}}}) \cdot (\hat{\mathcal{S}}(i) - \overline{\hat{\mathcal{S}}})}{\sqrt{\sum_{i=1}^{N}((\mathcal{S}_{\text{data}}(i) - \overline{\mathcal{S}_{\text{data}}})^2}\sqrt{\sum_{i=1}^{N}(\hat{\mathcal{S}}(i) - \overline{\hat{\mathcal{S}}})^2}} \quad \in [-1,1] \tag{4.6.1}$$

$\mathcal{S}_{\text{data}}$ is the reference waveform, and $\hat{\mathcal{S}}$ is our estimated signal, both composed of N bins. $\overline{\mathcal{S}_{\text{data}}}$ and $\overline{\hat{\mathcal{S}}}$ are their respective mean values. If the reconstructed signal perfectly fits with the lidar waveforms bins, ρ=1. The correlation coefficient is rather sensitive to outliers. KS is a normalized L_∞ norm, both used to detect missing echoes and local shifts between signals. It is defined as follows:

$$\text{KS}(\mathcal{S}_{\text{data}}, \hat{\mathcal{S}}) = \frac{\sup_K |\mathcal{S}_{\text{data}} - \hat{\mathcal{S}}|}{\max_K \mathcal{S}_{\text{data}}} \quad \in [0,1] \tag{4.6.2}$$

The L_∞ norm has been normalized to allow comparisons between waveform fitting results from different sensors. KS=0 means that every lidar bin perfectly matches with the reconstructed signal, whereas KS=1 means that the main echo has been missed. Setting a satisfactory KS upper bound thus mainly depends on the noise level of the lidar waveforms. The standard L_2 norm, already used in Chapter 3 to evaluate the fitting performance of the Generalized Gaussian decomposition, would have been another solution. However, such distance has already been introduced in the algorithm: this is the data term of our energy. Therefore, selecting a fitting quality measure for evaluating an algorithm which minimizes the optimized metric is extremely biased. The fitting process and the evaluation criterion must be independent.

4.6.1 Simulated data

Relevance of the optimized energy

Various simulations have been carried out to assess the relevance and the effectiveness of the proposed energy model. Figure 4.8 shows several reconstructions of a simulated signal composed of three pulses with two overlapping peaks with variations on the optimized energy. The simulated signal and the estimated one are represented by the dotted black line and the continuous grey line, respectively.
First, the data term U_d has been considered only (*i.e.*, we only minimize the difference according to Equation 4.4.4). It can be noticed in Figure 4.8a that the signal is correctly estimated, but the configuration is composed of a high number of echoes (eleven). Thus, the result is not realistic since not all echoes do represent a real target. Then, only the regularization term is considered (U_p, see Equation 4.4.5). Figure 4.8b shows that the proposed regularization energy constraint is useful since it provides a realistic lidar waveform:

one echo with a bounded energy. This is due to both the echo number limitation and the backscatter energy limitation terms. Finally, Figures 4.8c to 4.8f show the influence of the reconstruction complexity term U_m. It is first discarded on Figure 4.8c: the signal is perfectly reconstructed, but with the maximum number of echoes allowed by U_n. It does not correspond to reality since the echoes are too closely located to each other. Then, U_m is introduced and r is respectively set to 0.3, 0.75, and 3 m in Figures 4.8d, 4.8e, and 4.8f. The greater the radius r, the lower the number of peaks. It can be noticed that a reasonable value of r allows the reconstruction of the signal with the appropriate number of echoes (Figure 4.8e), whereas larger values lead to erroneous detections (Figure 4.8f).

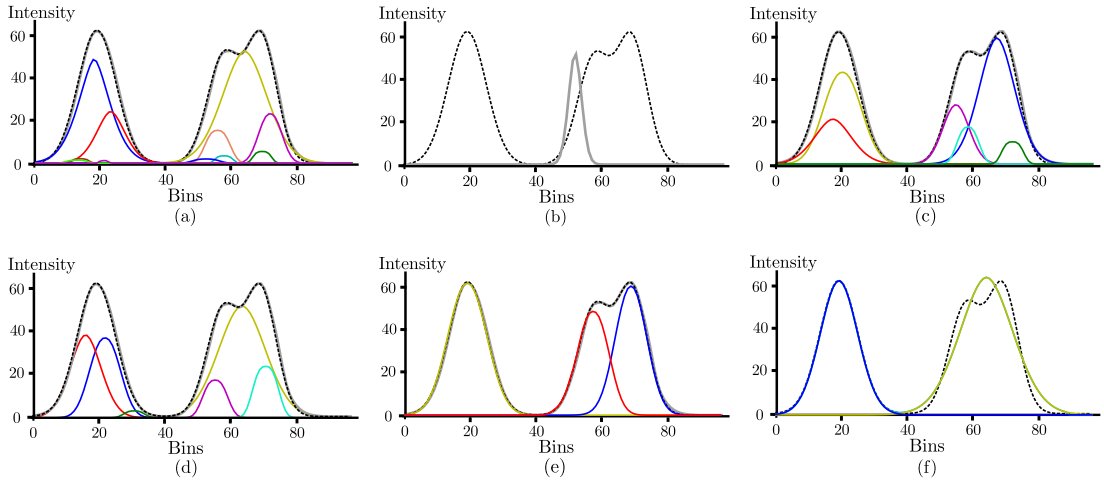


Figure 4.8: Various signal reconstructions with variations of the model energy. The dotted black line and the continuous grey one are respectively the raw and the estimated signals. The other colours correspond to the individual echoes that compose the estimated waveform. (a) No regularization term. (b) No data term. Both data and regularization terms are now considered. The reconstruction complexity term U_m is disabled in (c) (π_m=0). Then, U_m is introduced with increasing parameter values r: r is set to 0.3, 0.75, and 3 m respectively for (d), (e), and (f).

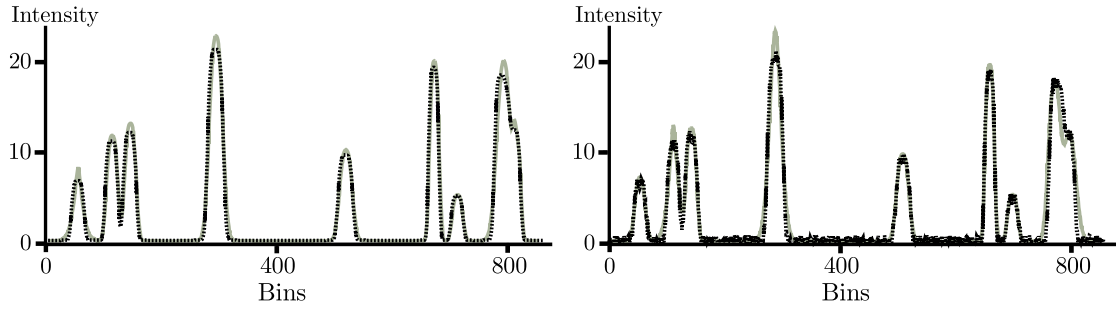


Figure 4.9: Results of complex waveform fitting on a simulated signal with nine echoes (**left**) and on the same signal with Gaussian noise (**right**). The dotted black line and the continuous grey one are respectively the raw and the reconstructed signals.

Target counting assessment

The algorithm has been first applied on signals with a higher complexity than real lidar waveforms. Longer signals with more echoes than physically expected have been fitted with our method, with distorted and overlapping modes as well as corrupted with noise. The interaction between objects can be changed by decreasing and increasing r and σ. To reconstruct signals with higher energy, E_{ref} can be tuned. To fit signals with more modes, the echo number limitation can be modified by accepting more echoes within the signal and with the same probability.
Figure 4.9 shows that good fitting results can be achieved on simulated waveforms, even corrupted with Gaussian noise, by tuning prior parameters. The nine echo locations are accurately found. However, small differences between the reference and the estimated signals can be locally noticed, especially with noisy signals where the algorithm has more difficulties to find the exact maxima and fit the upper parts of the modes (*e.g.*, 2$^{\text{nd}}$ and 4$^{\text{th}}$ echoes in Figure 4.9b).

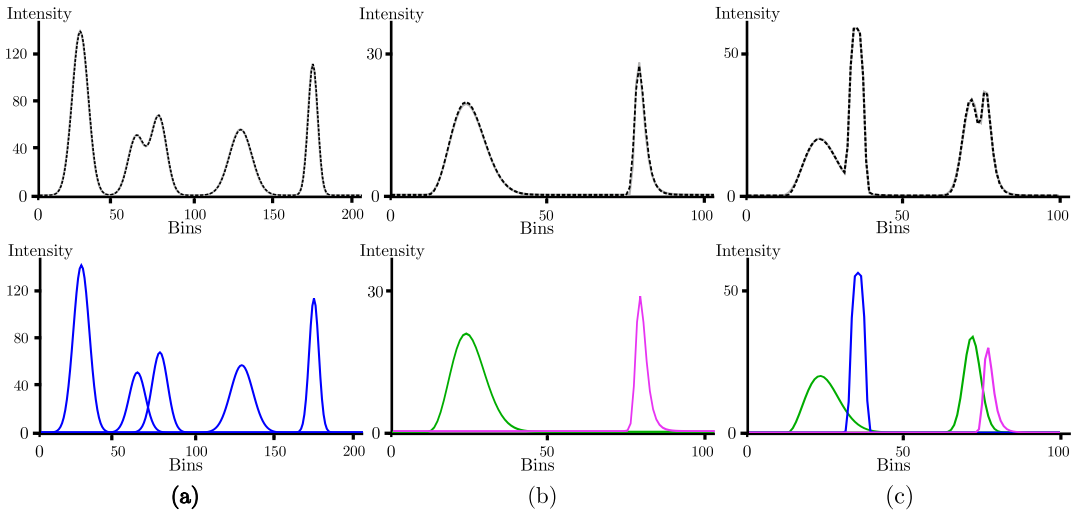


Figure 4.10: **Top**: simulated waveforms – **Bottom**: estimated waveforms and decomposition of the mixture (the colour indicates the type of the object, see Section III-A). **(a)** Five Gaussian pulses. **(b)** Two skewed pulses. **(c)** Two groups of two overlapping pulses with one Gaussian and one skewed pulse.

Parameter estimation accuracy and model selection assessment

Finding the right number of relevant peaks allows to count the appropriate number of targets. However, it is not sufficient for the generation and the subsequent classification of point clouds.
First, the parameters of each function should be accurately estimated. This is particularly true for the position of each peak (or the parameters of the function that allows to retrieve the peak mode) since such information is used to georeference the peak. For waveforms sampled at 1 ns, an error of half-a-sample in the peak location estimate leads to errors of ± 7 cm and ± 1 cm in altimetry and planimetry, respectively, for the derived 3D point. Besides, the correct analytical function must be retrieved for each peak. That is the genuine aim of the proposed method and crucial to perform 3D point cloud classification

afterwards.
Three experiments have been carried out with three simulated signals to assess the performance of the approach for these two issues (Figure 4.10). First, a signal composed of five Gaussians with overlapping peaks has been simulated, and estimated with our approach using only the Generalized Gaussian function with the α parameter set to $\sqrt{2}$ (*i.e.*, a Gaussian function, see Figure 4.10a). Table 4.5 shows the parameters are accurately estimated even for the two overlapping peaks. The peak modes are retrieved with an error below 1/10 of a bin which is satisfying for any kind of sensor. Besides, a signal with two single skewed peaks (one Nakagami and one Burr, see Figure 4.10b) has been simulated and estimated using the full library. One can see in Table 4.6 that the correct functions have been selected, and the parameters correctly retrieved with few errors in the mode location estimate (1/4 bin for the Burr pulse). Finally, two GG peaks are added to the waveform previously simulated to assess the impact of having overlapping pulses on the parameter estimates (Figure 4.10c). The functions are correctly selected, except for the third peak (see Table 4.7). However, the estimated parameters of the Nakagami pulse lead to a symmetric peak, similar to the simulated GG one. Moreover, the parameters of the two first peaks are correctly estimated but with an accuracy inferior to single peaks. The closeness of two peaks has therefore an impact on the estimation that cannot be neglected. The fourth peak is well fitted with the Burr model. This is due to the fact that several parameter combinations can lead to the same shape for this model, with the same mode location.

Parameter / Echo		I	μ	w
# 1	simulated / estimated	125 / 124.988	25 / 25.0006	5 / 5.001
# 2	simulated / estimated	45 /45.11	60 /59.92	5.3 / 5.23
# 3	simulated / estimated	60 / 59.988	74 / 73.962	4.8 / 4.845
# 4	simulated / estimated	50 / 50.32	125 / 124.985	7 / 6.969
# 5	simulated / estimated	100 / 99.82	170 / 170.005	3 / 3.0007

Table 4.5: Comparison between a simulated waveform composed of five Gaussians and the estimated signal (see Figure 4.10a).

Parameter / Echo		I	μ	ξ/a	$\omega/(b,c)$		mode
1	simulated (Nakagami)	230	10	15	2		22.99
	estimated (Nakagami)	230.67	10.198	14.797	1.899		22.89
2	simulated (Burr)	100	65	8.5	8.6	16.4	76.354
	estimated (Burr)	99.71	64.455	8.651	8.724	16.824	76.608

Table 4.6: Comparison between a simulated waveform composed of two asymmetric peaks and the estimated signal (see Figure 4.10b).

4.6.2 Medium and large footprint waveforms

The proposed approach and generally the content of this work are not dedicated to medium and large footprint sensors. We have however found relevant to assess the versatility of

Echo	Parameter	I	μ	$\xi/w/a$	$\omega/\alpha/(b,c)$		mode
1	simulated (Nakagami)	230	10	15	2		22.99
	estimated (Nakagami)	230.55	11.419	12.95	1.42		22.84
2	simulated (Generalized Gaussian)	50	35	5	2		35
	estimated (Generalized Gaussian)	49.903	35.001	4.82	1.991		35.001
3	simulated (Generalized Gaussian)	30	72	2.8	$\sqrt{2}$		72
	estimated (Nakagami)	201.65	52.62	19.48	13.11		71.72
4	simulated (Burr)	100	65	8.6	8.7	16.8	76.73
	estimated (Burr)	105.84	49.766	24.404	19.819	7.6809	76.73

Table 4.7: Comparison between a simulated waveform composed of a mixture of Gaussian and non-Gaussian pulses and the estimated signal (see Figure 4.10c).

our methodology. This is all the more interesting as the resulting waveforms can exhibit strong asymmetric behaviour due to the spread of the laser beam.

Waveforms from LVIS (Laser Vegetation Imaging Sensor) and SLICER (Scanning Lidar Imager of Canopies by Echo Recovery) NASA sensors have been decomposed and modeled with our approach (Figure 4.11). The sensor goals and specifications are described in Section 2.3. LVIS waveforms have been acquired in March 1998 over a 800 km^2 area of Costa Rica using 25 m-diameter footprints[3] [HRBD02]. Both fine and coarse fitting strategies have been tested. The fine strategy consists in selecting r such that each mode of the waveform will be fitted by a function ($r = 3$ m). It leads to almost perfect signal approximation, but conclusions are difficult to draw since the function selected for a given peak depends on the functions of the neighbouring echoes (Figure 4.11a). With the coarse solution, r is set to higher value (9 m) and σ to 0.001. This reduces the complexity of the reconstruction and therefore prevents overlapping or close echoes from being separated. A unique global peak is selected instead (Figure 4.11b), providing a general trend for the first part of the signal (in practice, the first tree canopy layer). SLICER elevation profiles come from in the BOREAS Northern Study Area in Canada[4], and have been acquired in July 1996 [Har00]. Table 4.8 shows that signals from both sensors are correctly decomposed, without significant errors. However, the KS distance values show that some small peaks are not retrieved. Indeed, LVIS and SLICER elevation profiles are very complex since the sensor laser beam integrates many distinct objects. Thus, even with the fine strategy, several close narrow pulses (as displayed on Figure 4.11b) cannot be all detected.

With medium and large-footprint waveforms, the Generalized Gaussian model is not preferred by the algorithm. The two functions allowing to simulate asymmetric peaks are preferred. The main noticeable results (see also Table 4.8) are that:

- the GG function is rarely chosen, mainly for peaks with a small amplitude. Thus, the lidar echo Gaussian assumption is no longer valid. This fact underlines the relevance of our approach for modeling lidar waveforms with a library of functions.

[3]Data set available at https://lvis.gsfc.nasa.gov/index.php
[4]Data set available at http://core2.gsfc.nasa.gov/research/laser/slicer/browser.html

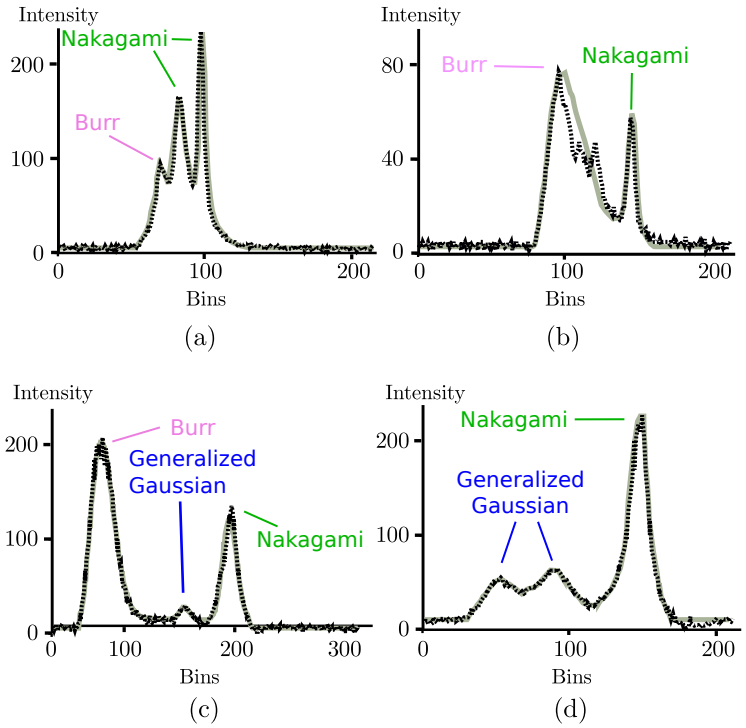


Figure 4.11: Examples of fitted **(a-b)** LVIS and **(c-d)** SLICER waveforms. Waveform **(b)** has been fitted setting r to a high value, to prevent small overlapping echoes to be individually detected.

Sensor (# waveforms)	ρ	KS	GG (%)	Nakagami (%)	Burr (%)
SLICER (76417)	0.949	0.11	6.5	51.2	41.8
LVIS (4001)	0.968	0.14	5.1	57.0	37.9

Table 4.8: Medium and large-footprint waveform fitting and modeling statistics. The fine solution has been adopted for the signal decomposition. The two first columns (ρ; KS) provide quality measures. The three last columns indicate the percentage of echoes that have been fitted with each of the three modeling functions.

- the Nakagami model is preferred to the Burr function, since its parameters allow for a higher flexibility. It is mainly selected for the last echo, which correspond to the ground and low above-ground objects, and is usually left-skewed.
- the Burr function is relevant for echoes that correspond to pulses backscattered from the tree canopy (first layer of the vegetation).

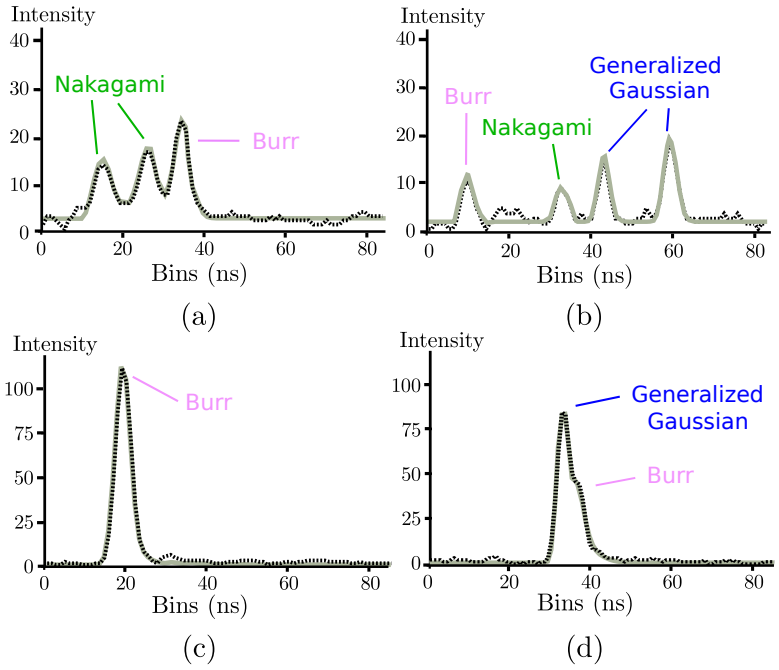


Figure 4.12: Decomposed and modelled waveforms on **(a-b)** trees, **(c)** a building roof, and **(d)** a hedge (Riegl LMS-Q560 sensor).

4.6.3 Small footprint waveforms

Both Riegl and Optech sensor waveforms have been processed with the proposed approach. For the RIEGL LMS-Q560 device over Biberach, some simple 1D results are presented in Figure 4.12. First, it can be noticed in Figure 4.12 that the algorithm performs well on complex waveforms. The correct number of echoes is found as well as the correct shape of the waveform: single and multiple overlapping echoes are retrieved, even in vegetated areas where the noise level is significant with respect to the echo amplitudes (Figures 4.12a and b). Moreover, for opaque solid targets like building roofs and ground, slightly asymmetric echoes are retrieved, and correctly adjusted: the Burr model allows to adjust them, especially when dealing with the second echo of two overlapping ones (Figures 4.12c and d).

Furthermore, approximately 6M waveforms acquired over Amiens for various kinds of (urban) landscapes have been fitted. First, six small patches have been selected, and more than 123,000 waveforms have been analyzed. The aim is to assess the reliability of the method in heterogeneous landscapes and to show its local stability in homogeneous areas.

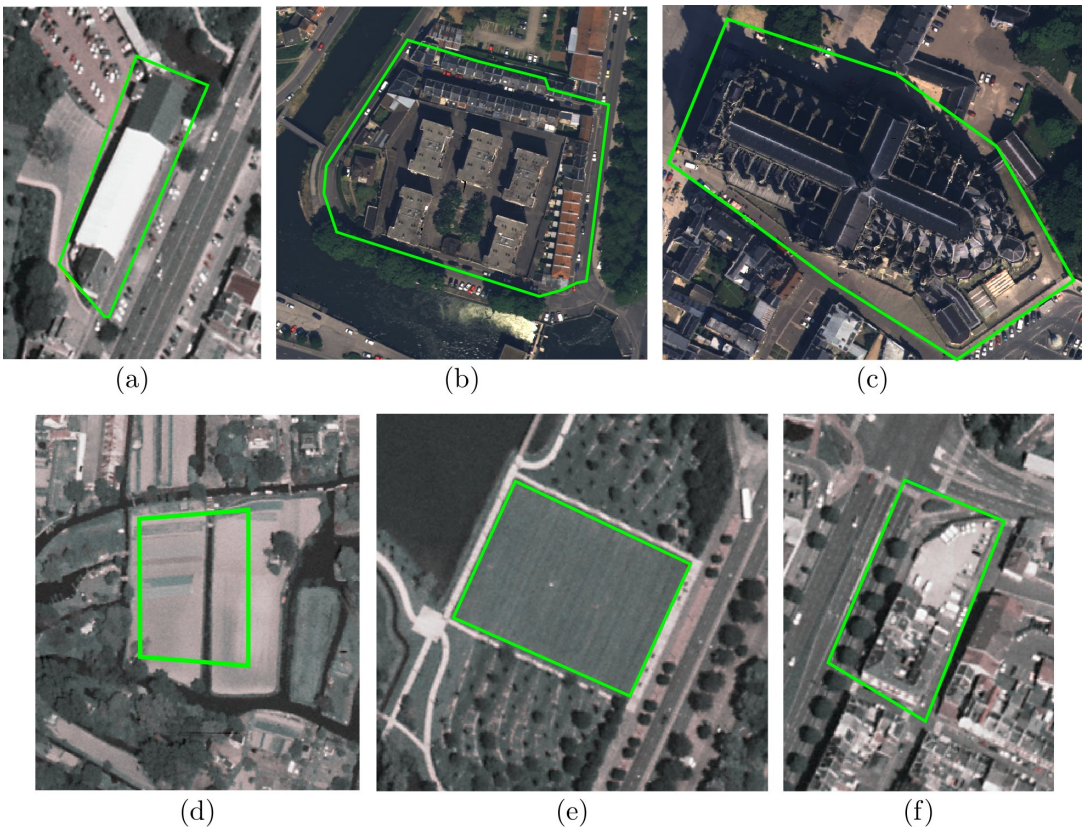
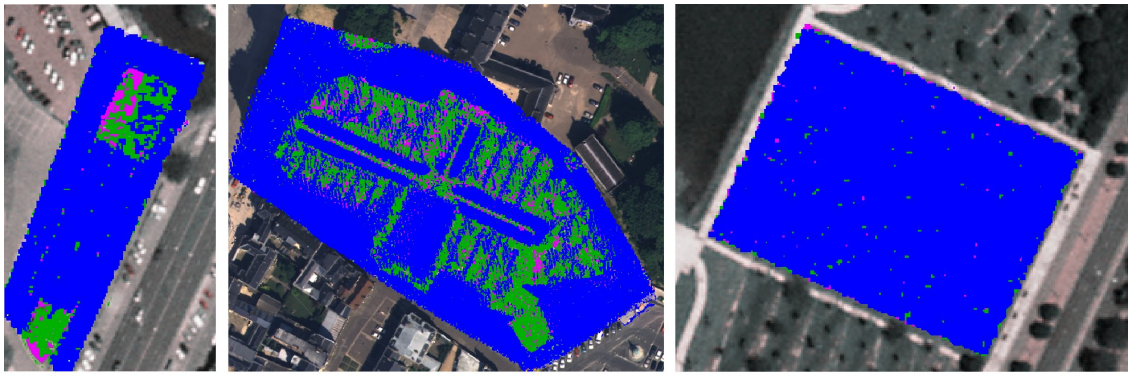


Figure 4.13: Orthoimages of the six areas of interest presented in Table 4.9 ©IGN (same order).

The six regions of interest are (see Figure 4.13): three simple aligned buildings with different slopes and materials (*Building #1* (a)); a complex area with high and low buildings with grass and trees (*Building #2* (b)); a Gothic cathedral (*Cathedral* (c)); a flat harvested field (*Field* (d)); a slightly sloped grass surface (*Grass* (e)); and a mixed set of buildings with a street, pavement, and trees (*Street* (f)). Furthermore, the echoes detected by the lidar system during the acquisition survey are provided (MP echoes). To assess the relevance of waveform processing and to solve a multiple mixture problem, the waveforms have been fitted with our proposed approach using the library of functions as well as based exclusively on the Gaussian model. Quality measures (ρ and KS) are provided for both. The percentages of echoes that have been fitted by each of the three modeling functions are indicated, as well as the percentage of echoes additionally retrieved, compared to the unknown hardware detection method. All the results are included in Table 4.9, and three model maps are presented in Figure 4.14.



(a) (b) (c)

Figure 4.14: 3D point clouds interpolated in 2D, and coloured with the selected function: Generalized Gaussian – Nakagami – Burr. **Left**: Building #1 – **Middle**: Cathedral – **Right**: Grass.

On a second time, larger areas have been processed to assess whether local scale behaviour is coherent with large scale results. We focus on two tiles composed 500,000 waveforms, namely Areas #1 and #2, presented in Figure 3.12. The generated 3D point clouds, coloured with the selected model, are displayed in Figure 4.16. Besides, Table 4.10 provide statistics about the fitting process: the distribution of the three models, the number of point that are retrieved, and the fitting quality both for the full library and the single Generalized Gaussian function. Since such results are averaged over the number of processed waveforms, further statistics are provided in Figures 4.15 and 4.17. Histograms for the two fitting quality measures ρ and KS are given in Figure 4.15. The distributions of the selected functions for each echo number are eventually displayed in Figure 4.17.

The main conclusions are:

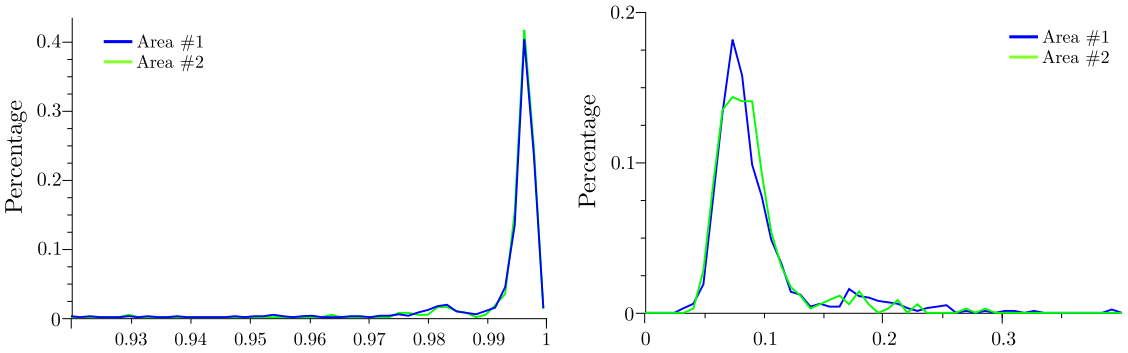
- Whatever the region of interest, the fitting accuracy is high (ρ >0.99 and KS<0.1) with our approach. One can notice that using a library of shapes slightly improved the fitting accuracy compared to only the Gaussian model or the Generalized Gaussian model. We have gained a higher flexibility in the fitting process with new models featuring more and distinct parameters. The fitting accuracy is higher than for

Area		Building #1	Building #2	Cathedral	Field	Grass	Street
# waveforms		9,943	38,565	43,563	10,035	9,790	11,770
# MP echoes		10,555	40,785	49161	10,035	9,790	12,428
Library	# echoes	11,054	43,385	50,638	10,035	9,790	13,033
	GG	81.2	60.3	62.5	91.2	99.5	61.7
	Nakagami	12.6	35.3	27.4	4.4	0.35	31.4
	Burr	6.2	4.4	10.1	4.2	0.15	6.9
	ρ	0.9947	0.9948	0.9948	0.997	0.999	0.994
	KS	0.098	0.0977	0.095	0.038	0.025	0.102
Gauss.	ρ	0.991	0.987	0.9824	0.992	0.995	0.981
	KS	0.109	0.125	0.113	0.087	0.057	0.134
Additional points (%)		+4.7	+6.4	+3	0	+0.5	+4.8

Table 4.9: Fitting results on six urban regions of interest of Amiens (see Figure 4.13). Our approach has been tested using both the full library of models and the single Gaussian function.

Area		Area #1	Area #2
# waveforms		500,001	500,001
# MP echoes		508,252	491,973
Library	# echoes	526,701	520,129
	GG	79.7	86.5
	Nakagami	12.1	8.9
	Burr	8.2	4.6
	ρ	0.9939	0.9927
	KS	0.092	0.095
GG	# echoes	526,996	520,533
	ρ	0.991	0.989
	KS	0.111	0.117
Additional points – MP (%)		+3.6	+5.7
Additional points – GG (%)		-0.06	-0.08

Table 4.10: Fitting results on Areas #1 and #2 (Optech 3100EA sensor, see Figure 4.13). Our approach has been tested using both the full library of models and the single Generalized Gaussian (GG) function. Quality measures (ρ and KS) are provided for both. The percentages of echoes that have been fitted by each of the three modeling functions are indicated, as well as the percentage of echoes additionally retrieved, compared to the unknown hardware detection method (MP) and to the method presented in Chapter 3 (GG).



(a) ρ. (b) KS.

Figure 4.15: Histograms of fitting quality for the two areas of interest presented in Table 4.10.

medium and large-footprint waveforms. However, the latter ones are much more complex.

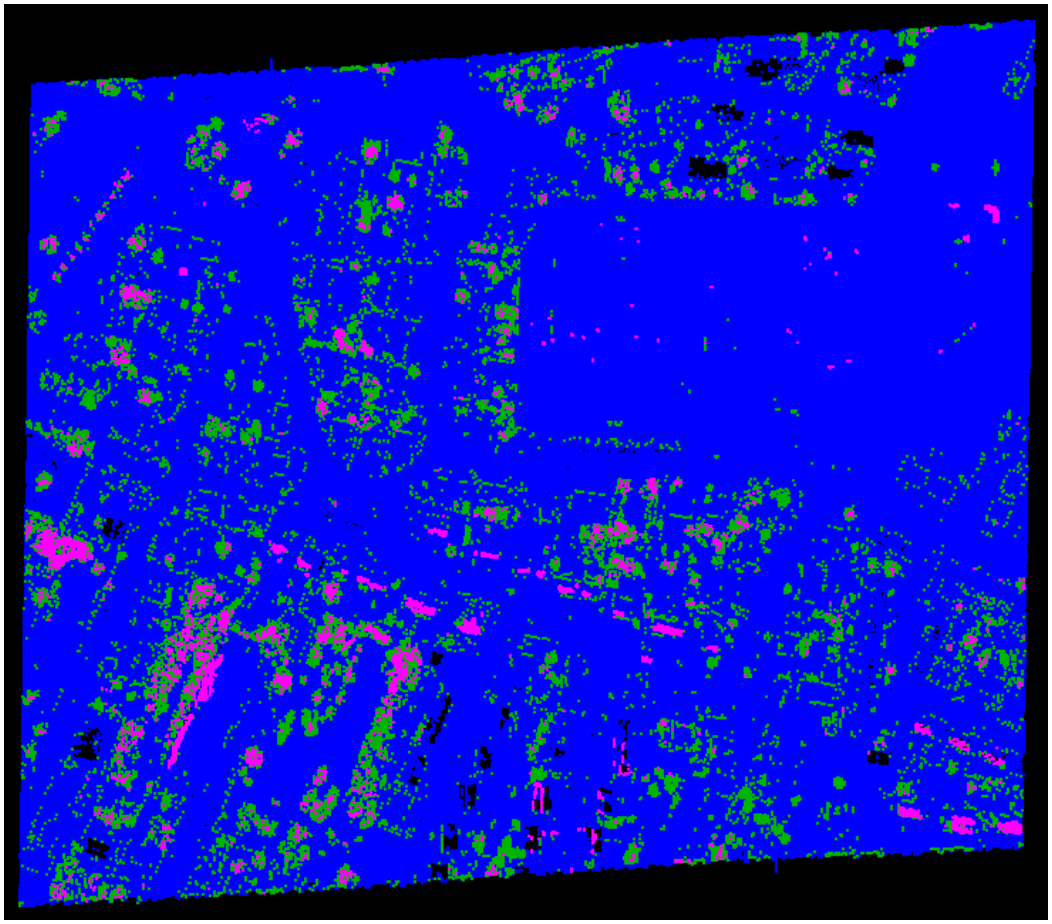
- The echo detection rate is still improved compared to the 3D point cloud provided by the lidar device. The figures are similar to those obtained in Chapter 3, using the Generalized Gaussian decomposition approach. This is an assessment of the effectiveness of our approach at large scales. Thanks to the introduction of basic physical knowledge, there is no overfitting, and the method does not generate false echoes anywhere. Such an issue was really at stake since the fitting procedure does not rely at all on any peak detection step. However, one can notice in Table 4.10 that compared to the latter method, the new method leads to a slightly inferior number of points. This may come from asymmetric pulses. On the one hand, they are fitted with one skewed peak. On the other hand, they are adjusted with two symmetric peaks, similarly to Figure 4.1.a. We do not consider they are conflicting results, but rather as two distinct fitting strategies.

- For flat areas, the echoes are symmetric and the Generalized Gaussian function is selected (*Field* and *Grass* areas, see Figure 4.14c, and streets in Figure 4.16). However, in some cases also the Burr and Nakagami functions have been selected because for some parameter set-ups they are very similar to Gaussian distributions.

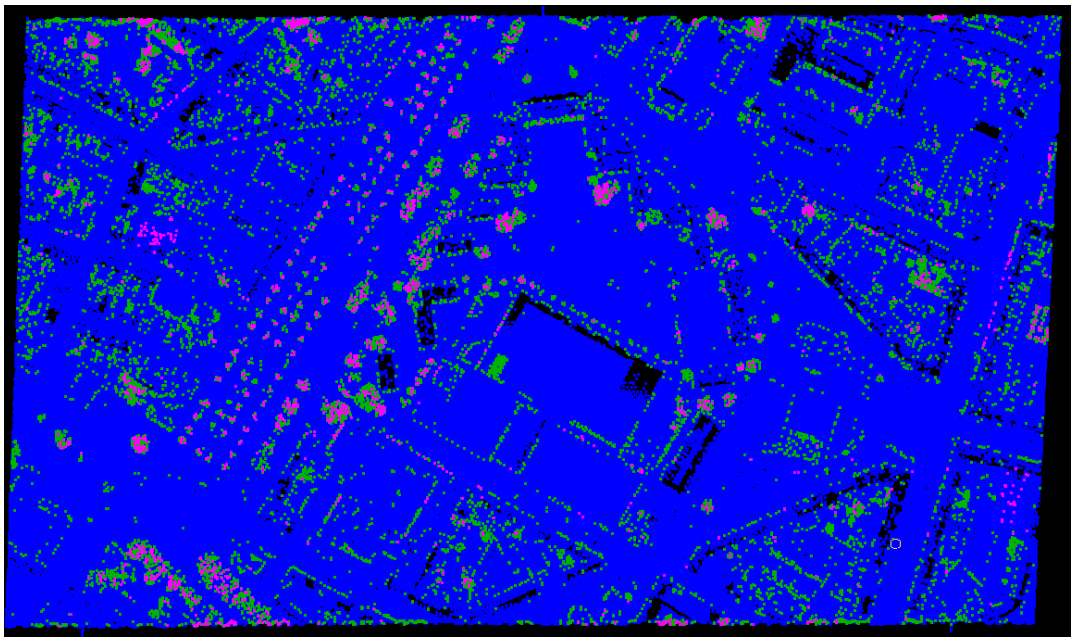
- In vegetated areas (trees), the algorithm does not preferably select a particular model (see Figure 4.18). The usefulness of asymmetric modelling functions is therefore difficult to assess for fitting echoes of small-footprint waveforms in forested areas, and the Gaussian function should be sufficient.

- For building regions, both symmetric and skewed peaks are retrieved. Asymmetric echoes can be found on building roofs and where surface discontinuities exist (buildings edges clearly appear on Figure 4.16.b). Such behaviour is also frequently observed for the Cathedral scene depicted in Figure 4.14b. When the target geometry becomes complex, the Nakagami and Burr functions are preferred.

- The reflectance of the targets also has an influence on the fitting algorithm: for high reflectance objects, the backscattered pulse has a significant amplitude and becomes narrower. In such cases, the Gaussian model is selected, as displayed in Figure 4.14a



(a) Area #1.



(b) Area #2.

Figure 4.16: 3D point cloud coloured with respect to the selected model (top view): Generalized Gaussian – Nakagami – Burr. The areas of interest are presented in Figure 3.12.

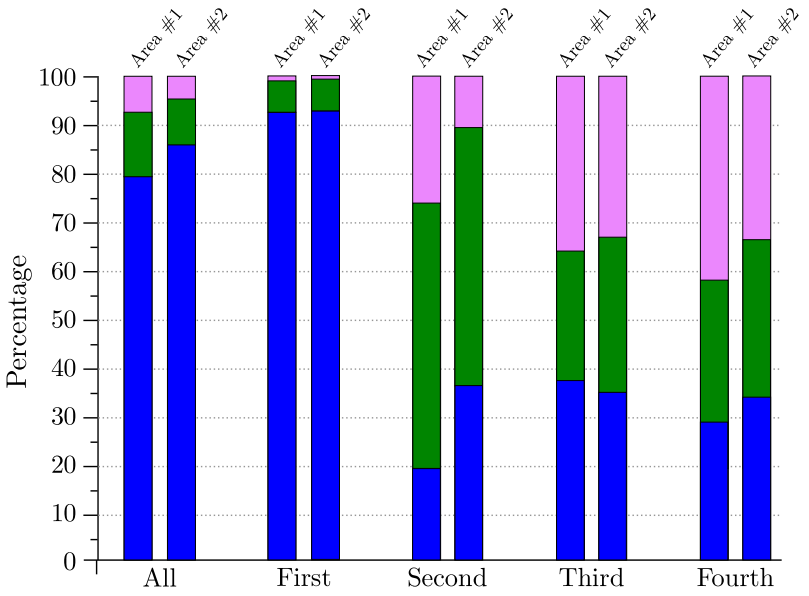


Figure 4.17: Distribution of the three modelling functions (Generalized Gaussian – Nakagami – Burr) for the two areas of interest according to the echo number. We have limited the plot to the fourth echo since there were not enough fifth and sixth echoes (< 15) for interesting statistics.

for the *Building #1* area. For roofs with lower reflectance, the peaks are smoother and the algorithm may favor slightly asymmetric peaks (*e.g.*, in Figure 4.16.a, where pink aligned patches correspond to small roofs with waveforms adjusted with the Burr model).

- Eventually, as enhanced in Tables 4.9 and 4.10, and in Figure 4.17, there are indeed asymmetric peaks, but in a relative low proportion. This is particularly true for first echoes, whereas for third and fourth echoes, there are no privileged solutions. This surely comes from the fact that such echoes have been influenced by the targets previously hit, and have relatively low amplitude. When dealing with signals with low SNR, the model selection step may not be straightforward since no real privileged solution exist.

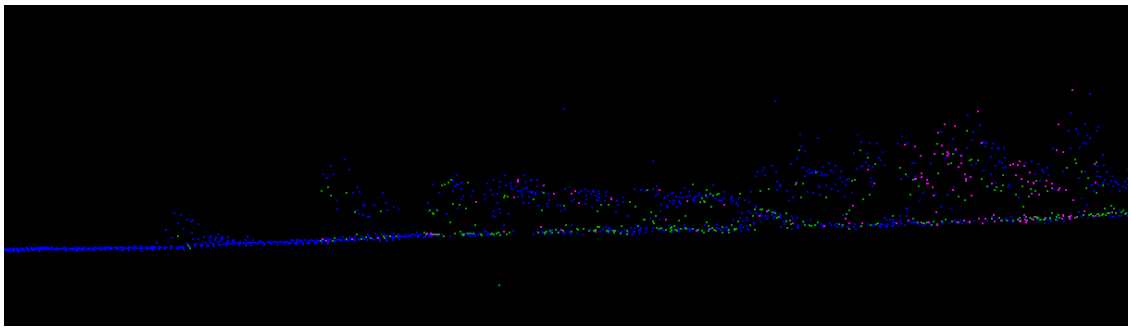


Figure 4.18: 3D profile of a vegetated place in Area #1 (Colour=selected model).

4.7 Conclusions and perspectives

We have proposed in this chapter a new and original methodology to process lidar waveforms, using a marked point process. Such a stochastic approach is well adapted to both locate echoes in signals, and accurately describe them with parametric functions taken from an extensible and tunable model library. Therefore, this allows to model lidar waveforms by complex parametric functions which results in a better fitting accuracy, compared to conventional Gaussian waveform fitting schemes. The signal is therefore better reconstructed, which also means that less loss in data compression. However this aspect has not been investigated, it seems to be a particularly interesting advantage of our method, due to the significant amount of data end-users have to deal with.
The algorithm has been successfully applied to waveforms from different lidar sensors, and at different spatial scales, showing its effectiveness and flexibility for various landscapes and resolutions. For that purpose, given a set of functions, one just has to change the maximal number of echoes that is accepted as well as the parameter r tuning the reconstruction complexity. The MDL term of the regularization energy is all the more interesting as it offers the possibility to free the waveform processing algorithm from the peak detection step, specific to each sensor. Indeed, the knowledge of the range resolution of a lidar sensor (and of course the digitization rate) is almost sufficient to blindly process the waveforms acquired with this sensor.

The multi-marked point process framework offers a very high flexibility for the given reconstruction problem. This is an evolving process which can integrate more or less parametric functions. The only condition to be ensured is the ability to evaluate the function for each value within the interval [0, #bins]. Thus, less restrictions are present that with a least-squares or EM approach. Few functions will provide more homogeneous results and it will be easier to assess their contribution, especially if they exhibit very distinct behaviours. A less restrictive library will generate more fuzzy results. The lidar waveforms are not very complex signals (even large-footprint ones), and inevitably several functions would produce almost identical curves. For instance, in addition to the Nakagami and Burr functions, the Weibull distribution has been introduced in early works [MLB+09]. It has been removed afterwards since its behaviour was finally too similar to the Nakagami, without any ability to simulate left-skewed peaks.

For medium and large footprint waveforms, the library functions allow to adjust the frequently occurring asymmetric peaks. Our approach is thus particularly relevant for such data: not all the peaks within the waveforms exhibit such asymmetry. We have processed these signals mainly in order to validate the approach and to show its versatility. It would be interesting to assess, in an analysis/synthesis framework, whether the segmentation of the waveforms in specific behaviours can be used in forest and large land-cover applications (LVIS and GLAS sensors, respectively).
For small footprints, the skewness of the echoes is less significant, and shows the present limitations of our model. For first echoes, the method is relevant and reveals us that most of the echoes are in reality symmetric. For slightly asymmetric echoes, all the asymmetric objects of the library are suitable and can be chosen, resulting in a sort of ambiguity. Even for symmetric echoes, the Burr and the Nakagami functions can be preferred to the Generalized Gaussian one since some combined ranges of parameters can produce almost symmetric curves. As mentioned above, the content of the library is therefore a key point

of the proposed approach. This is a limitation of the current version of our approach which should be targeted in future work.
Eventually, other non topographic waveforms have also been processed with the proposed approach, namely bathymetric green lidar waveforms [MLB+09]. They are ideally composed of two echoes corresponding respectively for the first and second echoes to the air-water boundary and the seafloor. They are both skewed, due to the scattering and spreading of pulse at the sea surface, and the water depth and the attenuation coefficient, respectively. Thus, they cannot be fitted with a Gaussian model and satisfactory results have also been found with such data. Since the second pulse can be represented by an exponential decaying function, in future works, such waveforms may be good candidates towards further demonstrating the versatility of the algorithm.

Many improvements and extensions of the proposed approach are conceivable. First, more specific proposition kernels may be introduced. For example, it should be interesting to specifically deal with overlapping echoes. Birth and death in a neighbourhood kernels may be a good solution for process complex shapes where several superimposed signals exist. We may therefore prevent some local signal misadjustments. Moreover, for two overlapping peaks, the selection of a model for the first peak has an influence on the fit of the second echo. Such influence has not been analysed. Dealing with such cases can come in useful for regularizing the proposed solution.
Besides, the automatic estimation of the parameters of the configuration energy is a well-known problem for marked point processes, and has been barely tackled in the literature. This is interesting in order to achieve fully unsupervised waveform reconstruction. Recent works have shown that embedding an algorithm within a Stochastic Expectation Maximization (SEM) framework is relevant [CDZ09]. Such technique has been applied in [HCDZ10] to solve various pattern recognition problems. Its main limitation is the significant computing time of the evaluation of the pseudo-likelihood (E step). 70% more time is necessary, which is, in our context, a crippling drawback.

Eventually, the RJMCMC technique is often adopted since the result is independent of the initialization step. On one hand, this can be considered as an appreciable feature of our approach. On the other hand, this is also a limitation. Indeed, as consecutive small-footprint waveforms along a scan line and in the orthogonal directions have very high probability to exhibit similar shapes, spatial interactions should also be included in the regularization term of the proposed model. One may want to take into account the output of the fitting procedure of the neighbouring waveforms. Such data-driven approach would smooth the selection of distinct models for consecutive waveforms and corresponding echoes. This is obviously the most intuitive knowledge we can integrate in the model. In urban areas, if we consider corresponding echoes of consecutive waveforms (along a scan line *i.e.*, in a temporal meaning), they will almost have the same location within the signals. The location will only slightly move depending on the slope of the surface *e.g.*, for a building roof or the ground. Additional echoes may sparsely spring whether a chimney or a low above-ground item are present. From a ground point of view, occlusions may appear in case of vegetated areas or buildings. Therefore, such regularization problem may be tackled as in classical visual tracking issues, using, for instance, Particle Filters. It has already been addressed in with a varying number of objects with surveillance video datasets [Smi07]. We consider such an approach as a very promising outlook, in addition

to the design on a global approach that is able to process all the waveforms at the same time.

Finally, Compressed Sensing [Don06] seems to be another promising dictionary-based method for waveform reconstruction.

Part III

Classification of urban areas

Chapter 5

Classification of 3D lidar point clouds

Subsequent to waveform processing, point-based classification is performed in order to discriminate *ground*, *vegetation*, and *building* points. For each 3D point, a feature vector is computed, gathering both spatial attributes and full-waveform attributes. Various kinds of spatial features are especially selected for this purpose, whereas all the previously computed FW features are introduced. Since achieving a high classification accuracy is coupled with assessing the relevance of the FW attributes, a feature selection is performed jointly with the selection process. Figure 5.1 sketches the workflow that will be set up.

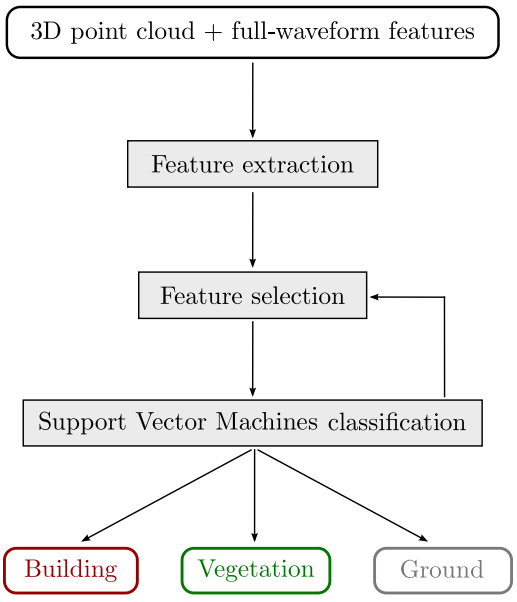


Figure 5.1: Overview of the classification process.

5.1 Introduction

5.1.1 Context

Part II has addressed the problem of waveform processing (decomposition and modelling) with two distinct methods. For each lidar signal, the outputs of both approaches are:

- A series of ranges that correspond to the positions of the modes of the signal. The ranges may be one of the parameters of the modelling function (as in Chapter 3), or has to be computed afterwards, once the parameters have been all estimated (*e.g.*, Burr and Nakagami models in Chapter 4).
- For each range, several shape features, parameters of the modelling functions.

The range values give estimates of the location of the targets: they are the required inputs for the georeferencing process, which allows to generate the 3D point cloud (see Section 3.2.5).
The shape features or so-called **full-waveform** features, such as amplitude, echo width or cross-section are additional information along with spatial attributes. They provide knowledge on the scattering properties of the targets, and consequently, they may be used for classification. This is especially relevant in urban areas since most of the echoes are single echoes: their shape (and therefore the associated features) is only influenced by the target, and not by the targets previously hit by the laser beam, as in forested areas. Such interactions are really difficult to model [WHBD08], and this is not yet possible to correct such disturbance.

The objective of this chapter is to perform point cloud (3D) classification in urban areas, subsequently to a waveform modelling step, *i.e.*, with the introduction of the full-waveform features. As mentioned in Chapter 1, three classes of interest have been selected: *ground*, *vegetation*, and *building*. We have noticed in Chapter 3 (Section 3.4) that it is not possible to classify the types of ground and road surfaces using the waveform properties alone. For assigning each return pulse to a specific surface type, additional information is required, such as optical images, but cannot result from the 3D geometric relationships of the returns within the point cloud. Such issue will therefore not be addressed in this chapter. Conversely, the classification of tree species mainly relies on the amplitude feature, and spatial attributes. It might be tackled at the same time. However, as we focus on urban areas and full-waveform features, this problem is considered out of our scope.

5.1.2 Related works

A point-wise classification strategy aims to label the 3D points resulting from the waveform modelling process. Each 3D point is associated to a feature vector mixing geometric features and full-waveform features. Geometrical features are spatial attributes relying only on the 3D coordinates and the echo number of the lidar point, whereas FW features are those extracted from the early modelling step. Feature selection is important since meaningful features facilitate accurate classification of the data. Indeed, for instance, buildings may have quite different appearances both with respect to their geometric dimensions and their reflectance properties. Such heterogeneous appearance may not be handled easily by

combining only planar- and amplitude-based attributes.
Segmentation followed by classification would be another possibility. It has been adopted, for instance, in [FNSZ06; SZ07; RHHP08] where region-growing, random walk, and normalized-cut segmentations precede the labelling task, respectively. In order not to aggregate distinct objects in the same segment, over-segmentation is often preferred as preliminary step. On the one hand, an object-based approach may lead to more homogeneous results than a point-based approach since (1) spatial coherency is enforced, and (2) segment-based features can be computed (area, compacity, etc. [Tóv06]). On the other hand, one may find the most appropriate measure to turn point features to a single segment attribute (mean? standard deviation?). This is not necessarily a straightforward task. Furthermore, the local discrepancies in the point-based labelling process may also be corrected subsequently with a contextual regularization step [CML04].

Several classification methods have already been applied to lidar data for urban scenes. Specific works on full-waveform data have already been presented in 2.6.3. Both unsupervised (*e.g.*, Mean-Shift algorithm in [Mel07]) and supervised (*e.g.*, cascade of binary classifiers in [CGCZ09]) classifications are conceivable. Rule-based methods have been favored since many features (such as the echo number or the height above ground) have a physical meaning, and may be discriminative using simply thresholds. Decision trees (DT) have therefore been adopted in [DHU+06; RHHP08; ATK+10]. On the one hand, DT may allow to discriminate a large number of classes (*e.g.*, trees, shrubs, grass, roofs, and road are labelled in [DHU+06]). On the other hand, such heuristic method may suffer from area-specific and sensor-specific thresholds [ATK+10]. Besides, DT can be easily computed with few number of features (but high classification accuracy may not be achieved), whereas a large feature vector may lead to very complex trees where the optimal splitting order may not be easy to find. Consequently, supervised approaches have been selected to deal with such issues.
For such methods, in addition to the feature vector, a set of data samples have class labels associated with them. This set is called the training dataset and is used to estimate the parameters of the classifier. An important underlying assumption is that the whole dataset is similar in terms of distribution of features to the training dataset. This means that the classifier must have observed similar features in the training in order to perform a good classification (the so-called generalization performance). A Gaussian Mixture Model approach, solved with the Expectation-Maximization algorithm, is adopted in [CML04], whereas the Random Forest classifier is chosen in [GCMB10] (see Section 5.6 for further details). Furthermore, since classifiers based on statistical learning theory have shown remarkable abilities to deal with both high-dimensional data and a limited training set, Support Vector Machines (SVM) have been naturally investigated for the classification of airborne lidar data in urban areas [LKHF06; SZ07; SBR10]. This method is well adapted to deal with high-dimensional feature space since the algorithm complexity does not depend on the data dimension. Furthermore, SVM belong to the non-parametric classification techniques, *i.e.*, no parametric probability density functions are required. Moreover, the training step is reduced to a convex optimization problem, and eventually the kernel trick allows to find a linear classifier in some feature space.

5.1.3 Strategy

A point-based classification based on Support Vector Machines is adopted. SVMs have been selected because a simple feature vector can be used as input for the discrimination task. As mentioned above and in Section 1.6, the feature vector partly stems from the modelling step. Therefore, even if the modelling functions are modified, our proposed workflow will not have to be adapted. This would not be the case for Decision Trees. The choice of Support Vector Machines is rather objective. Other methods, such as Classification and Regression Trees (CART) or ensemble classifiers (*e.g.*, Random Forests), would be as suitable as SVMs, and are known to provide the same level of performance.
The principle of SVM is presented in Section 5.2, as well as the classification strategy. Both spatial and full-waveform features are computed for each 3D lidar point. Afterwards, they are introduced into the supervised classifier. The analysis is also performed on the local neighbourhood level for the geometric attributes. One may find the set of selected features in Section 5.3.
The aim of this chapter is two-fold. In addition to 3D classification, we aim to assess the relevance of full-waveform features in urban areas. We would like to quantify the genuine contribution of the modelling step in comparison to simple on-board methods *i.e.*, traditional multiple pulse datasets. For that purpose, a feature selection has to be carried out in conjunction to the labelling process. Two main strategies are conceivable, and are presented in Section 5.4.
The full classification process is carried out both on Biberach and Amiens datasets on the basis of manually selected ground truth. The results are presented in Section 5.5.

Eventually, in parallel to the 3D analysis of urban areas, another approach involving full-waveform data has been developed with other authors. We also aimed at assessing the contribution of FW sensors with respect to multiple pulse lidar data. Random Forests have been adopted to perform 2D classification with feature selection in urban areas. Such an approach will be briefly presented, and the results will be put in parallel with our method.

5.2 Support Vector Machines and feature selection

5.2.1 Principle of SVMs

The general mathematical formulation of Support Vector Machines (SVMs) is briefly recalled in this section. SVMs can be considered as the merge of the optimal separating hyperplane theory [Vap95], with a modified margin for mislabelled examples [CV95] (linear classification), and the kernel trick applied to create non-linear classifiers [BGV92]. In-depth theoretical and practical aspects of SVMs can be found, for instance, in [SS02].

Supervised classification and statistical learning theory

A supervised algorithm aims to label a set of elements given a training set $\mathcal{S}$, such as $\mathcal{S} = \{(\mathbf{x}_1, y_1), .., \mathbf{x}_l, y_l)\} = \mathcal{D} \times \mathcal{L}$. $\mathcal{D}$ is the data space, and $\mathcal{L}$ is the label space. For instance, $\mathcal{L} = \{-1, +1\}$ for a two-class problem. The problem is therefore to find a

function $f \in \mathcal{F}$ that assigns label $+1$ (resp. -1) to the $\mathbf{x}_i$ such as $f(\mathbf{x}_i) \geq 0$ if $y_i = 1$ (resp. $f(\mathbf{x}_i) < 0$, if $y_i = -1$). f is a function from the vector space $\mathcal{D}$ to the label space $\mathcal{L}$. The decision boundary is $\ker(f) = \{\mathbf{x} | f(\mathbf{x}) = 0\}$, and a classifier g is derived by considering the sign of such a function to solve problem $y = g(\mathbf{x}) = \text{sgn}(f(\mathbf{x}))$.
One solution may be to minimize the training error, called empirical risk $R_{\text{emp}}(f)$, *i.e.*, the rate of misclassified samples over $\mathcal{S}$.

$$R_{\text{emp}}(f) = \frac{1}{l} \sum_{i=1}^{l} L(f(\mathbf{x}_i), y_i) \tag{5.2.1}$$

where L is a loss function, *e.g.*, $L(\mathbf{x}, y) = \frac{1}{2}|y - f(\mathbf{x})|$.
The so-called "Empirical Risk Minimization" principle has two main limitations. First, there is no unique solution. Second, minimizing the error over a subset is not equivalent to minimize the error over all the elements: f may have not a good generalization ability, *i.e.*, we could face to over-learning, due to the fact that the surface $\ker(f)$ is very complex, and sticks to the training set. The statistical learning theory uses only simple assumption about the complexity of the class of possible functions $\mathcal{F}$: simple functions are favored so that to achieve a good generalization ability. The Vapnik-Chervonenkis (VC) theory detailed in [Vap98] allows to measure such complexity with the VC dimension h, and provides a relation between the empirical risk and the total risk, *i.e.*, the misclassification over $\mathcal{D}$. h is defined as the largest number of points that can be separated with $\mathcal{F}$, whatever $\mathcal{L}$. Hence, the structural risk minimization principle allows to select the function f that minimizes a given upper bound error.

$$R(f) \leqslant R_{\text{emp}}(f) + V(h, l) \tag{5.2.2}$$

with $V(h, l)$ detailed in [Vap98], and only dependent of $\mathcal{F}$ (now limited). Such inequality should be minimized for the training step of the classifier, according to VC theory.
A specific case of functions f are linear functions. They are described in the next section, and leads to the definition of the SVM classifier.

Linear SVMs

The classification is carried out using a linear discriminant function f. The decision function has the following expression:

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle_{\mathcal{D}} + b \tag{5.2.3}$$

where $\mathbf{w} \in \mathcal{D}$, $b \in \mathbb{R}$, and $\langle , \rangle$ is the inner product. For a supervised classification problem, we define the separation hyperplane H as a linear decision function of $\mathbf{w}$ and b which separates the space into two half-spaces (each half-space corresponds to a given class), *i.e.*, $\forall\ (\mathbf{x}_i, y_i) \in \mathcal{S},\ y_i = \text{sgn}(\langle \mathbf{w}, \mathbf{x}_i \rangle_{\mathcal{D}} + b)$.
In the following, we consider real n-dimensional vector space *i.e.*, $\mathcal{D} = \mathbb{R}^n$, and $\mathcal{L} = \{-1, 1\}$ in a two-class problem. In Section 5.3, we will see that n is close to 30, which is, for remote sensing applications between multispectral ($n \in [3, 10]$) and hyperspectral data ($n \gg 10$). If we assume that the data are linearly separable, thus the learning problem is to find the parameters $(\mathbf{w}, b)$. However, many separating hyperplanes are acceptable, and the optimal separating hyperplane is found so that to minimize the upper bound of the risk R (see

Equation 5.2.2). In order to correctly classify the training data set, we should have:

$$\langle \mathbf{w}, \mathbf{x} \rangle + b \geq 1 \quad \text{if} \quad y_i = +1 \tag{5.2.4}$$

$$\langle \mathbf{w}, \mathbf{x} \rangle + b < -1 \quad \text{if} \quad y_i = -1 \tag{5.2.5}$$

In the statistical learning theory, the VC dimension is estimated by the margin for a class of separating hyperplanes. The margin is defined by the minimum distance of a training sample from the decision boundary (not necessarily correctly classified). It is given by the distance ζ between the two samples that are the closest to the hyperplane $(\mathbf{x}_{+1} - \mathbf{x}_{-1})$, projected onto the unary vector normal to the hyperplane.
$\mathbf{w}.(\mathbf{x}_{+1} - \mathbf{x}_{-1}) = 1 \Rightarrow \zeta = 2\|\mathbf{x}_{+1} - \mathbf{x}_{-1}\| \Rightarrow \zeta = \frac{2}{\|w\|}$. Finally, by considering Equation 5.2.2, the learning strategy is to both minimize:

The empirical risk : the hyperplane parameters are constrained to give perfect classification of $\mathcal{S}$. It means that: $\forall\, (\mathbf{x}_i, y_i) \in \mathcal{S}, y_i \langle \mathbf{w}, \mathbf{x}_i \rangle + b \geq 1$.

The complexity term : by minimizing $\|\mathbf{w}\|^2$. Please see [Sch97; MMR+01] for more details.

Hence, we have to solve the following constrained quadratic optimization problem:

$$\text{minimize} \quad \frac{\langle \mathbf{w}, \mathbf{w} \rangle}{2} \tag{5.2.6}$$

$$\text{subject to} \quad \forall\, (\mathbf{x}_i, y_i) \in \mathcal{S}, \quad y_i \langle \mathbf{w}, \mathbf{x}_i \rangle + b \geq 1 \tag{5.2.7}$$

Such a problem can be solved by introducing the Lagrange multipliers. One dual variable α_i is introduced for each constraint, *i.e.*, for each training sample. The gradient of the function must be orthogonal to the tangent plane of the constraints. That is the projection of the gradient of f onto the space of directions tangent to the constraint "surface" is zero. The problem is now:

$$L(\mathbf{w}, b, \alpha) = \frac{\langle \mathbf{w}, \mathbf{w} \rangle}{2} + \sum_{i=1}^{l} \alpha_i \left(1 - y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b)\right) \tag{5.2.8}$$

The Lagrange multipliers have to be minimized with respect to $\mathbf{w}$ and b (primal parameters), and maximized with respect to the parameters α_i (dual parameters). At the optimal point, we have the following gradient expressions:

$$\frac{\partial L}{\partial \mathbf{x}} = \mathbf{w} - \sum_{i=1}^{l} \alpha_i y_i \mathbf{x}_i = 0 \tag{5.2.9}$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{5.2.10}$$

Combining the two latter equations with Equation 5.2.8 gives us the dual quadratic problem with bound constraints, and a linear equality constraint, with only one variable (the α_i):

$$\max_{\alpha} g(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \tag{5.2.11}$$

$$\text{subject to} \quad \alpha_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{5.2.12}$$

Calling α_i^* the solution of this problem, the solution parameters $\mathbf{w}^*$ are obtained from Equation 5.2.10:

$$\mathbf{w}^* = \sum_{i=1}^{l} \alpha_i^* y_i \mathbf{x}_i, \tag{5.2.13}$$

and b^* can be derived from the constraints of Equation 5.2.7:

$$\alpha_i^* [1 - y_i(\langle \mathbf{w}^*, \mathbf{x}_i \rangle + b^*)] = 0 \tag{5.2.14}$$

If $\alpha_i^* \neq 0$, the corresponding $\mathbf{x}_i$ lies exactly on the boundary, *i.e.*, at a distance equal to the half margin from the optimal separating hyperplane. Otherwise, $\alpha_i^* = 0$, and $\mathbf{x}_i$ lies on the "good side" of the boundary.
One can note that $\mathbf{w}^*$ is a linear combination of the vectors $\mathbf{x}_i$ for which Lagrange multipliers are not null. Such vectors are the so-called **support vectors** (see Figure 5.2.a). They are the closest elements to the optimal separating hyperplane. Moreover, the optimal set of parameters ($\mathbf{w}^*$, b^*, α_i^*) remain the same if non-support vectors are removed from the training set.
The final decision is given by:

$$g(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{l} \alpha_i^* y_i \langle \mathbf{x}, \mathbf{x}_i \rangle + b^* \right) \tag{5.2.15}$$

Such decision rule has been obtained assuming that the data are linearly separable, which is far too restrictive for remote sensing problems. To solve this problem, slack variables ξ_i are introduced (see Figure 5.2.b). A slack variable is a nonnegative variable that turns an inequality into an equality constraint.
We therefore now consider soft-margin constraints: $\forall\ (\mathbf{x}_i, y_i) \in \mathcal{S}, y_i \langle \mathbf{w}, \mathbf{x}_i \rangle + b \geq 1 + \xi_i$. Errors are allowed during the training process, and that avoids over-fitting the classifier to the training samples, which would result in poor performance. The optimization problem described in Equation 5.2.7 has now changed to:

$$\text{minimize} \quad \frac{\langle \mathbf{w}, \mathbf{w} \rangle}{2} + C \sum_{i=1}^{l} \xi_i \tag{5.2.16}$$

$$\text{subject to} \quad \forall\ (\mathbf{x}_i, y_i) \in \mathcal{S},\ y_i \langle \mathbf{w}, \ \mathbf{x}_i \rangle + b \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0 \tag{5.2.17}$$

C is a constant which controls the number of training errors and therefore determines the trade-off between margin maximization and training error minimization. C is called the regularization parameter. A larger C assigns a higher penalty on the training error, whereas a smaller value tends to increase the margin and the training errors. Equation 5.2.8 becomes:

$$L(\mathbf{w}, b, \xi, \alpha, \beta) = \frac{\langle \mathbf{w}, \mathbf{w} \rangle}{2} + \sum_{i=1}^{l} \alpha_i (1 - \xi_i - y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle + b)) - \sum_{i=1}^{l} \beta_i \xi_i + C \sum_{i=1}^{l} \xi_i \tag{5.2.18}$$

The dual problem described in Equation 5.2.12 is finally slightly modified with now: $0 \leq \alpha_i \leq C$. The method of Lagrange multipliers is only valid for equality constraints. For inequality constraints, it is generalized by the Karush–Kuhn–Tucker conditions, that require additional constraints. A new derivative can be computed:

$$\frac{\partial L}{\partial \xi_i} = -\alpha_i - \beta_i + C = 0 \tag{5.2.19}$$

By substituting the three partial derivatives into Equation 5.2.18, we obtain:

$$\max_{\alpha} g(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \tag{5.2.20}$$

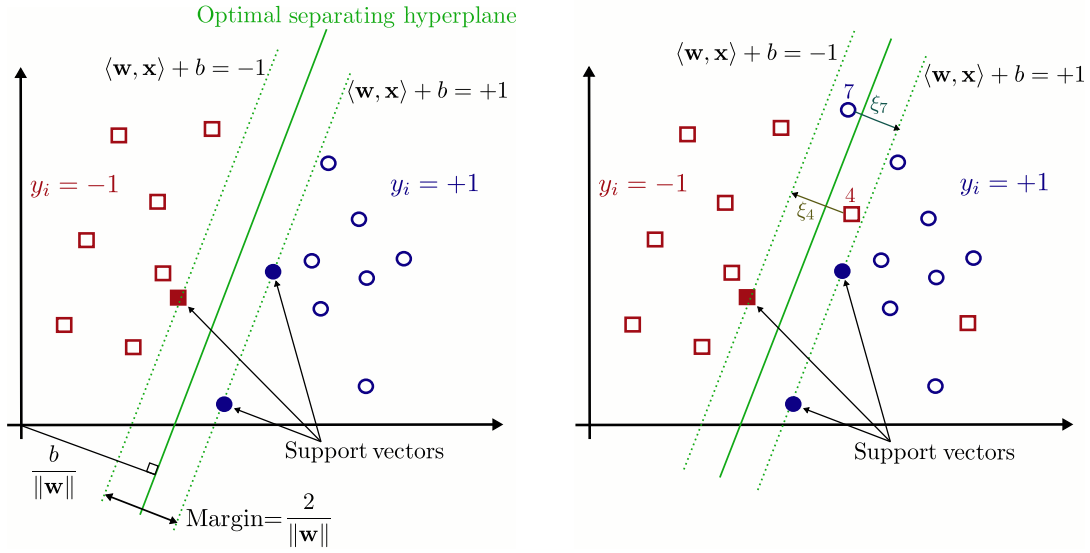
$$\text{subject to} \quad 0 \leq \alpha_i \leq C \quad \text{and} \quad \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{5.2.21}$$

Furthermore, the Karush–Kuhn–Tucker conditions, that allow to compute b, are:

$$\alpha_i \left(1 - \xi_i - y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b)\right) = 0 \tag{5.2.22}$$

$$\text{subject to} \quad \beta_i \geq 0 \quad \text{and} \quad \beta_i \xi_i = 0 \tag{5.2.23}$$

Equation 5.2.23 requires that $\alpha_i = 0$ or $1 - \xi_i - y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b) = 0$. It means only some of the α_i are not equal to 0. The $\alpha_i \neq 0$ define the separating hyperplanes. One can notice that the $\alpha_i = 0$ are not involved on the formulation of the optimal classifier given in Equation 5.2.15. Therefore, the two separating hyperplanes given by $\mathbf{w}$ are supported by the training samples with $\alpha_i^* \neq 0$ (the support vectors).



(a) Linear case.

(b) Linear case with soft margins.

Figure 5.2: Classification of separable and non separable data by SVMs using the optimal separating hyperplane.

Non-linear SVMs

However, as underlined in Section 3.4 for full-waveform features, classes of interest are partially overlapped. Therefore, a linear classifier would not be optimal. When the classification problem is not linearly separable, one solution consists in changing the feature

space. The **kernels methods** are designed to both have the effectiveness of a linear training model, and the discriminant ability of a non-linear model. The data is projected in a higher dimension space using a non-linear mapping function $\Phi : \mathcal{D} \rightarrow \mathcal{H}$, in which the new distribution of samples enables the fitting of a linear hyperplane. Kernels methods provide non-linear hyperplanes, and improve classification abilities. The same margin optimization method can then be performed.
The optimization problem can be formulated as follows:

$$\max_{\alpha} g(\alpha) = \sum_{i=1}^{l} \alpha_i - \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle \tag{5.2.24}$$

$$\text{subject to} \quad 0 \leq \alpha_i \leq C \quad \text{and} \quad \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{5.2.25}$$

which gives:

$$g(\mathbf{x}) = \text{sgn}\left(\sum_{i=1}^{l} \alpha_i y_i \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}_i) \rangle + b \right) \tag{5.2.26}$$

One can notice that projecting $\mathcal{D}$ in a higher dimension space do not complexify the algorithmic resolution of the problem since the parameters to be optimized are still of size l.
Finding Φ is a difficult problem. In practice, the $\mathbf{x}_i$ points are implicitly projected in $\mathcal{H}$ by defining a kernel $K : \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}$ such as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle \tag{5.2.27}$$

This is the so-called *kernel trick*. The kernel function represents an inner product in the new representation space. Hence, it characterizes a certain similarity of two data in the initial space.
Two approaches are possible to define a kernel. On one hand, K can be directly defined. On the other hand, indirectly, one can use the Φ function and the inner product in the new representation space. In fact, the knowledge of K is sufficient to compute the optimal classifier. SVMs work implicitly on the feature space. The kernel has only to fulfill Mercer's condition [SBS98]: it should be a symmetric definite positive function. Therefore, the problem defined in Equation 5.2.25 is still convex.
One main property of kernels is the possibility to compose them without knowing both the direct and indirect expressions of the two kernels. For instance, knowing the direct expression of the first one, and the indirect expression of the second one is enough to know the direct expression of their product.

Multi-class SVMs

SVMs are intrinsically binary classifiers. When having $n \geq 3$ classes of interest, various approaches are possible to address the problem. Different possibilities are:

- Multi-class optimization: the design of the SVM can be modified in order to incorporate the multi-class learning directly in the quadratic programming solving [WW98];

- Two-class optimization: a set of binary classifiers is combined. In such a case, two strategies are possible [SS02]:
  - One versus the rest. This is the simplest solution of the decomposition methods. n binary classifiers are applied on each class against the others (*i.e.*, n hyperplanes are constructed). Each sample is assigned to the class with largest output of the decision function: $\forall\, i,\ y_i = \underset{k\in[1,n]}{\arg\max}\ f_k(\mathbf{x}_i)$ [BSB+96].
  - One against one. For such pairwise classification, $n\,(n-1)/2$ binary classifiers are computed on each pair of classes. Each sample is assigned to the class getting the highest number of votes. A vote for a given class is defined as a classifier assigning the sample to that class [PV97]. Such decision rule may vary according to the problem and the number of classes.

The number of classes to be discriminated simultaneously can be large. Therefore, the number of parameters to be estimated increases considerably in a multi-class optimization formulation. The method tends to be less stable and, accordingly, affects the classification performances in terms of accuracy. For this reason, multi-class optimization has not been as successful as the approach based on the two-class optimization.
According to a comparison study [WW98], correctly tuned SVMs with "one versus the rest" solution have shown performances slightly inferior but close to the "one-against-one" approach. However, one can notice the performance of the first technique can be compromised due to unbalanced training datasets. However the "one-against-one" approach is more computationally intensive (larger number of classifiers), the classification task is divided in much simpler ones.
Other strategies have been proposed such as hierarchical tree-based approaches [MB04]. In spite of an improved computing time, the results exhibit lower accuracy since errors can be propagated inside the branches of the trees. Furthermore, such methods also require a far more complex training step.
We choose the "one-against-one" approach mainly because it has been shown to be more suitable for large problems [HL02].

Implementation

The **LIBSVM** software[1] is used to implement the SVM algorithm. Slack variables are introduced (soft-margin classifier). Then, the parameter C has to be optimized with the kernel hyperparameters (see Section 5.2.2).
There are several standard methods to solve the above mentioned convex quadratic programming problem, with linear and box constraints. However, fast algorithms have been developed to cope with the important number of training samples, which may turn these methods ineffective. Vapnik has, for instance, proposed a chunking algorithm [Vap98], which removes from the problem the non-support vectors (since the problem remains the same). Then, his algorithm breaks the large problem into a series of smaller sub-problems, which aims to identify the non-null Lagrange multipliers and remove the other ones.
The matrix Q such as $Q_{i,j} = y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)$ is usually fully dense and may be too large to be stored. Decomposition methods are designed to handle such difficulties (*e.g.*, [OFG97; Pla98a; CFL06]). Unlike most optimization methods which update the whole vector α

[1] Available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

in each step of an iterative process, the decomposition method modifies only a subset of α per iteration. This subset leads to a small sub-problem to be minimized in each iteration. A specific case is the Sequential Minimal Optimization (SMO) [Pla98b], which restricts the subset to have only two elements. Therefore, in each iteration, one does not require any optimization software in order to solve a simple two-variable problem. Indeed, the 2-dimensional sub-problem can be solved analytically, *i.e.*, without time-consuming numerical quadratic optimization. The SMO strategy has been adopted in LIBSVM.

5.2.2 Kernel selection

Without sufficient a priori knowledge of the influence of geometric and radiometric parameters on the pulse shape, the design of a kernel dedicated to our specific purpose given our cues is a very difficult task. Therefore, a classical kernel is selected, the Gaussian kernel, defined as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\vartheta^2}\right) \qquad \text{with } \vartheta > 0 \tag{5.2.28}$$

ϑ tunes how similar to the training data the test data is expected to be. $\vartheta \to 0$, for instance, leads to over-fitting, and consequently reveals a low generalization ability of the classifier.
The polynomial kernel has also been tested but have not shown better performance. Moreover, since this kernel has more hyperparameters, the model selection is more complex (see below). The grid search becomes a 4-dimension tensor search (see below). Moreover, the polynomial kernel has more numerical difficulties because kernel values may go to the infinity or zero, while the degree is large (in contrast, $0 < K_{\text{Gauss}}(\mathbf{x}_i, \mathbf{x}_j) \leqslant 1$). The other reason why adopting the Gaussian kernel is that many authors have stated that more sophisticated kernels do not mean better results.

Hyperparameter tuning

There is no criterion concerning the choice of the parameters involved in the optimization problem. A simple way to choose them is a *trial-and-error* strategy *i.e.*, to take several values and test their efficiency in order to select the most suited ones. [Fau07] tackles this problem by considering some a priori to choose a proper value for the penalty factor C. In their experiments, C is set to 200. They have a high penalty for non-separable points and impose no training errors, considering the data are linearly separable in the feature space. The grid search therefore becomes a line search.

Because optimal values of C and ϑ are not known beforehand, we prefer a simple variation of the latter method. A grid search exhaustively sweeps the parameter space, and for each point the cross-validation accuracy (CVA) is computed. In a v-fold cross-validation procedure, the training data are divided in v subsets of equal size. The classifier is trained on $v-1$ subsets and ran on the remaining one. The CVA represents the percentage of samples correctly classified averaged over all the subsets when they were used as the testing subset (mean between the True Positive Rate and the True Negative one). The (C,ϑ) grid is composed of exponentially growing values of C and ϑ, for instance, in our study $C, \vartheta = 2^{-15}, 2^{-13}, .., 2^{15}$. After the coarse grid search, a finer one is computed in a smaller range around the optimal parameters found in the first step. Such grid search is

necessary since the CVA over (C,ϑ) set is not convex. The values of the kernel parameters can have a significant impact on the learning capacity but their optimum values are not critical: there is a range of values which gives almost the same cross-validation accuracies. Consequently, this is not necessary to perform in-depth grid search.
The CVA has one main drawback. The SVM needs to be trained many times. For instance, if $v = 4$ and 15 values of two hyper-parameters are tested, each binary classifier requires 900 training steps. Therefore with more than 2 hyper-parameters, since the complexity of such grid search is exponential with respect to the number of optimized parameters, training and test become intractable even for small-scale problems. This is the main justification of choosing the Gaussian kernel for our problem. Concerning hyper-parameter tuning, the CVA has been shown to be a suitable performance measure among other solutions [DKP03]. We choose $v = 5$, whatever the area of interest.

More advanced methods have also been proposed. The VC dimension h, which defines an upper bound of the risk (Equation 5.2.2) can be used. [Kee01; CVBM02] find the optimal hyperparameters by minimizing an error bound, based on the risk upper bound and function of the kernel parameters. The solution is found with a gradient descent, or more effective techniques [BTB05]. These training strategies are compared with the CV method in [Fau07] for hyperspectral data. The main conclusion is that the gradient-based approaches are able to fit a much larger number of parameters (more than 100 according to [CVBM02]), with higher classification accuracies. However, such methods are mainly useful for advanced kernels or composition of kernels, and is not necessary for our (simple) classification task. More discussions on cross-validation techniques and optimality with respect to the given problem can be found in [AC10].

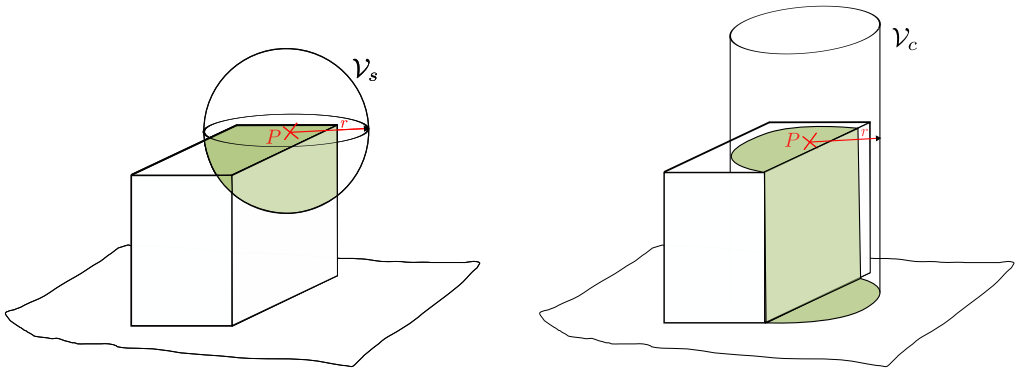
5.3 Features of interest

5.3.1 Various kinds of features (and computational strategies)

Our feature vector for each 3D lidar point has 27 components. All the attributes are lidar based. Part of the vector is composed of purely geometric features (15), and attributes derived from the echo number (4). The remaining part comes from the two waveform modelling methods that have been proposed (8).

Geometrical (spatial) features are computed with the single knowledge of the 3D coordinates, *i.e.*, without any waveform recording (multiple pulse sensors). However, they are indubitably improved by the waveform processing step. Retrieving additional points allows a better 3D representation of the scene and a better computation of several attributes. Three groups of features can be computed. They are described below.

Echo-based features The echo number or the type of echo (*single*, *first of many*, *intermediate*, or *last*) may be discriminant, especially for vegetated areas. Multiple-pulse sensors nowadays provide such information, but it is demonstrated that on-board echo detection methods may fail in case of waveform of complex shapes. As a consequence, an efficient peak detection after waveform digitization may supply more reliable echo numbers, and derived features.



(a) Spherical neighbourhood. (b) Cylindrical neighbourhood.

Figure 5.3: Two kinds of 3D environments for computing geometric features (adapted from [GJT07].

The **full-waveform features** come from the processing of the waveforms described in Chapter 3 and 4.
Since the Generalized Gaussian model provides interesting features, its three morphological parameters are fed into the feature vector. Two other attributes derived from the cross-section are also computed since recent works have highlighted their relevance for classification purposes [ATK+10; Wag10].
Furthermore, we also want to benefit from the stochastic approach. The function selected in the predefined library is therefore introduced. Eventually, the three models have not the same parameters and the same number of parameters. When dealing with a fluctuating number of parameters, one solution is to design a specific kernel that can handle such cases. However, this is not enough for us since, even if with a fixed number of parameters, this is not straightforward to compare the features of two models. Therefore, we have decided to synthesize all these parameters in a single common one, the peak skewness, that enhances the asymmetry, or not, of the echoes.

Geometrical and echo-based features are computed using a volumetric approach, *i.e.*, lidar points included in a restricted 3D neighbourhood $\mathcal{V}$ are used in addition to the current lidar point P. $\mathcal{V}$ is called *environment*. It can be a sphere centered on the point P ($\mathcal{V}_s$), or a vertical cylinder ($\mathcal{V}_c$). $\mathcal{V}_c$ includes all the points inside a 3D environment without restriction on the vertical position. It favors the measure of the dispersion in height of the neighbours. These two environments have a single parameter, the radius r (see Figure 5.3).

5.3.2 Description

The features may be separated as follows:

- **Height features**: they are related to the altitude of the current 3D point. They may discriminate both above-ground points and points with a significant altimetric scattering.

- Δz: height difference between the lidar point and the lowest point found in a large cylindrical volume whose radius has been experimentally set to 20 m.
- Δz_{fl}: the height difference between the first and the last pulses of the waveform.
- σ_z: the height variance of the 3D points included in $\mathcal{V}_c$.

$$\sigma_z^2 = \frac{1}{|\mathcal{V}_c|}\sqrt{\sum_{i\in\mathcal{V}_c}(z_i - \bar{z})^2}$$

where $\bar{z}$ is the mean height in $\mathcal{V}_c$.

- Eigenvalue features: a covariance matrix of the 3D coordinates is computed in $\mathcal{V}_c$. Such matrix can be considered as a 3D structure tensor [TMMT04] that gives information about the repartition of the points within $\mathcal{V}_c$. The three eigenvalues λ_1, λ_2, and λ_3 (descending order) are a direct link to the dimensions of the ellipsoid that really represents the local 3D structure [Sze10]. A flat ellipsoid ($\lambda_1 \simeq \lambda_2 \gg \lambda_3$) means the point cloud is locally plane. When $\lambda_1 \gg \lambda_2; \lambda_3$, we face a linear structure (*e.g.*, building edge), whereas $\lambda_1 \simeq \lambda_2 \simeq \lambda_3$ corresponds to volumic structures such as vegetated areas. Many combinations of these features are conceivable and may provide discriminant features, especially in urban areas [JG09a; CGCZ09]. A visual assessment has allowed us to keep the most promising ones.

- Σ_λ: the sum of the three eigenvalues.
- e_1, e_2, and e_3: the three eigenvalues normalized by Σ_λ ($\forall\, i \in [1,3],\ e_i = \lambda_i/\Sigma_\lambda$).
- A_λ, O_λ, P_λ, E_λ, S_λ: anisotropy, omnivariance, planarity, eigentropy, and scatter, respectively. They are eigen-based features describing the spatial local distribution. Their formulation are taken from [GT06; TMT10], with slight variations (normalizing coefficient).

$$A_\lambda = \frac{e_1 - e_2}{e_1} \quad ; \quad O_\lambda = \sqrt[3]{e_1 e_2 e_3} \quad ; \quad P_\lambda = 2(e_2 - e_3)$$

$$E_\lambda = -\ln(e_1)^{e_1} - \ln(e_2)^{e_2} - \ln(e_3)^{e_3} \quad ; \quad S_\lambda = \frac{e_3}{e_1}$$

- Local plane features: they are related to the 3D local plane Π. Such plane has been estimated by a robust M-estimator with norm $L_{1.2}$. The selection of an optimal p value for L_p norms has been investigated in [XZ96], and for p around 1.2, a good estimate may be expected. The deviation angle is useful for discriminating ground from building roofs ($N_z \simeq 0$ and $N_z \in [-45^o, 45^o]$, respectively). The two other features are good hints for enhancing the altimetric scattering around P.

- N_z: deviation angle of the local normal vector from the vertical direction.
- σ_{N_z}: variance of the deviation angles, computed for all points in $\mathcal{V}_s$.
- $\mathcal{R}_z$: residuals of the local plane Π.

$$\mathcal{R}_z = \sum_{i\in\mathcal{V}_c}\frac{(d_i)^p}{p}$$

where d_i is the distance between the lidar point $i \in \mathcal{V}_c$ and the plane Π, and p the norm (here $p = 1.2$).

- **Echo features**: they enhance the multiple scatterings of the laser pulse.

  - n: the echo number.
  - N: the total number of echoes within the waveform of P.
  - N_e: the normalized number of echoes. $N_e = \frac{n}{N}$.
  - PDR: the point density ratio. This is ratio between the number of points within $\mathcal{V}_s$ and within $\mathcal{V}_c$ [RHHP08].

- **Full-waveform features**: their relevance has already been discussed in Section 3.4.

  - A, w, σ, α: echo amplitude, width, cross-section, and shape, respectively. They have been retrieved using the General Gaussian model (see Chapter 3).
  - σ^0: the cross-section per illuminated area. We have: $\sigma^0 = \dfrac{\sigma}{A_i}$, where A_i is the illuminated surface area (see [WHU+08; Wag10]).
  - γ: the backscatter coefficient, which is σ^0 corrected from the incidence angle. If we assume a circular transmitter aperture,

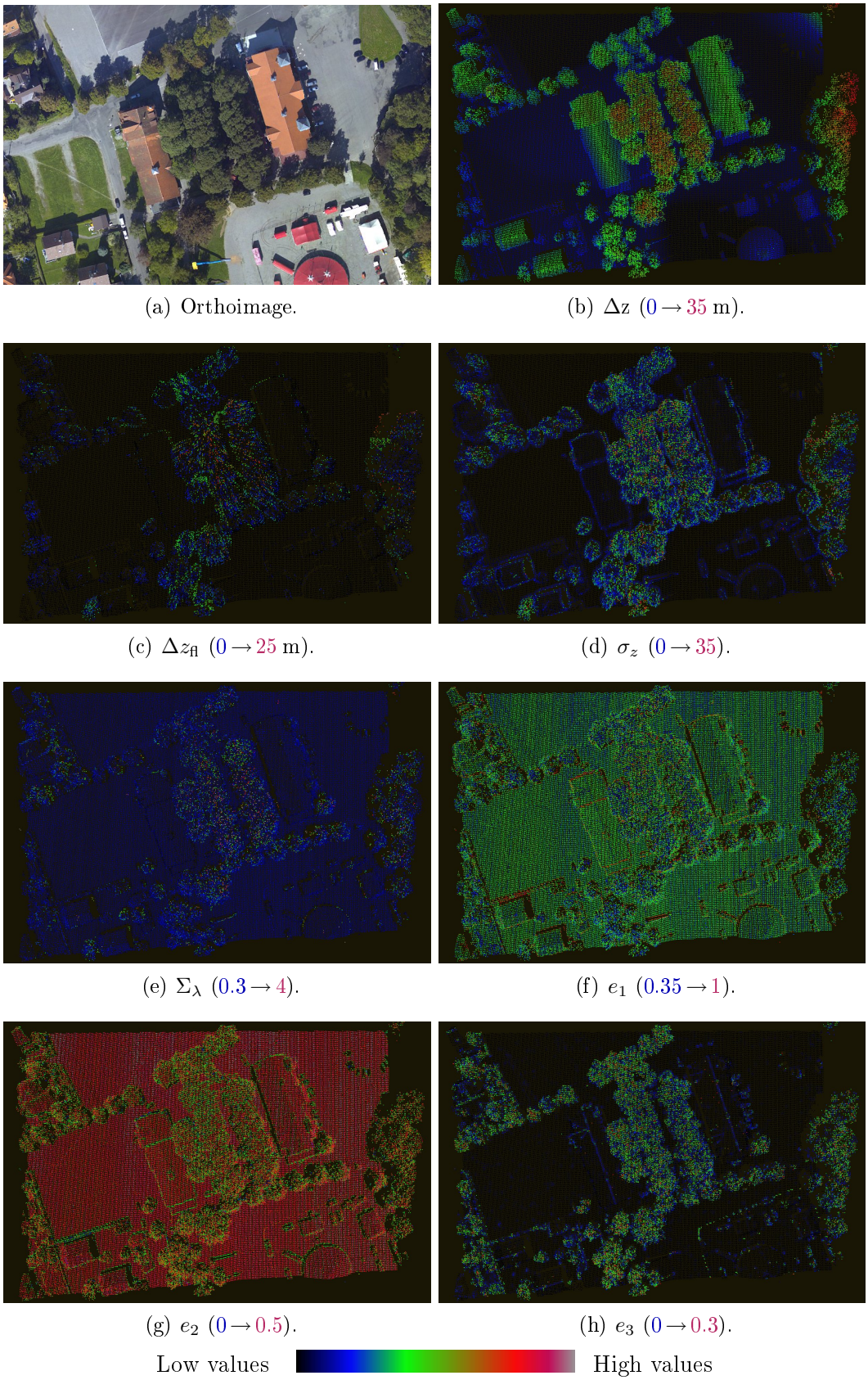
    $$\gamma = \frac{4\sigma}{\pi R^2 \beta^2}$$

    where R is the range between the point and the sensor, and β is the beam divergence.
  - $\mathcal{M}$: the modelling function selected in the marked point process approach (Chapter 4).
  - s: the echo asymmetry. This is the ratio between the widths at half-maximum that are located before and after the peak mode. $s < 1$ means that the echo is left-skewed, whereas $s > 1$ reveals a right-skewness.

Finally, the **feature vector** f_v can be written as follows:

$$f_v = \begin{cases} \Delta z, \Delta z_{\text{fl}}, \sigma_z; \\ \Sigma_\lambda, e_1, e_2, e_3, A_\lambda, O_\lambda, P_\lambda, E_\lambda, S_\lambda; \\ N_z, \sigma_{N_z}, \mathcal{R}_z; \\ n, N, N_e, PDR; \\ A, w, \sigma, \alpha, \sigma^0, \gamma, \mathcal{M}, s \end{cases} \tag{5.3.1}$$

Many other geometric features are conceivable and may be relevant for discriminating our three classes of interest. A large body of literature has indeed addressed the specific problem of building detection and some attributes may also be useful in our context. For instance, border gradients or texture features can be computed from raster data such as Grey Level Co-occurrence Matrix features [MKH07] or slope indices [EJL+01; Tóv06].



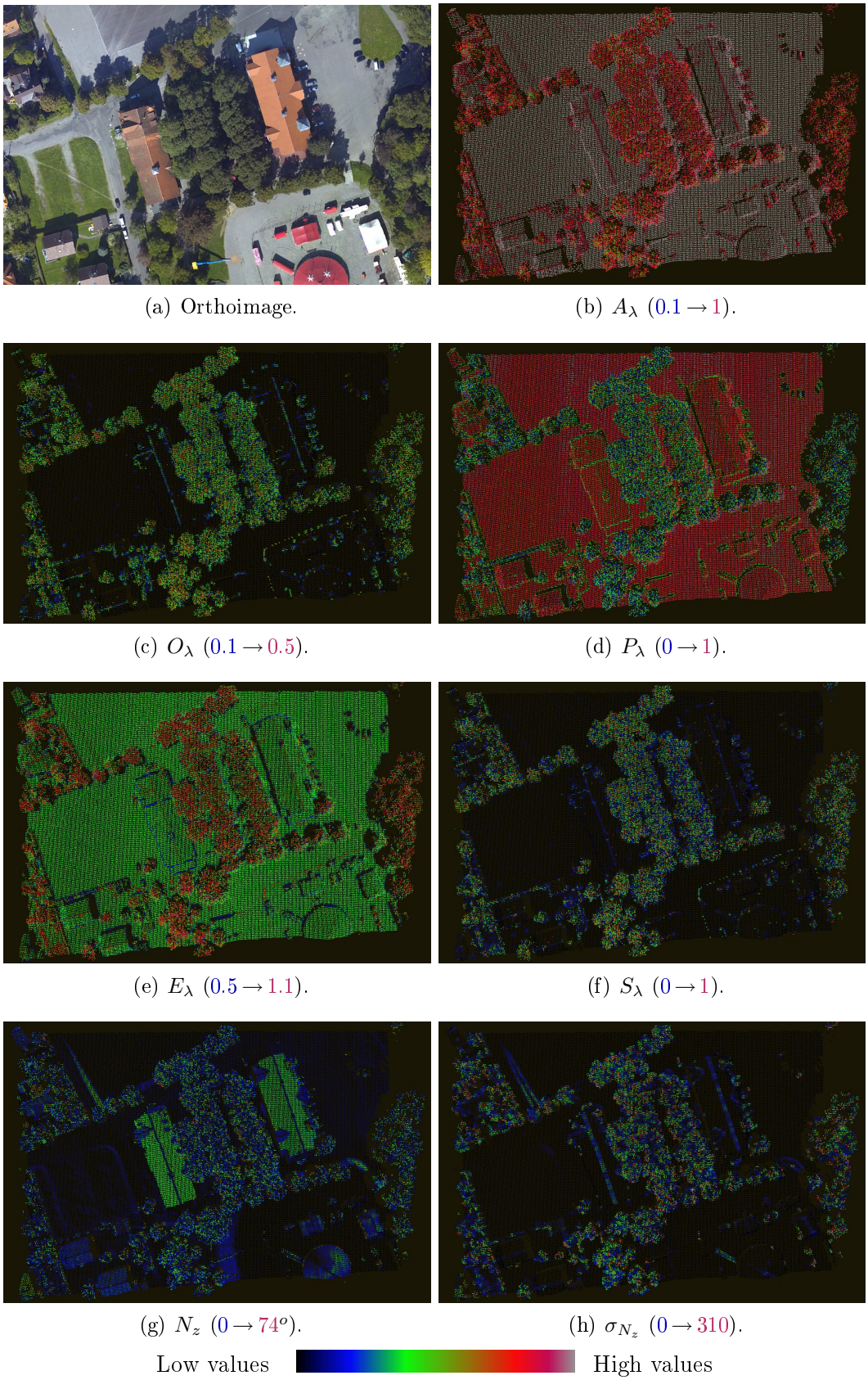
(a) Orthoimage. (b) Δz ($0 \rightarrow 35$ m).

(c) Δz_{fl} ($0 \rightarrow 25$ m). (d) σ_z ($0 \rightarrow 35$).

(e) Σ_λ ($0.3 \rightarrow 4$). (f) e_1 ($0.35 \rightarrow 1$).

(g) e_2 ($0 \rightarrow 0.5$). (h) e_3 ($0 \rightarrow 0.3$).

Figure 5.4: Features of interest over the area of Biberach (Part 1).



(a) Orthoimage.

(b) A_λ ($0.1 \rightarrow 1$).

(c) O_λ ($0.1 \rightarrow 0.5$).

(d) P_λ ($0 \rightarrow 1$).

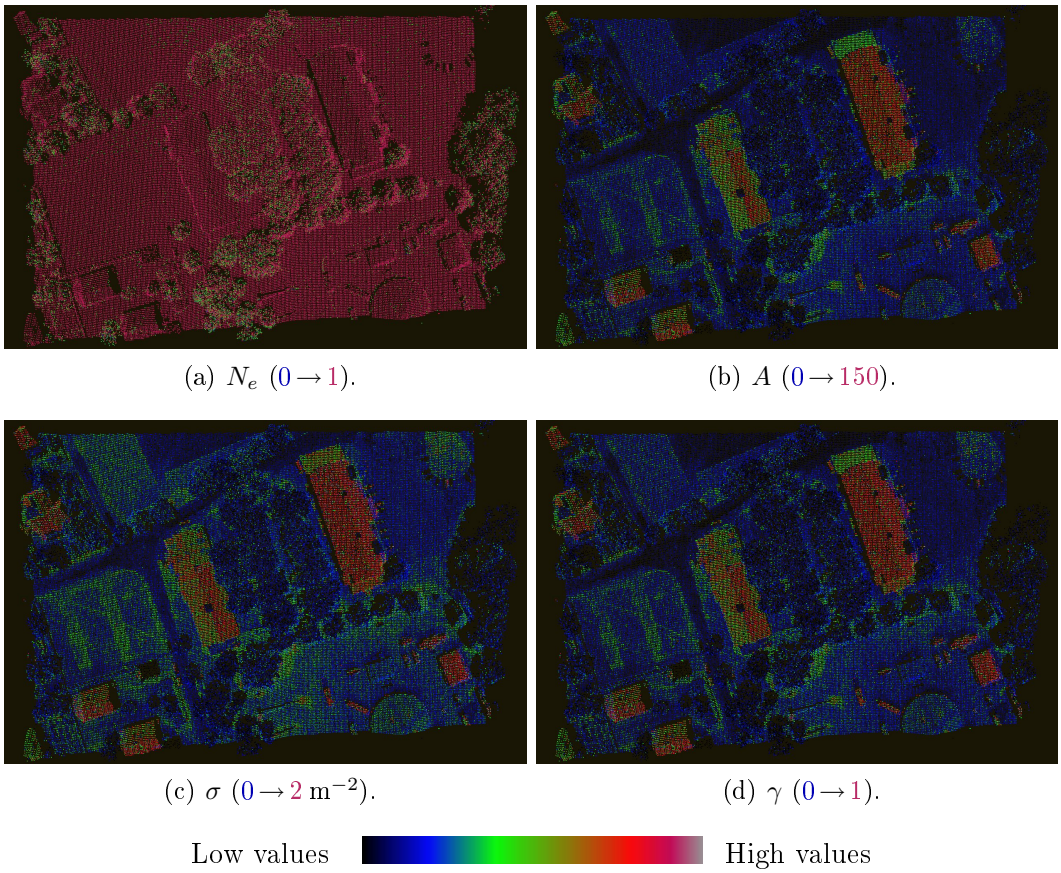
(e) E_λ ($0.5 \rightarrow 1.1$).

(f) S_λ ($0 \rightarrow 1$).

(g) N_z ($0 \rightarrow 74^o$).

(h) σ_{N_z} ($0 \rightarrow 310$).

Figure 5.5: Features of interest over the area of Biberach (Part 2).



(a) N_e ($0 \rightarrow 1$).

(b) A ($0 \rightarrow 150$).

(c) σ ($0 \rightarrow 2\ \mathrm{m}^{-2}$).

(d) γ ($0 \rightarrow 1$).

Figure 5.6: Features of interest over the area of Biberach (Part 3).

Raster data may be, for instance, Digital Surface Models computed from both first or last pulses or the normalized DSM, which is the subtraction of the DTM from the DSM [EM00; RTCK05]. Nevertheless, such features rely on the computation of advanced lidar products (DTM-DSM). Since they require interpolation or filtering methods, which is not in the framework of the proposed approach, we do not further consider these attributes.

5.3.3 Optimal neighbourhood analysis

For the search of neighbouring 3D points, parameter r can be set up empirically, according the point density, in order to gather a minimal set of points for reliable statistics and plane estimate. Another solution is its automatic tuning. On the one hand, the neighbourhood environment can be adaptative *i.e.*, may be specific to each point. Indeed, the neighbourhood will depend on the object the point lies on [FP05]. Such approach has been developed in [LUVH05] for terrestrial autonomous navigation, in order to compute the normal vectors on the largest number of points. However, the approach is focused in planar surfaces and may not be straightforward to adapt to the aerial context with our classes of interest. On the other hand, one may want to find a single value for a whole point cloud. A low value will provide a small set of points and unreliable statistics, whereas a high value will smooth the feature values, which will be highly correlated for adjacent 3D points. The analysis of the feature entropy $E_{\mathcal{V}_c}$ is one solution to find the best r between these two bounds. $E_{\mathcal{V}_c}$ is defined as follows, using eigenvalue-based features:

$$E_{\mathcal{V}_c} = -\ln(P_\lambda)^{P_\lambda} - \ln(c_\lambda)^{c_\lambda} - \ln(s_\lambda)^{s_\lambda} \tag{5.3.2}$$

where $P_\lambda = 2(e_2 - e_3)$, $c_\lambda = e_1 - e_2$, and $s_\lambda = 3e_3$ are the planarity, cylindricity, and sphericity features. Therefore, $E_{\mathcal{V}_c}$ allows to indicate whether $\mathcal{V}_c$ has one predominant behaviour (plane, line, or volume) among these three ones. A value of r for which $E_{\mathcal{V}_c}$ is low means that the current scale of analysis allows to easily describe the point local geometry. The mean value of $E_{\mathcal{V}_c}$ for all the point of the dataset is computed between a predefined lower and upper radius bounds (*e.g.*, 0.5 m→10 m, with a step of 0.1 m). Then, the optimal r value is obtained as the global minimum of this mean value. In our experiments, we have found a value of r=1.25 m for Biberach dataset, and r=1.75 m for Amiens area. More details can be found in [Dem10].

5.4 Classification strategies with feature selection

5.4.1 Classification strategies

Although theoretically considered to be insensitive to the dimensionality of the discrimination problem, SVM classifiers have proved to be influenced by the redundancy and by the irrelevance of input features. Such statements have lead towards the development of various feature selection strategies. Feature selection is the technique of selecting a subset of relevant attributes (v features among D), and may not be confused with feature extraction. This is another method for dimensionality reduction that aims to combine features also in order to obtain a relevant representation of the data in a lower dimensional space. Such method is often used for data compression, whereas, for our purpose, our goal is

first variable ranking. Moreover, the elimination of the less relevant features would also improve class separability. Both unsupervised (such as the Principal Component Analysis) and supervised (such as the Linear Discriminant Analysis) methods exists. Since training datasets are available, the selection step is performed according to the properties of the training set.
Three main feature selection strategies exist. Some of them are filter-type approaches, *i.e.*, feature selection is performed before the classification task (intrinsic discriminate power), some are wrapper-type methods *i.e.*, the process is included in the SVM classifier (which is therefore modified to choose important features as well as conduct training and testing steps) and assessment is made on its performance, whereas the **embedded** models incorporate selection into the training process, and feature relevance is obtained from the objective of the learning model. A very large body of literature has addressed the issue with both strategies, and many Feature Selection Challenges exist on various kinds of datasets [GGNZ06]. Reviews on selection algorithms may be found in [BL97; KJ97; GE03].

Three distinct classification strategies have been carried out in order to evaluate the relevance of the FW features.

- No selection: The strategy is to directly use SVM without selection. In such case, the classification accuracies achieved with the full set of features and with a subset including only multiple-pulse features (*i.e.*, discarding FW ones) will be compared.

- Filter selection: Such technique may not be optimal since there is no link between the selection step and the classifier. Intrinsic data properties are used without reference to any application. However, this is an efficient strategy since it requires only the computation of d scores (where d is the number of features) and their sorting. This is also statistically robust to overfitting. Two techniques will therefore be tested in our context: F-score and ReliefF. They are described below.

- Wrapper-based selection: SVM is directly used to conduct feature selection. Wrapper algorithms perform selection, while maintaining or even enhancing the application performance. When using classification error rates as evaluators for selection processes, it is important to note that selection algorithms must be applied inside the cross validation loop. For this kind of family, the Recursive Feature Elimination (SVM-RFE) has been adopted [GWBV02].
The SVM-RFE does not provide a variable importance metric. Contrary to filtering methods, the features will not be weighted and no comparison will be performed.

Usually SVM suffers from a large number of features. In our case, the number of features is not really significant, and we have found that a direct use of SVM works well. Nevertheless, the variable ranking provided by the three above-mentioned methods allows to sequentially feed the classifier, one feature by one, starting from the best ranked attributes. The analysis of the evolution of the classification accuracy will lead to the selection of a subset of features, providing performances similar or better to the full feature vector. This is interesting for enhancing the generalization capability of the SVM as well as for speeding up the learning process.

In practice, the available ground truth is divided into two sets. The first one is used for feature selection, and the SVM learning phase. The second subset is employed for evaluating the classification accuracy. In a 5-fold cross-validation procedure, the respective ratio are 4/5 and 1/5.

5.4.2 Feature selection methods

F-score

F-score, or Fisher score, is a technique measuring the discrimination of two sets of real numbers. $x_{k=1,..,l}$ are the training vectors, and n_+ and n_- are the number of positive and negative instances, respectively. Then, the F-score of the i$^{\text{th}}$ feature is defined as:

$$F(i) = \frac{\left(\overline{\mathbf{x}}_i^{(+)} - \overline{\mathbf{x}}_i\right)^2 + \left(\overline{\mathbf{x}}_i^{(-)} - \overline{\mathbf{x}}_i\right)^2}{\frac{1}{n_+ - 1}\sum_{k=1}^{n_+}\left(\mathbf{x}_{k,i}^{(+)} - \overline{\mathbf{x}}_i^{(+)}\right)^2 + \frac{1}{n_- - 1}\sum_{k=1}^{n_-}\left(\mathbf{x}_{k,i}^{(-)} - \overline{\mathbf{x}}_i^{(-)}\right)^2} \tag{5.4.1}$$

where $\overline{\mathbf{x}}_i$, $\overline{\mathbf{x}}_i^{(+)}$, $\overline{\mathbf{x}}_i^{(-)}$ are the average of the i$^{\text{th}}$ feature of the whole, positive, and negative datasets, respectively; $\mathbf{x}_{k,i}^{(+)}$ and $\mathbf{x}_{k,i}^{(-)}$ are the i$^{\text{th}}$ feature of the k$^{\text{th}}$ positive and negative instances, respectively. The numerator indicates the discrimination between the positive and negative sets, and the denominator indicates the one within each of the two sets.
A large value of the F-score means that this feature is more likely to be discriminative [CL06].

ReliefF

This algorithm is the improvement of the Relief algorithm presented in [KR92]. It has been adapted to the multi-class problem in [RvK03]. The key idea of ReliefF is to estimate the quality of the attributes according to how well their values distinguish between the instances that are near to each other. It does not just eliminate the redundant features but defines a relevance criterion. This is the ability of each feature to gather the instances with the same label and to discriminate those with distinct labels. The method is described in Algorithm 5.1.
The function diff(I_1, d, I_2) computes the difference between the values of the attribute d for two instances I_1 and I_2. It is defined by:

$$\text{diff}(I_1, d, I_2) = \frac{|\mathbf{x}_{I_1,d} - \mathbf{x}_{I_2,d}|}{\max(d) - \min(d)} \tag{5.4.2}$$

$\max(d)$ (respectively $\min(d)$) is the maximal value (respectively the minimal) of feature d in $\mathcal{S}$, and $\mathbf{x}_{i,d}$ is the value of $\mathbf{x}_i$ for the feature d. The function *diff* is used also for calculating the distance between instances to find the k nearest neighbours. The total distance is the sum of distances over all attributes. Moreover, the contribution for each class of the misses is weighted with the prior probability of that class $P(c)$, which is estimated from the training set (see [RvK03] for more details).
If the data from the same label have close values, and the data coming from different classes are well discriminated, the feature will exhibit a significant weight.

Algorithm 5.1: Description of the ReliefF algorithm.

Data: $\mathcal{S}$= the training set,
D= the number of features,
M= iteration times,
k= the number of nearest neighbours.

Result: $\forall\, d \in [1, D]$, weights w_d.

Initialization: $\forall\, d \in [1, D], w_d = 0$

for *i=1* **to** M **do**

1. Randomly choose a $(\mathbf{x}_i, y_i) \in \mathcal{S}$;
2. Find the k nearest neighbours with label y_i ($hits$ – H_j);
3. For each class $c \neq y_i$, find the k nearest neighbours with label c ($misses$ – $M_{j,c}$);
   for *d=1* **to** D **do**
   Update w_d such as:

$$w_d = w_d - \sum_{j=1}^{k} \frac{\text{diff}(\mathbf{x}_i, d, H_j)}{m \times k} + \sum_{c \neq y_i} \frac{P(c)}{1 - P(y_i)} \cdot \sum_{j=1}^{k} \frac{\text{diff}(\mathbf{x}_i, d, M_{j,c})}{m \times k}$$

For our experiments, we choose $k = 10$. Since it controls the locality of the estimation, it should not be set to high values. Otherwise, with increasing k values, informative features may become less and less distinguishable from useless attributes [RvK03].

The Data Mining software **Weka** [HFH+09] has been used to compute the ReliefF metric.

Recursive Feature Elimination

The selection algorithm has been initially presented in [GWBV02], and extended in [Rak03]. The goal is to find a subset of a given size v among D variables which maximizes the performance of the predictor. The method is based on a backward sequential selection. One starts with all the features, and removes one feature at a time until v features are left. At each iteration, the variable d to be removed is the one whose removal minimizes the variation of $\|w\|^2$. The cost function $C^{(d)}$ is therefore:

$$C^{(d)} = |\|\mathbf{w}\|^2 - \|\mathbf{w}^{(d)}\|^2| = \frac{1}{2}|\sum_{k,j} \alpha_k^* \alpha_j^* y_k y_j K(\mathbf{x}_k, \mathbf{x}_j) - \sum_{k,j} \alpha_k^{*(d)} \alpha_j^{*(d)} y_k y_j K^{(d)}(\mathbf{x}_k, \mathbf{x}_j)| \tag{5.4.3}$$

where $K^{(d)}$ is the matrix of the training data when variable d is removed, and $\alpha_k^{*(d)}$ is the solution of Equation 5.2.25. In practice, [GWBV02] assume $\alpha_k^{*(d)} = \alpha_k^*$ in order to reduce the computational complexity.

Algorithm 5.2: SVM Recursive Feature Elimination (SVM-RFE).

Data: $\mathcal{S}$= the training set,
D= the number of features,
d= the number of remaining features.
Result: f_v^* the feature vector sorted with respect to the feature importance.

while $d > 0$ **do**

1. Train the SVM linear classifier;
2. $\forall\ \delta \in [1, d]$, compute the $\|w^{(\delta)}\|^2$;
3. Remove the feature with the lowest $C^{(\delta)}$ value.

end

Contrary to the F-score and ReliefF strategy, the Recursive Feature Elimination method does not provide any relevance metric but a simple variable ranking. One can see this is a rather greedy algorithm.

Weka has also been employed to compute variable ranking with the SVM-RFE approach. The method implemented in Weka is restricted to the linear case (however, it has been extended to non-linear cases). Therefore, C is the only parameter to be optimized. C is arbitrarily set to 100, in order to impose no training errors.

5.5 Results

5.5.1 Some preliminary remarks

Quality measures

The confusion matrix is used to estimate the classification accuracy. This is a visualization tool widely adopted in supervised learning, where each column represents the instances in the predicted class, while each row gives the instances of the actual class. Several measures can be extracted from the confusion matrix.

- The **Overall Accuracy** (OA) is the percentage of correctly classified pixels for all the classes.

$$OA = 100 \times \frac{\sum_{i=1}^{N_c} C_{ii}}{\sum_{i,j=1}^{N_c} C_{ij}} \tag{5.5.1}$$

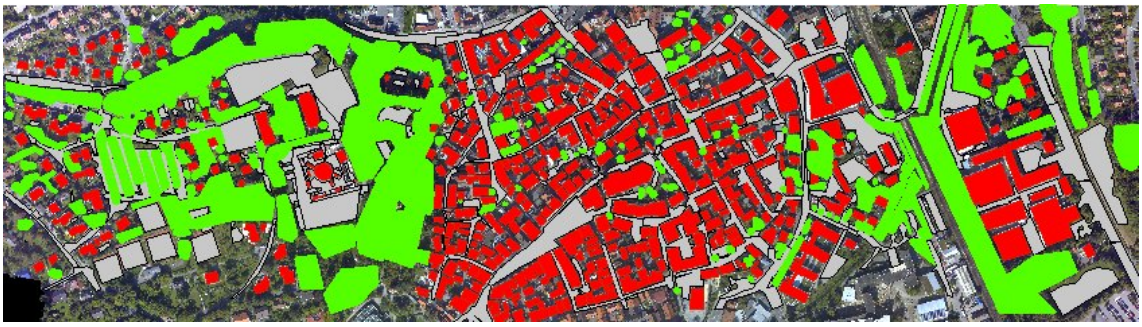
  where N_c is the number of classes (N_c=3 in our case), and C_{ij} is the number of pixels assigned to the class j by the SVM classifier, which are referenced as class i. When the OA reaches 100%, this means that the classification is almost perfect. When the OA is close to $1/N_c$ (in our case, ∼33%), this signifies the classifier randomly chooses the class for each point with the same probability. Nevertheless, when the ground truth is unbalanced, the OA may not reveal the true performance of the classifier.

In fact, even if a class with few reference points is totally misclassified, the OA value may still be very high. Indeed, its influence will be very low in the computation whether other classes with a larger number of points are correctly labelled.

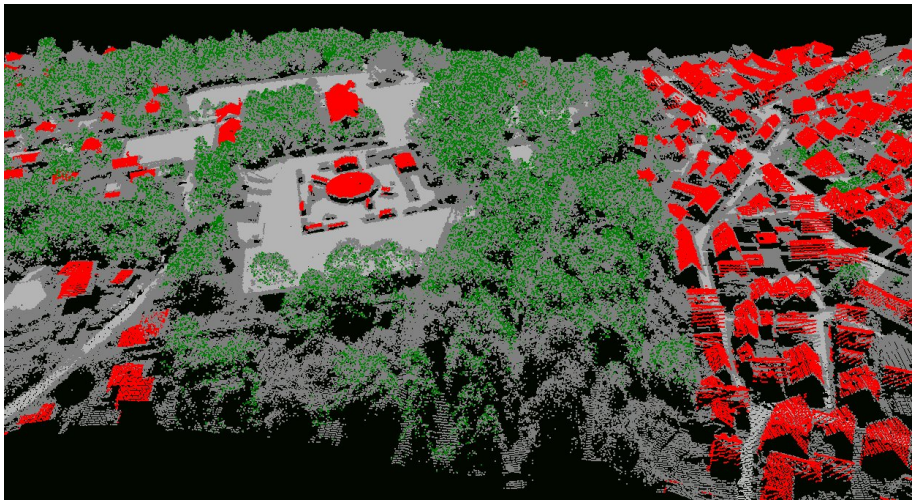
- The **Average Accuracy** (AA) is the mean of class accuracy for all the classes.

$$AA = 100 \times \frac{\sum_{i=1}^{N_c} CA_i}{N_c} \qquad \text{with} \quad CA_i = \frac{C_{ii}}{\sum_{j=1}^{N_c} C_{ij}} \tag{5.5.2}$$

CA_i is the Class Accuracy (or Producer's accuracy), and the percentage of correctly classified points for the class i.
A high difference between the OA and AA values may indicate that a high proportion of instances of a specific class is wrongly labelled.



(a) Defining the ground truth with manually selected 2D polygons (full area of Biberach).



(b) One limitation of the 2D labelling: building edges and inside tree canopy points are not classified (dark grey points).

Figure 5.7: Principle and limitations of 2D-based ground truth definition (Biberach). Classification: Building – Vegetation – Ground.

Defining a suitable ground truth

A suitable ground truth is necessary for two main reasons. First, random instances will be selected for the training step of the SVM classification. The priority is therefore to avoid mislabelled instances. Furthermore, ground truth is the basis of the computation of the quality measures. Consequently, it should be as exhaustive as possible. Achieving both objectives is not straightforward.

The manual selection of 2D polygons on an orthoimage, as presented in Figure 5.7.a, is a simple and efficient approach. 3D points lying with $\{x, y\}$ coordinates lying on such polygons will be labelled according to the polygon semantic. It works well for *ground* and *building* regions. For the *vegetation* class, we may cope with the third dimension problem by labelling only echoes that are not last echoes (*e.g.*, using the normalized number of echoes – $N_e \neq 1$). Such process is sufficient for the SVM training step, but has one main drawback. A large majority of 3D points located on the tree canopy or on the understory, as well as those lying on the building edges, are not labelled (dark grey points in Figure 5.7.b), whereas they are likely to be misclassified. As a consequence, it leads to an overestimation of the accuracy of the classification process, which is, objectively speaking, not desirable. However, this approach has been selected: approximately 1/3 of the 3D points over Biberach have been labelled.

The optimal solution would be to first provide a coarse classification, and secondly, refine it with 3D tools. This can be achieved using for instance Pointools [Poi10] or TerraSolid's TerraScan software [Soi10]. However, even with a good initialization, such task is tedious, and is not applicable at large scale, *e.g.*, for the area of Amiens (more than 67 million points).

The figures on feature selection and classification accuracies will therefore be presented only over Biberach.

The ground truth is not error-free, and the mislabelled points has therefore a negative or positive influence on the quality results, as discussed in [Car09; Foo10]. However, this should correspond to a very large minority of the manually labelled samples.

5.5.2 Feature selection

Figures 5.8, 5.9, and Table 5.1 present the feature ranking for the F-score and ReliefF metrics, and the SVM-RFE strategy respectively. The behaviours of the three selection strategies are compared in terms of classification accuracy using a forward approach. For each method, the features are introduced one by one in the SVM, starting by the most relevant one. The training step with hyperparameter tuning is first carried out, and, afterwards, the classification is performed. Results are presented in Figure 5.10.

Since we have not first worked on synthetic datasets, we are not able to evaluate the relevance of the three adopted feature selection methods as pure selection methods. However, some conclusions can be drawn:

- F-score does not reveal mutual information (*e.g.*, S_λ/e_3 in Figure 5.8);
- ReliefF favors in fact the addition of redundant characteristics (*e.g.*, $A/\gamma/\sigma^0/\sigma$ or S_λ/e_3 in Figure 5.9);

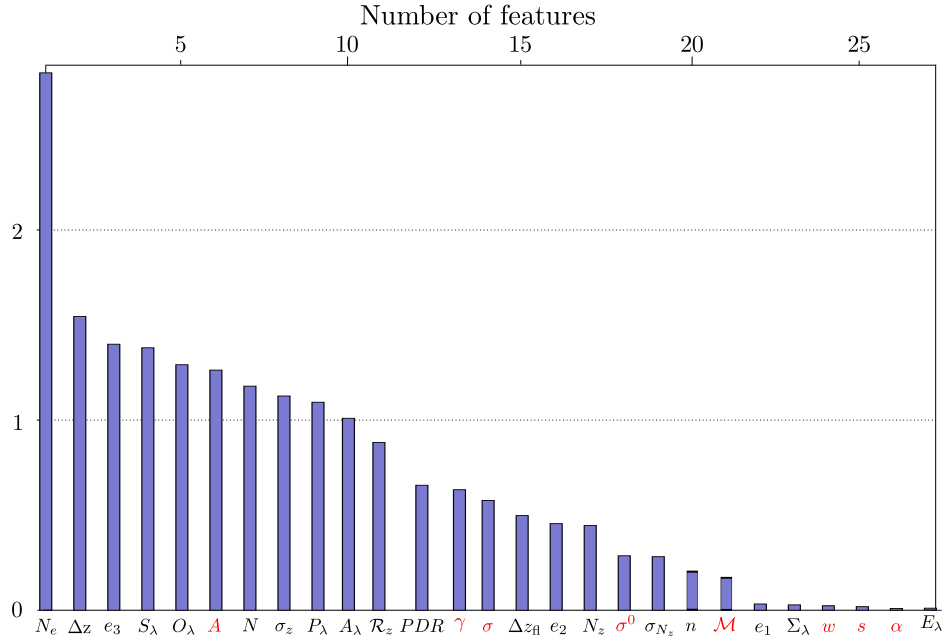


Figure 5.8: F-scores computed with 100,000 instances. The full-waveform features are displayed in red.

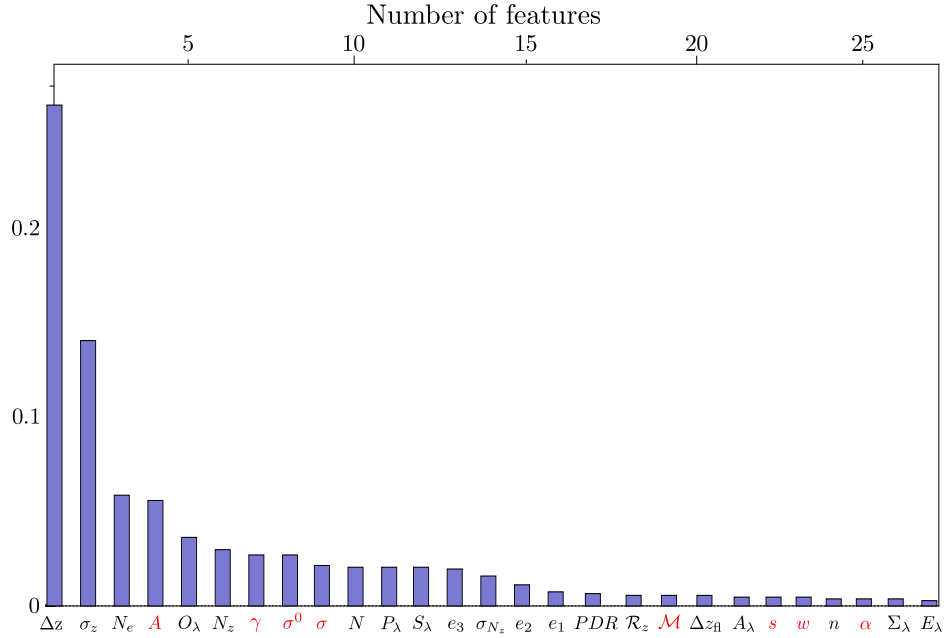


Figure 5.9: ReliefF metric computed with approximately 150,000 instances (number of neighbours equal to 10). The full-waveform features are displayed in red.

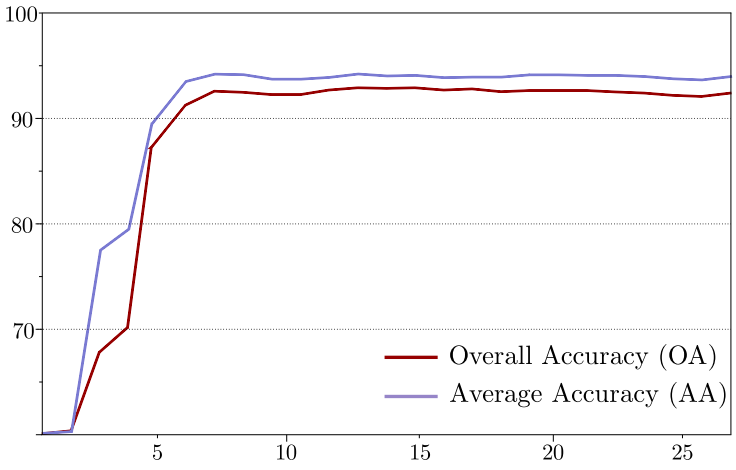
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Feature	Δz	PDR	A	n	N_e	γ	σ	σ_z	N	w	S_λ	$\mathcal{M}$	E_λ	Δz_{fl}
Rank	15	16	17	18	19	20	21	22	23	24	25	26	27	
Feature	e_3	P_λ	A_λ	O_λ	N_z	σ_{N_z}	σ^0	α	e_1	e_2	s	Σ_λ	$\mathcal{R}_z$	

Table 5.1: Variable ranking using the SVM-RFE method. The full-waveform features are displayed in red.

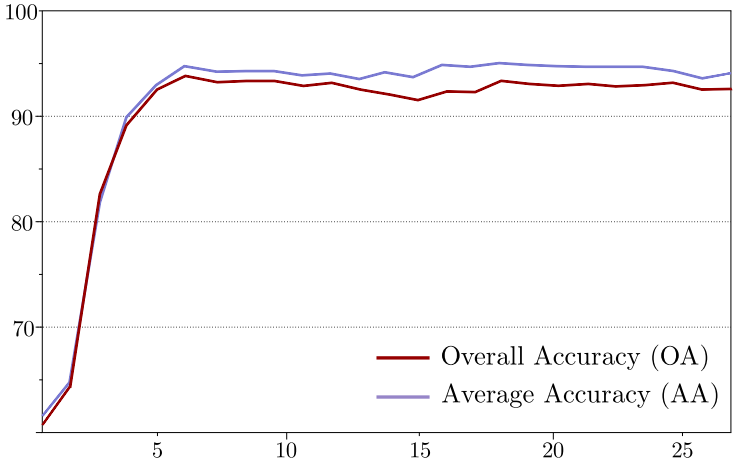
- Δz, σ_z, and N_e are always among the most relevant features: both spatial and echo-based features are discriminative. One can remind that N_e is clearly improved using FW data.
- Several attributes are always ranked among the less discriminant features (E_λ or Σ_λ). Thus, they can be removed in future works.
- Some features can behave very differently according to the selectors. For instance, the point density ratio (PDR) is ranked second, and in the second half of the set for the embedded method and the filters, respectively. Some eigenvalue-based features, such as the omnivariance, are correctly ranked for the filters, whereas for the SVM-RFE, they do not appear on the ten first attributes.
- Full-waveform features can be clearly classified in two groups. The first one corresponds to relevant features (A, γ, and σ), whereas the second one gathers non-discriminative attributes (w, s, and α). We can conclude that:
  - feature selection assesses the relevance of modelling lidar echoes by a Gaussian model;
  - radiometric calibration provides meaningful but also discriminant attributes;
  - the Generalized Gaussian is not necessary in the scope of *building*, *ground*, *vegetation* classification;
  - taking the asymmetry of echoes into account does not prevail towards improving the classification accuracy.
- The complete set of features does not give the best result (see Figure 5.10). Adding features can degrade the classification performance; this also justifies our interest for selection procedures.

5.5.3 Classification accuracy

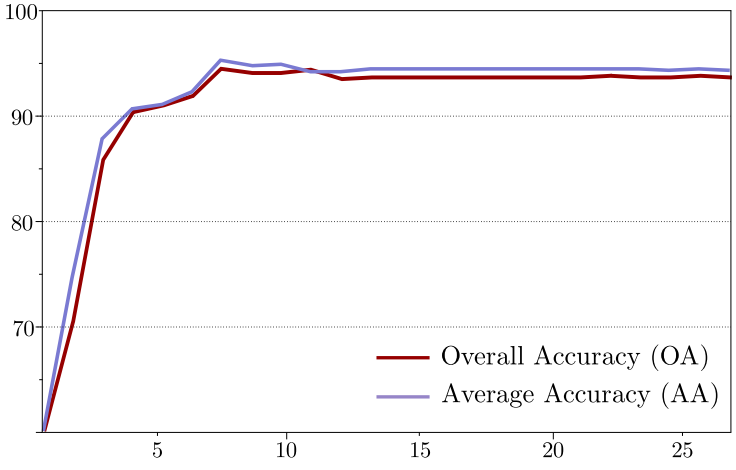
The training step has been performed, for both Biberach and Amiens datasets, using Biberach ground truth. 100 samples per class have been randomly selected. This corresponds to 0.04% of the training set, and approximately to 0.01% of the test set. Another 0.05% of the samples have been randomly chosen in the ground truth to find the optimal values of C and ϑ in the Cross-Validation procedure. For all the tests carried out, the correct classification rate for the training step oscillated between 80 and 90%. It illustrates that the



(a) F-score.



(b) ReliefF.



(c) SVM-RFE.

Figure 5.10: Evolution of the classification accuracy with the three methods: forward introduction of features, according to their ranking .

SVM classifier does not over-fit, but is able to generalize, and has been trained sufficiently.

The full feature vector introduced in the SVM classifier allows to achieve an accuracy of 94.35%. A first comparison is performed with a reduced set of 19 "multiple-pulse" attributes. The eight full-waveform features are discarded: the average accuracy reaches 91.4%. Consequently, this is another simple assessment of their beneficial impact for our classification task.

Moreover, in addition to the estimation of the relevance of features, variable ranking helps to find:

1. the subset that achieves the highest accuracy;
2. the minimal subset necessary to reach a suitable accuracy. As mentioned before, the less features you have, the less complicated the decision function will be, and the more important the generalization capability of the SVM will be. Such minimal subset can be retrieve by analyzing the evolution of the accuracy by forward selection. Figure 5.10 enhances the logarithmic behaviour of the classification performance while adding less and less discriminative attributes.

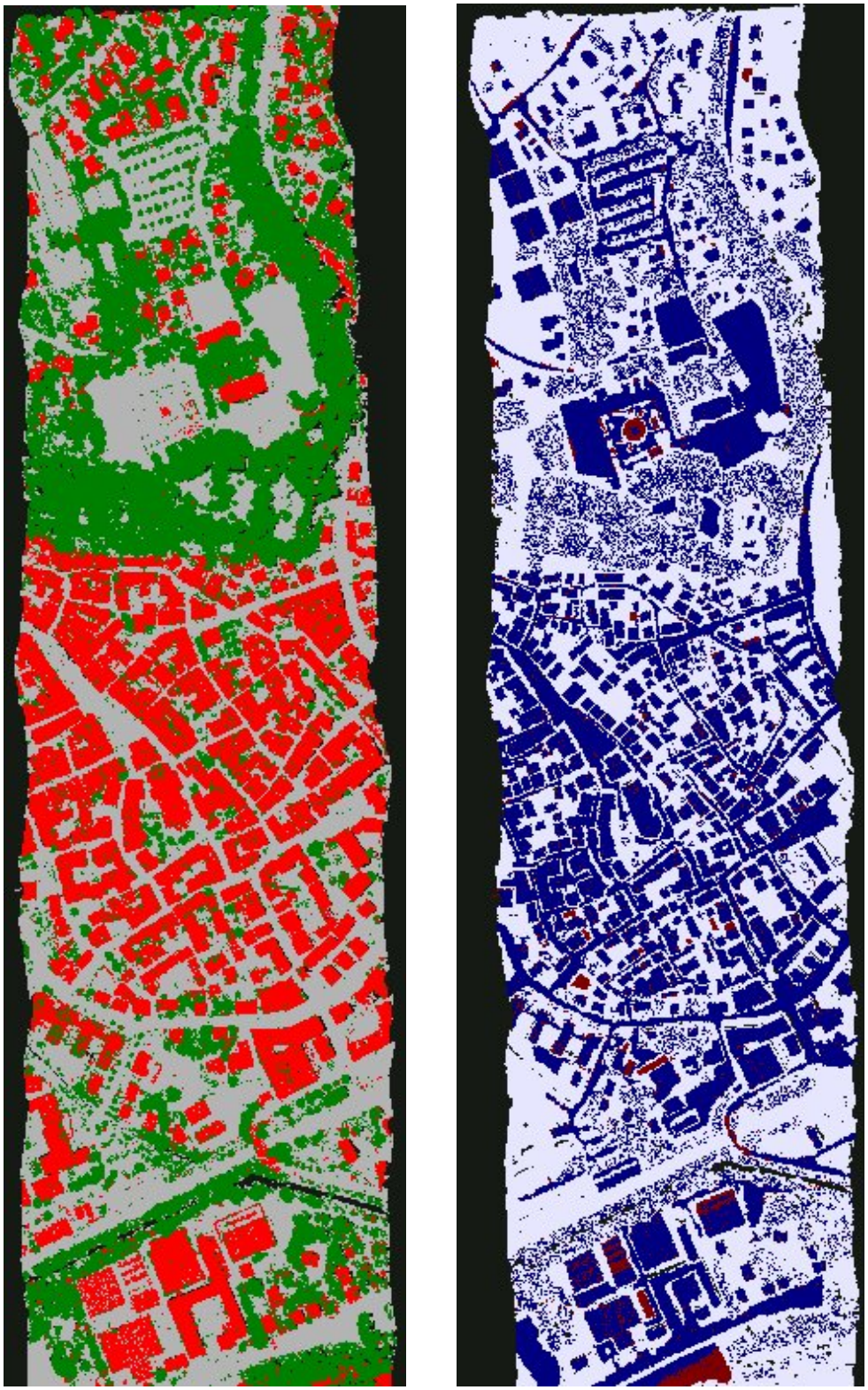
For the F-score strategy, a correct accuracy is found for seven attributes (93.8%), and the highest score is reached for the full set (Figure 5.10.a). This reveals the low performance of such feature selection method. Nevertheless, the full set of features may also be weighted by the F-score, which allows an improvement of the accuracy (94.4%).
The highest accuracy is reached for a subset of 18 features using ReliefF (94.9%), whereas suitable accuracy is found with only six features (94.6%). Besides, the SVM classifier also benefits from the weighting the variables with the ReliefF metric since a slight improvement can be noticed (94.7%).
The Recursive Feature Elimination strategy allows to achieve the best accuracy (95.3%), and to top it all, with a small subset of features (8). This enhances the performance of wrapper-based methods in relation to filtering methods. The final subset that can be conserved is:

$$\{\Delta z, PDR, A, n, N_e, \gamma, \sigma, \sigma_z\}$$

. Three full-waveform features are included in such an optimal subset, demonstrating their significant impact for solving our 3-class problem.
Table 5.2 summarizes the classification performances achieved with the various strategies that have been tested.

Figures 5.11.a and 5.12 show the results in 3D for the full area of Biberach, and two regions of interest (one residential area, and a dense urban center). The classification has been achieved with the optimal solution we have found: selection of a subset of features using the SVM-RFE, and afterwards application of SVM over this subset. Not so many errors exist, the results are globally coherent: thus, the labelling process is very satisfactory. The confusion matrix is reported in Table 5.3.

Furthermore, Figures 5.11.b and 5.13 present the comparison between our classification and the manually selected ground truth. One can see that misclassified points are found first on building edges. They are labelled as *vegetation* since such points feature similar behaviours

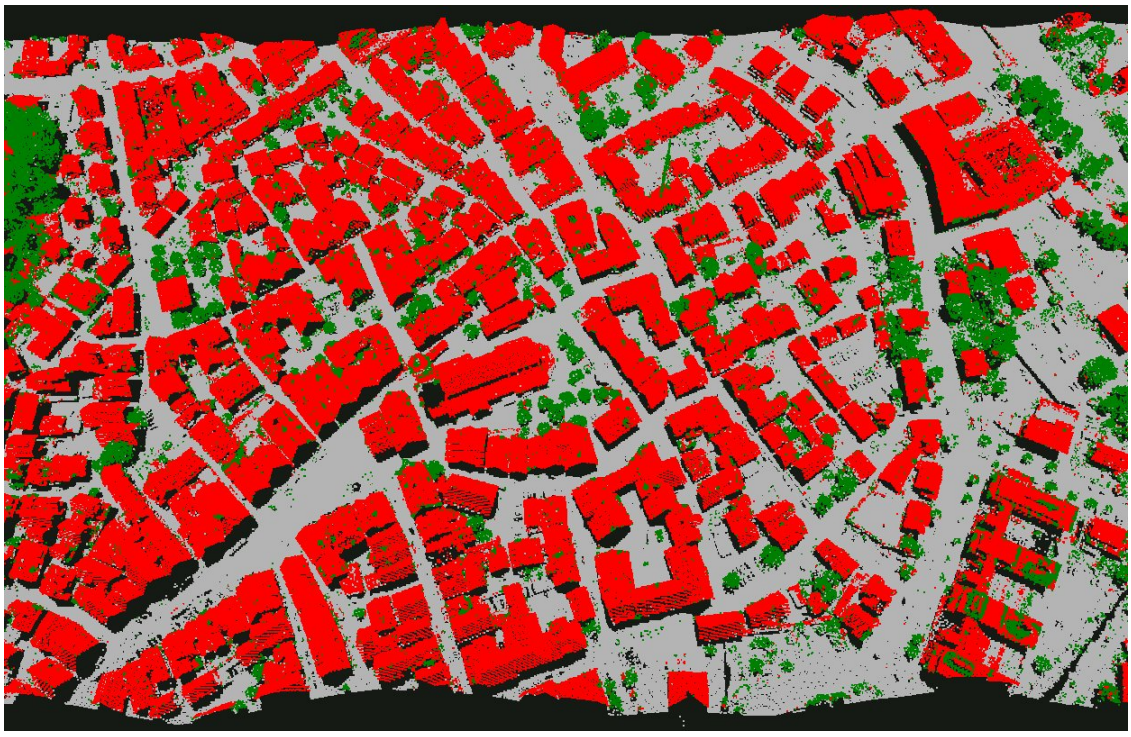


(a) Classification.

(b) Comparison with ground truth. Blue: correct – Red: incorrect – White: no ground truth.

Figure 5.11: Classification results for Biberach dataset (2.3 million points) using the RFE+subset selection+SVM strategy.

(a) Residential area.



(b) Downtown.

Figure 5.12: Focus on the classification results for two areas of Biberach.

Strategy	**Average Accuracy** (%)
Full-waveform features (8)	64.8
Geometrical + echo-based features (19)	91.4
All features (27)	94.35
F-score + subset selection (7) + SVM	93.8
F-score + weighted SVM	94.4
ReliefF+ subset selection (6) + SVM	94.6
F-ReliefF + weighted SVM	94.7
SVM-RFE+ subset selection (8) + SVM	**95.3**

Table 5.2: Average accuracies with distinct classification strategies.

Figure 5.13: Qualification of the classification over Biberach downtown. White: correct – Red: incorrect – Pink: no ground truth.

for many attributes, especially the most relevant ones: low amplitude and large echo width, equivalent point density ratio, and multiple reflections. Conversely, many erroneous points can be found on vegetated areas (even if this does not appear with respect to the ground truth). Inside canopy points are often labelled as *buildings*. This can be due to the very heterogeneous behaviour of the tree points. In case of non multiple scatterings, a point may locally appear as a building point: the echo-based features are similar in such cases, and the most important feature Δz ranges in the same value intervals. Besides, errors appear very locally on roof tops and on the ground, where points are labelled as *vegetation* and *buildings*, respectively (see Figure 5.13). They mainly correspond to low-rise objects (chimneys, cars or poles) that are not (and cannot be) taken into account in the classification process, and which do not behave as the large majority of the 3D points belonging to their class. Finally, the major source of errors comes from the labelling of whole buildings as *ground* points. Such buildings are not very elevated, and their roofs are composed of metal. Their scattering properties are very similar with streets, and since full-waveform features A, γ, and σ are included in the optimal subset, this leads to a certain confusion (see Figure 3.17.c).

Eventually, the classification of ground points is very satisfactory, and our proposed approach is valuable as an alternative method for performed lidar point filtering in urban areas, *e.g.*, for Digital Terrain Model generation. The minor errors, as well as those on building roofs, may be easily corrected with a regularization step. The main remaining issue is the misclassification of full building roofs, that also can be solved using some contextual knowledge at the object level.

# ground truth points	Class	Building	Ground	Vegetation
275,058	Building	**94.4**	3.8	1.8
333,506	Ground	1.7	**98.1**	0.2
157,281	Vegetation	5.0	0.1	**94.9**

Table 5.3: Confusion matrix for the strategy RFE+ subset selection+SVM. $AA = 95.3\%$ with 765,845 points.

"Historical" evolution of the accucary

Finding the most relevant features among an unordered set of attributes is the first and most objective way to assess the contribution of full-waveform data. Furthermore, such assessment can also be carried out by adding progressively the attributes within the feature set, and performing the classification with the available set. We have decided to insert them according to their "historical order of appearance".

- We first start with all the attributes that can be computed with the simple knowledge on the $\{x, y, z\}$ triplet (*i.e.*, a simple 3D point cloud which corresponds to 14 features – scenario 1). The echo number and the amplitude are then successively added (five new features – scenario 2, and one for scenario 3, respectively), enhancing the development of multiple sensors providing an amplitude value. This first group of 20 features does not need any waveform recording process.
- Later, with the development of full-waveform sensors, the Gaussian decomposition method has provided another feature (namely the echo width), as well as those com-

puted for the calibration procedure (three other features – scenario 4). Afterwards, our work has allowed to extract the echo shape (scenario 5), and jointly the best-fit model and the peak asymmetry (2 features – scenario 6).

For each of these six scenarii, a SVM classification is performed without any feature selection step. The evolution of the classification Average Accuracy is reported in Figure 5.14. Both multiple-pulse (MP) and full-waveform data have been tested. One can note that MP data only allows to reach scenario 3. MP data correspond to the point cloud that is provided with the full-waveform datasets.
When comparing MP and FW data, one can first note that the simple knowledge of the 3D position of the points leads to slightly better results with MP data than with FW data. This can be explained by the fact that the waveform processing step allows to retrieve points in complex locations, such as building edges, facades, or inside the tree canopies. But, as reported in the previous sections, such points typically correspond to confusing areas, where the classification procedure may fail. This explains why the results are getting worse. However, the results are reversed with the addition of the amplitude feature: the classification process clearly benefits from its correction procedure, with the knowledge of the survey metadata and the emitted pulses.
From scenario 1 to scenario 4, *i.e.*, until the waveform Gaussian decomposition process, one can note in Figure 5.14 a constant improvement of the AA (from 85.2% to 94.75%). Then, the addition of the shape feature α only allows to reach an AA of 94.9%. It shows that the Generalized Gaussian model is at present time more relevant for improving the waveform reconstruction. Finally, the same conclusion can be drawn with the scenario 6, since the knowledge of the asymmetry of the echoes provide worse results (94.9%→94.35%). However, this can only be stated for our 3-class problem, using a standard SVM classifier, with our limited ground truth.

5.6 An alternative approach: the Random Forests

5.6.1 Motivation

In parallel to our work, another approach dealing with full-waveform lidar data classification in urban areas has been carried out in conjunction with the EGID Institute/GHYMAC lab of the University Bordeaux III. In particular, this is part of the thesis work of Li Guo. More details on this work can be found in [CGM09; GCMB10]. The relevance of full-waveform features is also assessed with a feature selection step. Consequently, we have decided to briefly present the strategy and related results, in order to enlighten similarities with the conclusions that can be drawn with our approach.
Two distinct strategies have been adopted in Li Guo's work: relevance of FW data versus various multiple pulse (MP) lidar features, and versus MP features and Red-Green-Blue channels of an aerial orthoimage. Since our thesis is focused on lidar data, we only present the first strategy.
In comparison with our SVM-based approach, there are four main distinctions:

- four classes are selected: *building*, *vegetation*, *artificial ground*, and *natural ground*;
- an image-based approach has been adopted (2D classification);
- the lidar feature set is not exactly the same;

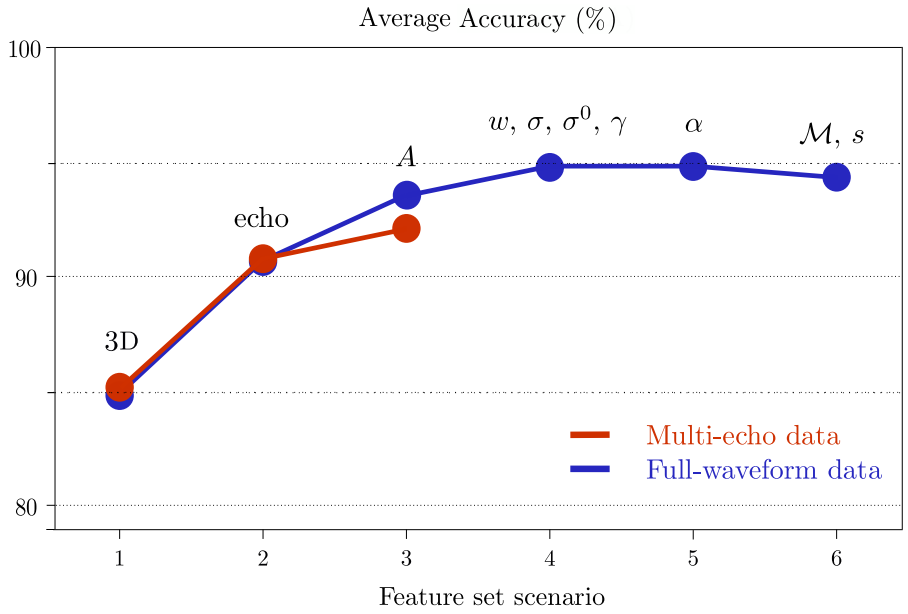


Figure 5.14: Evolution of the Average Accuracy of the classification procedure according to the progressive insertion of the 27 features of interest. The "historical" order of appearance has been selected for the introduction order. The two plots correspond to the two distinct kinds of lidar data than have been tested (multiple pulse versus full-waveform data).

- the Random Forests classifier is selected (the supervised approach is kept).

In this case, the 3D point cloud is interpolated in 2D. The output of such process is therefore a 2D land-cover classification. It allows to use afterwards 2D based applications, and may be a valuable input for subsequent 3D classification, especially in urban areas. Consequently, this solution has been adopted.
For each pixel, the lidar features are computed using the 3D points included in a given cylindrical neighborhood (see Figure 5.3). The raster cell spacing and the cylinder radius r are chosen with respect to the 3D point density, and the contrast between objects we aim to retrieve. Interpolation has one main limitation: edges between classes are smoothed. Indeed, with a cylindric environment, one pixel that fails at the border between classes will include mixed information from these two classes, and its labelling may be biased. However, we can cope with this issue by applying more advanced interpolation processes involving adaptative neighbourhoods, but this is not the topic of this dissertation (see [Dem10]). The impact of the interpolation process is further discussed in [CGM09].

5.6.2 Strategy

The input of the algorithm is again a feature vector f_{v2}. The Random Forests classifier is applied on it, and provides a feature importance measure. Such approach has been tested on the Biberach dataset with 21 features.

Random Forests

Random Forests (RF) is a variant of bagging proposed by Breiman [Bre01]. It is a decision tree based ensemble classifier that can achieve a classification accuracy comparable to SVMs [Zhu08a]. It belongs to the family of the multiple classifiers *i.e.*, several classifiers are trained and their results combined through a voting process. It has been successfully applied to multispectral data [Pal05], multitemporal SAR images [WB09], hyperspectral data [HCCG05], or multi-source data (spatial images and topographic date in [GBS06] or SAR and multispectral images in [WB07]).

RF are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. In training, the algorithm creates T multiple bootstrapped samples of the original training data, then builds a number of no pruned Classification and Regression Trees (CART) from each bootstrapped samples set. Only a randomly selected subset of the input features is considered to split each node of CART. The feature that minimizes the Gini impurity is used for the split [Bre01]. For classification, each tree gives a unit vote for the most popular class at each input instance. The final label is determined by a majority vote of all trees. The RF classifier has two parameters: the number of trees T, and the number of variables M randomly chosen at each split.
When the training set for a particular tree is drawn by sampling with replacement, about one-third of the cases are left out of the sample set. These samples are called *Out-of-Bag* (OOB) data, and are used to estimate the features importance as detailed hereby. More details can be found in [GCMB10].

Feature importance measure

Aside from classification, Random Forests provide measures of variable importance based on the permutation importance measure. It was shown that it is a more reliable indicator than the mean Gini importance [SBAT07]. The importance of the variable f can be estimated by randomly permuting all the values of the f^{th} variable in the OOB samples for each classifier. The measure of feature importance is the difference between prediction accuracy (*i.e.*, the number of observations correctly classified) before and after permuting feature f, averaged over all the trees. A high prediction accuracy decrease denotes the importance of that feature. An increasing OOB error indicates the importance of that variable [GBS06].

The feature vector

21 lidar features, both spatial and FW, are selected. It has been designed in our early works to compare lidar features. Most of the features (19) are the same than those selected for Support Vector Machines classification (see Section 5.3). The two other ones are described below. The full-waveform features are those extracted from the Generalized Gaussian decomposition approach described in Chapter 3. We have:

$$f_{v2} = \{\Delta z\ \Delta z_{\text{fl}}\ \sigma_z^2\ \mathcal{C};\ \ \lambda_1\ \lambda_2\ \lambda_3\ A_\lambda\ P_\lambda\ S_\lambda\ L_\lambda;\ \ N_z\ \sigma_{N_z}^2\ \mathcal{R}_z\ D_\Pi;\ \ N\ N_e;\ \ A\ w\ \sigma\ \alpha\} \quad (5.6.1)$$

$\mathcal{C}$ is the local curvature. This is the maximum value of the altimetric gradient differences,

which are computed in four main directions. Please see [SV01; Tóv06] for more details. D_Π is the distance from the current point to the local estimated plane Π.

5.6.3 Results and discussions

The ground truth is directly derived from the 3D ground truth of Biberach, described in Section 5.5. The feature selection was run with $M = 4$ and $T = 1000$ trees.

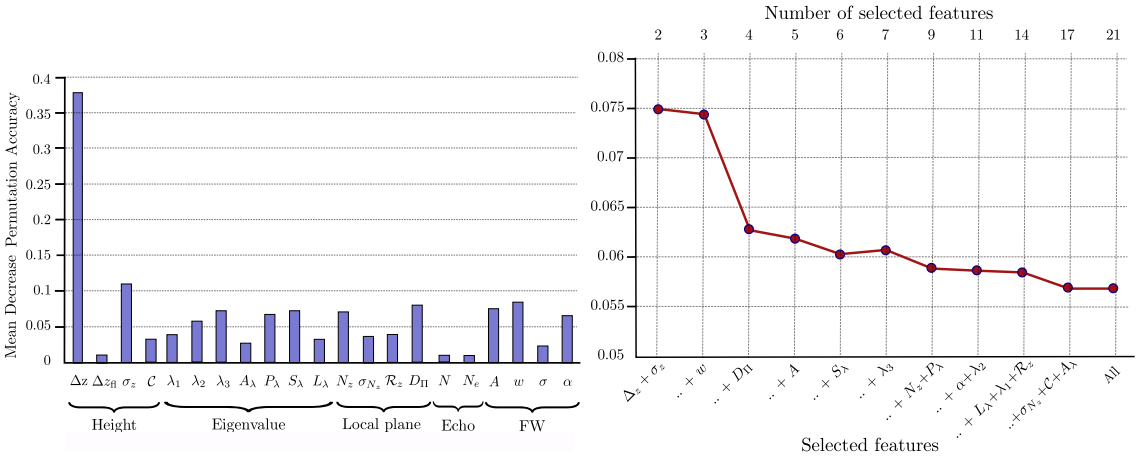
Feature importance and selection

The most important features are height-based: the height difference, and the height variance. Echo-based features are not important when using other attributes that describe more accurately the local distribution of 3D points, such as eigenvalue-based or 3D plane-based features. Moreover, the first-last height difference is not important to classify urban objects since it is used with the height variance. In fact, both variables can be correlated and the latter has more values which allows to distinguish rooftops building and ground for instance. λ_3 is the most important eigenvalue, and the sphericity S_λ shows a high importance, whereas the correlated anisotropy shows a lower one. This illustrates the advantage of permutation accuracy measure since redundant features should be less important. Among 3D-plane based features, the distance to plane D_Π seems to be the most important one. Finally, for full-waveform features, echo amplitude, and width are the most important for all classes. The FW cross-section σ is less important as it is correlated to the former features.

Figure 5.15.b is obtained by the backward iterative elimination of features using OOB errors. The graph is shown in a forward way to illustrate the more relevant features. As mentioned previously, the selection resolution increases when the number of features becomes smaller. The smallest error rate returns a feature vector of 17 attributes, where Δz_{fl} , σ, N and N_e are eliminated. However, we can also keep the smallest set of features that returns a low error rate which corresponds to $\{\Delta_z, \sigma_z, w, D_\Pi, A, S_\lambda\}$. One can observe that four feature groups are represented: the height based group is the most important one, then two FW features are selected, which confirms the contribution of full-waveform lidar data for urban scene classification.

Classification results

The Random Forests classifier is run with the 17 best selected features. The confusion matrix is given in Table 5.4. The training dataset is highly imbalanced with two major classes (building and artificial ground), that are more than 10 times larger than vegetation and natural ground classes. We can notice that artificial ground and buildings are well classified. However, the algorithm has more difficulties in classifying natural ground and vegetation which suffer from smaller training sets. Errors essentially occur between building and artificial ground since (1) these two classes have similar colors on optical images, and the presence of shadows increases these errors; (2) lidar features are ambiguous for building facades that are transitions between both classes. In addition, vegetation classes can be confused with artificial ground classes as the laser pulse can reach the ground under sparse vegetation. This confusion is increased by the 2D lidar data interpolation.



(a) Variable importance.

(b) Iterative elimination feature selection.

Figure 5.15: Feature importance and selection for the entire lidar feature vector f_{v2}. The vertical axis corresponds to the Mean Decrease Permutation Accuracy.

Finally, classification errors may occur between natural and artificial grounds since MP lidar features do not allow to discriminate them.

# pixels		Building	Vegetation	Artificial Ground	Natural Ground
188015	Building	**92.3**	0.3	7.4	~ 0
15723	Vegetation	3.6	**79.8**	16.6	0
192945	Artificial Ground	1.7	0.6	**97.7**	~ 0
2149	Natural Ground	0.9	0.3	23.3	**75.5**

Table 5.4: Confusion matrix for test data using a RF classifier with 17 best features of f_{v2}, 60 trees, and 4 split variables. $AA = 94.35\%$.

5.6.4 Outlook

Although performed in 2D, another feature selection approach embedded in the Random Forests classifier, similar to the SVM-RFE, has assessed the relevance of full-waveform lidar data for land-cover classification in urban areas. In addition, it has also allowed to find the optimal feature selection, achieving the highest classification accuracy.
The SVM and RF lidar feature sets are not exactly the same but almost similar conclusions can be drawn: spatial features are important, especially the height above the ground, but the echo amplitude estimated from the approach described in Chapter 3 has also a significant impact on the discrimination between classes. Finally, as for the SVM-based classification, misclassified pixels can be found on the border between classes (building edges, limits of vegetated areas), and punctually on low-rise ground objects (cars, poles etc.). A regularization step would also be beneficial for improving the global accuracy.

5.7 Conclusions and perspectives

5.7.1 Conclusions

This chapter has addressed the issue of the classification of 3D point clouds in urban areas, using full-waveform features derived from a previous modelling step. The three classes of interest were *building*, *ground*, and *vegetation*. Even if choosing more classes would have been possible, as presented in Section 5.6, we have considered the segmentation of roof materials, ground surfaces, and tree species as a subsequent step of this 3-class problem. Our land-cover classification has been designed to be "sufficient" for many applications: Digital Terrain Model generation with ground points, focusing step for building reconstruction or tree modelling.
The FW features are extracted from the two distinct approaches described in Chapters 3 and 4. The supervised Support Vector Machines classifier has been adopted because it performs well for large datasets, it has a good generalization performance, and can be fed with a simple feature vector. Such feature vector is computed for each point. It gathers the FW attributes but also features that only rely on the 3D coordinates of the point, its local neighborhood and its echo number. They can be therefore computed using only multiple pulse data. However, as extensively demonstrated in Part II, such spatial features benefit from the off-board waveform processing chain. 27 features are finally introduced in the SVM, which allows to reach a classification accuracy close to 95% for the three classes of interest, with very few training samples. This is very satisfactory. The mislabelled points correspond mostly to building edges and vegetated areas. This can be explained by the fact that such regions have complex behaviours, which are all the more difficult to learn for the SVM than such points are barely included in the ground truth (and therefore in the learning step). Errors are also noticed for flat low-rise buildings exhibiting radiometric properties similar to the ground.

In order to objectively assess the relevance of FW data for land-cover analysis, the supervised classifier has been coupled with a feature selection step. Two distinct strategies have been chosen: variable ranking before classification, and selection based on the decision taken by the SVM (SVM-RFE). For variable ranking out of the classifier, two methods have been adopted (F-score and ReliefF metrics). The performance of the these three methods have not been evaluated nor compared. Nevertheless, common trends emerge: (1) several full-waveform features, namely the amplitude, the cross-section, and the backscatter coefficient are among the most discriminative features; (2) the shape and the skewness attributes are among the less discriminative ones.
Moreover, such a variable ranking has two other advantages: firstly, it may permit to weight the features in the SVM classifier, which improves the accuracy; secondly, the analysis of the evolution of the accuracy with forward selection enables to find the minimal feature set allowing to achieve suitable accuracy. All of the three reduced feature set include the point amplitude.

Eventually, the main positive point is that some full-waveform features are very relevant for 3D classification. This is true in 2D as well. An image-based strategy based on Random Forests have also been applied on FW lidar data with almost the same set of features. Similar conclusions can be drawn.
The main negative point comes from the relatively low contribution of the model feature $\mathcal{M}$

and the poor ranking of the echo shape (extracted from the Generalized Gaussian modelling of Chapter 3), and the skewness attribute (computed from the results of Chapter 4). This may lead us to conclude that for the scope of classification, the Gaussian model is sufficient to provide additional useful and discriminative information.

Contrary to waveform decomposition and modelling, there is no classification method with a unique set of features that has been adopted by research teams as a standard approach. However, we have tried to design a rather objective workflow for full-waveform data contribution evaluation. This has motivated the adoption of SVM, instead for instance, of Decision Trees, that, however, represent a simple and efficient method. The limitation of the problem to three classes that can be both discriminated with MP and FW data has also "facilitated" the comparison: the additional introduction of various kinds of ground or roof materials, difficult to classify with spatial lidar features, would have artificially demonstrated the relevance of radiometric properties extracted from the waveform modelling step. Finally, we have proceeded to a very selective choice of spatial features that we also tried to compute as efficiently as possible. The good discriminative behaviour of FW data in relation to such an optimal MP feature set also demonstrates their relevance in our context.

5.7.2 Perspectives

We have only initiated land-cover classification works in urban areas, and many perspectives are conceivable.
Firstly, only point-wise labelling has been performed. Adding contextual information would indubitably improve the results. Indeed, even if several spatial features are currently computed using a local neighborhood analysis, many isolated points are misclassified. Regularization algorithms (*e.g.*, Markov Random Fields, Graphical Models, or Graph Cut-based methods) are suited to cope with such a problem.
Secondly, when dealing with medium scale classification, labelling small urban objects is not interesting: street items, such as cars, traffic lights, poles etc. have been put with the *ground* class, however they do not feature similar behaviours. Therefore, the introduction of a "rejection class" should cope with such objects, that are likely to be misclassified. In such a case, one has to pay attention to the learning step, especially if performed in 2D using an orthoimage. Diachronism may exist, and training areas for the **ground** class should be carefully checked to ensure valid training samples.
To deal with both issues, an interesting perspective would be to rely on the estimate, for each point, of the probabilities to belong to each of the classes of interest. Such probabilities can be computed using one of the approaches described in [WLW04], and which are improvements of the seminal work presented in [HT98]. 3D points with a high probability in one class may have a strong influence on its neighbours, whereas an instance without peremptory decision would be more prone to be reclassified.

We have restricted our problem to three labels, arguing that it was the fairest strategy to evaluate the impact of FW data with respect to multiple pulse data. Once such a coarse classification achieved, it would be relevant to test whether full-waveform features may help to discriminate ground surfaces, roof materials or tree species, without necessarily involving spatial features.

Furthermore, as enlightened in Section 5.5.1, the performance of the learning stage of a supervised classifier heavily relies on a correct manual ground truth. However, the acquisition of such a ground truth is a time-consuming and tedious task, often 2D-based. Besides the existence of misclassified training samples should not be neglected. Consequently, in order to reduce this task, and to learn using simultaneously a small training set with a large set of unlabelled instances, a large variety of semi-supervised methods have been developed [CSZ06; Zhu08b]. In our context, it would be interesting to test, for example, transductive Support Vector Machines [Vap98]. Moreover, to encode some knowledge on geometry and data distribution, an unsupervised classification should be performed first to understand how the data set can be separated.

Eventually, we have presented in this chapter two distinct classification methods. They have been carried out in parallel, but a merge would also be conceivable: decision fusion may use the outputs of both classifiers, whereas a SVM ensemble strategy may embed SVM and multiple classifier systems directly at the decision level [WvdLB+10].

Part IV

Conclusions

Chapter 6

Conclusions and perspectives

6.1 Conclusions

6.1.1 Summary

Full-waveform lidar systems can provide a more in-depth description of ground topography than standard multiple pulse systems.
The analysis of small-footprint data over urban areas was addressed in this thesis. The objective was to propose a methodology to process lidar waveforms in order to carry out land-cover classification. Since systems with waveform digitizers have very recently appeared, and its potential for urban classification is barely evaluated, the goal was to assess at the same time the genuine contribution of full-waveform data, in relation to multiple pulse systems.
The proposed methodology was composed of two phases. The first one consisted in a signal processing task. From the raw waveforms, a sequence of ranges values was derived, as well as shape features revealing the particular morphology of each echo composing the waveform. The output of the first step was a 3D point cloud, coupled with these shape attributes, named *full-waveform* features. The second phase tackled the issue of classifying the resulting point cloud into *ground*, *vegetation*, and *building* labels.

Two general strategies have emerged from the waveform processing step:

1. The first methodology consists in dividing the analysis of the waveforms into two stages: decomposition and modelling, *i.e.*, echo detection and fitting. Both steps are commonly tackled in the literature, and we only proposed slight variations to the widely adopted workflow. Each peak detected in the waveform is adjusted by the same parametric function, the Gaussian model, using a Least-Squares technique. Two parameters are estimated, namely the amplitude and the width.
   Our contribution in this topic concerned the use of an iterative procedure for echo finding in case of complex waveforms, and the introduction of the Generalized Gaussian function, which is an extension, still symmetric, of the Gaussian curve.
   This lead to the detection of a higher number of echoes compared to multiple pulse data, and improvements in the fitting procedure. Flattened and peaked echoes are better handled, the echo amplitude and width are better estimated, and to top it all,

a new feature, namely the echo shape, was from now on provided.
The benefit of such an advanced echo detection method can be assessed whatever the subsequent application. The position of previously undetected objects was retrieved. Closely located targets were discriminated. This should be helpful for vegetation characterization, and any pattern recognition problem involving 3D lidar point clouds. The performance of our proposed detection algorithm (as well as any other "off-line" methods) is now ruled by the waveform digitizing sampling rate and the characteristics of the transmitted pulse.
Full-waveform features were analyzed in the scope of land-cover classification. Echo amplitude and width have revealed high potential for the discrimination of our three classes of interest (as well as the cross-section, derived from these two ones), conversely to the shape feature. The latter one only demonstrated that the fitted echoes are not perfect Gaussian curves. Various kinds of ground surfaces and roof materials were finally selected: the amplitude and the cross-section may be discriminant for their segmentation. However, more in-depth analysis is still necessary to really draw such conclusions.

2. The Generalized Gaussian decomposition enhanced the fact that some waveforms, even those corresponding to flat targets, exhibited significant residuals with the estimated signals. A close look to the waveforms revealed the asymmetry of many peaks. Consequently, an innovative strategy has been developed to cope with waveforms composed of echoes with distinct behaviours. We therefore aimed to find, for each peak of each waveform, the most appropriate parametric function among a given library. The stochastic framework of marked point process has been adopted due to its ability to deal with echo detection, parameter estimation, and model selection jointly in a single paradigm. Three functions, including the Generalized Gaussian model, have been selected.
Experiments on simulated data, large-footprint and small-footprint waveforms shown the performance and versatility of this new method. This is due to the introduction of a priori physical knowledge on the fitted waveforms.
Despite a significant computing time, it yielded satisfactory results in term of echo detection (similar to the first strategy), and fitting accuracy (slightly better). For small-footprint data in urban areas, statistics showed a very large majority of echoes were fitted with the Generalized Gaussian function, validating the standard Gaussian assumption. Asymmetric peaks are found mainly on building edges. For vegetated areas, all three models are identically selected: we can conclude symmetric models are sufficient.

Subsequent to the processing of the waveforms, the second main part of the thesis dealt with the classification of the derived 3D point cloud.
Both spatial and newly estimated full-waveform features were considered as discriminant attributes for our purpose. They were introduced into a supervised Support Vector Machines classifier. Besides, since we aimed to assess the contribution of the FW technology in relation to multiple pulse system, a special attention has been focused on feature selection. Selection have been performed before the classification task as well as embedded into it. Both strategies revealed that the amplitude and the cross-section features are among the most relevant for discriminating between *ground*, *vegetation*, and *building* classes. Conversely, the echo model and skewness, extracted from the stochastic approach did not prove to be very useful. Eventually, the embedded strategy yielded more satisfactory results in

terms of accuracy than the filter and the "full feature vector" approaches, and in terms of optimal subset selection.
Another classification strategy, based on Random Forests, involving also echo amplitude and width showed the relevance of full-waveform features, for urban areas.

Some final conclusions can be drawn:

- For waveform decomposition and modelling:
  - With 1ns digitization sampling and transmitted pulses of 4/5ns, slight improvements of standard echo detection techniques suffice to retrieve all the necessary peaks;
  - A symmetric function is enough for echo modelling;
  - In urban areas, the standard Gaussian model should always be adopted since its parameters are discriminant for many objects and surfaces of interest.
  - Superior geometric information is provided by full-waveform sensors;
  - The stochastic approach is suited for testing a larger variety of models;
  - The stochastic approach is better suited for adjusting medium and large footprint data.
- For 3D classification:
  - The SVM classifier performs very well, even with small training sets;
  - A feature selection (if possible embedded in the classifier) is necessary to improve the classification accuracy and select a small set of relevant attributes;
  - The feature selection step demonstrates the relevance of the amplitude feature and the backscatter coefficient in urban areas;
  - The two previous conclusions are not limited to Support Vector Machines;
  - Since the amplitude exhibits highly discriminative information, calibration techniques have to be set up to be able to carry out classification over several distinct areas.

6.1.2 Contributions and limitations

We review here the contributions of the proposed approach and highlight its main limitations, which leads to mention some of its possible extensions in the next Section.

Contributions

▶ **Complete workflow**: we proposed a complete workflow for analyzing small-footprint lidar waveform in urban areas, starting from raw 1D signals, and finishing with a 3D classification. It allowed us not to move away from the standard products of multiple-pulse systems, the 3D point clouds, and therefore perform more easily comparisons at several levels: how many additional 3D points are we able to find ? what is the contribution of the newly extracted morphological parameters for classification purposes ?

► **Versatile approach**: our approach is designed to be flexible. For one or several modelling functions, one may change it (them) easily, and assess its impact in terms of classification accuracy, without modifying any step in the workflow. Furthermore, however focused on urban areas, the approach can be adapted to other landscapes: one has just to train the SVM classifier with other classes of interest.

► **Full automation**: simple knowledge on lidar sensor specifications are sufficient to set up the parameters of the two processing methods. Furthermore, SVM hyperparameters are automatically selected as well as the most discriminative features.

► **Ability to process a large variety of signals**: out of the scope of land-cover classification, the proposed marked point process approach has revealed a high versatility for handling lidar waveforms. It can be reused to analyze other kinds of remote sensing signals.

► **Accurate classification**: satisfactory labelling results have been reached in terms of accuracy. However noisy the final results are, we prevent the use of a filtering step that may be not be always transferable to other landscapes.

► **Objective assessment of the contribution of FW data**: the feature selection step coupled with a supervised classifier has shown that FW features, especially the amplitude, are relevant for classification purposes.

► **New tools**: the development of such a workflow has eventually allowed us to create several tools for handling, processing, and visualizing full-waveform lidar data. These tools are not specific to urban areas, and may serve as a basis for future developments and research on the subject.

Limitations

► **No classification regularization**: no post-processing step has been developed in order to cope with isolated misclassified points. Adding contextual information *i.e.*, reaching the object-based level would finalize this part, and improve the classification accuracy.

► **No integration of waveform similarity**: each waveform, as well as each 3D point, is processed separately. Taking into account the closeness of the shapes of successive waveforms, or the dissimilarity between them, has been left out.

6.2 Perspectives

We first consider that the development of fully new methodologies for waveform decomposition and modelling is not necessary. Simple peak detection methods are suited, and even modifications in the sensor specifications (higher digitizing rate and lower transmitted pulse width) will not drastically change the problem. Concerning echo modelling, it may be of interest to test other symmetric parametric functions, one by one, in the perspective of land-cover classification. For that purpose, our proposed framework is still valid since it does not require any modification (except in the learning stage of the SVMs).
One potential improvement of both waveform processing methods would be to perform fitting regularization between consecutive waveforms. For instance, the knowledge of the overlapping ratio between two successive footprints might be an interesting a priori knowledge on how similar these waveforms are likely to be.

From a classification point of view, many improvements are conceivable. Some of them directly stem from the proposed workflow. They have been presented in Section 5.7.
Full-waveform classification is always considered from a traditional point of view. Standard ALS products are point clouds, that is why almost all the research teams first perform waveform processing in order to generate a 3D point cloud. However, it should be interesting to directly classify the 1D signals. The advantages would be threefold: 1/ no modelling function is required; 2/ new classes can be introduced (*e.g.*, building edge, dense vegetation, open vegetation etc. depending for instance on the presence of ground returns and the complexity of the waveforms); 3/ and finally 3D labelling can still be derived after peak detection. The classification of sequences that are themselves composed of smaller sequences exist in many domains, such as bioinformatics or speech analysis. It would be interesting to assess whether an adaptation of existing approaches is suited in our context. Moreover, classification of FW data can be performed separately on the point cloud (in 3D), on the waveforms (in "1D"), and in 2D (image-based approach). Subsequently to the three labelling processes, a challenging task would be to fuse the decisions. In case of final point-based classification, the 3D approach would benefit from the spatial regularity of the 2D process, and the topological context of the 1D classification. If one aims to derive 2D thematic maps, image-based classification can take advantage of 1D and 3D labels to enrich the existing classes (*e.g.*, different kinds of vegetation).

The above-mentioned perspectives are not specific to urban areas. Nevertheless, many other interesting works can be carried out on these ones.
As mentioned in Section 2.6.6, developing pattern recognition algorithms directly on the 3D volume avoids to perform any peak detection, and allows to directly benefit from a spatio-temporal analysis of the data, which is not straightforward to achieve afterwards.
Furthermore, building edges cannot be easily retrieved on 3D point clouds but may be facilitated keeping the raw waveforms. Waveforms lying on building edges have a typical shape (two modes). For instance, if we assume the waveforms have temporal and spatial Gaussian shapes, fitting these two distinct echoes indicates the proportion of energy that has felt on the roof, and on the ground. Eventually, a good hint on the localization of the edge is retrieved, and can be refined using several waveforms lying on the same border.
For building characterization, another relevant field of research is the direct estimation of the roof normal vector without any point-based analysis. As already performed by several authors, an hypothesis-and-verify approach should help to carry out such estimation. Several superimposed waveforms from all hypotheses are compared to verify the estimate. Moreover, waveforms from different strips (*i.e.*, different points of view) can also be used to refine the normal vector computation.
Eventually, the issue of building roof material and ground surface discrimination has to be tackled. Several possibilities are conceivable. The first would be to carry out simulation studies in order to assess whether surface micro-structures (*e.g.*, tiles for a roof, or low grass for the ground) have an impact on the shape of the waveforms. In order to enhance surface micro-characteristics, and in order to mitigate the local discrepancies that may appear at fine scales, it should be better to perform such analysis at the object level, *i.e.*, after a first classification of ground and building regions. Besides, we have shown that the amplitude may be a relevant cue for such segmentation. The forthcoming development of sensors with multiple wavelengths, *i.e.*, not restricted to the infra-red domain, may help to carry

out such task. Conversely, by assuming the type of surface known, lidar waveforms may be used synergistically with optical images to better estimate the BRDF of these surfaces.

To top it all, whatever the employed signal processing and classification methods, when one has to deal with such a significant amount of data, many issues raise on the scalability of the data and the processes. Efficient data parsing and processing methods have to be developed, and the actual constant improvement of the performance of the computers (multiple cores etc.) is definitively an advantage for such tasks.

6.3 Outlook on the real contributions and limitations of FW data

In Chapter 2 (Section 2.2.3), we have detailed, rather objectively, the advantages and limitations of using full-waveform data from a beginner end-user point of view. They were based on results presented in the literature, as well as on practical issues that we faced, without any assumption on the subsequent processes.
The same outlook can be performed, still from an end-user point of view, having in mind the 2-step workflow that we have proposed, as well as the latest technological sensor improvements:

- Advantages:
  - **No minimum pulse separation limitation** (+++ → +++): the performance of echo detection is limited by sensor specifications and, due to the asymmetry of several peaks, this is not straightforward to convert an "echo assumption" into an "echo detection".
  - **Additional information retrieved** (+++ → ++): digitizing the waveforms allows to model them as one wants. However, the proposed stochastic approach enlightened the prominence of symmetric peaks. On the one hand, it can be considered as a limitation since this means the echo shapes are homogeneous, and not really specific to objects. On the other hand, there is no concluding that other modelling functions, with other parameters than amplitude and width may not reveal other discriminant features.
  - **Data accuracy improvement** (++ → ++): the qualification of the peak detection step is still a significant improvement compared to on-board methods. However, its is limited in case of overlapping peaks since the two modes may not always be visually distinguishable.
  - **Metadata available** (++ → +++): they are irreplaceable for georeferencing the point cloud, but have also revealed to be mandatory for amplitude correction or amplitude and echo width normalization.
  - **Amplitude data exploitable** (++ → +++): being able to calibrate and correct amplitude values over strips and surveys is all the more important than the feature is among the most relevant ones for urban land-cover classification.
  - **Detection of ground reflections** and **Sensor geometry processing possible**: these two topics have not been tackled in this thesis. Their evaluation is not considered again.

	Full-waveform data properties	Final evaluation
Advantages	No minimum pulse separation limitation	++
	Detection of ground reflections	(+++)
	Additional information retrieved	++
	Data accuracy improvement	++
	Metadata available	++
	Amplitude data exploitable	+++
	Sensor geometry processing possible	(++)
Limitations	Lower PRF	=
	Lower Signal-to-Noise Ratio	–/–––
	Entire waveforms ?	=
	Huge datasets?	–
	Limited acquisition time	=
	No dedicated software	––

Table 6.1: Final evaluation of advantages and limitations of small-footprint full-waveform topographic data. Marks within parenthesis correspond to aspects that have not been re-evaluated.

- Limitations:
  - **Lower Pulse Repetition Frequencies** (– → =): several sensors are now able to reach PRF superior to 100 kHz, and, with multiple overlapping strips, point densities higher than 25 points/m^2 can be reached. Consequently, we no longer consider such feature as a limitation.
  - **Lower Signal-to-Noise Ratio** (-- → –/–––): this depends on the calibration procedure. If the waveforms are acquired over 6 or 7 bits instead of 8 or 12 bits, the background noise may have an impact on the recording process and therefore on all the processing chain. Holes may appear in the point cloud for slopped targets, spatial feature computation will be affected, and be less effective for final classification. However, such problem can be solved with calibration flights.
  - **Entire waveforms** (– → =): the recording process is indeed constrained with multiple thresholds. However, a trade-off has to be found between the recording of useless information and the volume of acquired data. We consider data volume as a genuine limitation to the process of FW data, conversely to the supposed insufficient length of the waveforms.
  - **Huge datasets** (–– → –): this is still a significant limitation. Nevertheless, parallel processing is possible, since computers are now equipped with numerous cores.
  - **Limited acquisition time** (– → =): the acquisition time is bound by hard-drive capacities. There is no doubt that their performance will increase and such limitation will no longer be valid.
  - **No dedicated software** (––– → ––): we believe this point as one of the most crucial issue. This is perhaps the only limitation that may prevent an end-user from adopting the full-waveform technology. However, several tools

have appeared during the three last years, which now allow to process raw signals. Some companies have developed sensors with "on-line" Gaussian decomposition facilities, directly providing a 3D point cloud with amplitude and echo width. Preliminary engineering task is thus facilitated. Finally, the question of a full-waveform standard file format has arisen, especially in the open-source community, enhancing the increasing interest for such kind of data. This also may favor the rapid development of tools for end-users.

6.4 Final outlook

We have proposed a full-workflow for the processing and the analysis of small-footprint airborne full-waveform lidar data over urban areas. It lead us to satisfactory results in term of waveform processing and point cloud classification. To top it all, this thesis has allowed us to figure out the genuine potential of FW data and interesting fields of research.

Full-waveform features (amplitude, width, and backscatter coefficient) should be integrated in traditional applications dealing with 3D lidar point clouds to enhance them. This is the most popular issue at stake at the moment, and there is no doubt that an increasing number of research teams will be involved on the subject. The retrieval of new scattering parameters will necessarily go through the development of standardized radiometric calibration techniques.

Finally, we also advocate the overtaking of standard echo-based approaches by adoption a real "full-waveform" processing strategy. The decomposition and modelling steps are not mandatory. Keeping the waveform topology, and working on the 3D data volume directly seems to be for us most promising field of research.

Part V

Appendices

Appendix A

Thesis Publications

This chapter presents the publications dealing with full-waveform laser scanning that we have co-authored during the PhD thesis. All are not focused on urban areas.
Non peer-reviewed papers are pointed out with the symbol †.

Book chapter

- F. Lafarge, C. Mallet. *Quelques applications à la reconnaissance de formes.* In Applications de la Géometrie stochastique à l'analyse d'images, pp. 235-263, Descombes X. (Ed.), Editions Hermès, 2010, to appear (*in French*).

Journal papers

- L. Guo, N. Chehata, C. Mallet, S. Boukir. *Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests.* ISPRS Journal of Photogrammetry and Remote Sensing, vol. 66(1), pp. 56–66, 2011.
- C. Mallet, F. Lafarge, M. Roux, U. Soergel, F. Bretar, C. Heipke. *A Marked Point Process for Modeling Lidar Waveforms.* IEEE Transactions on Image Processing, vol. 19(12), pp. 3204–3221, 2010.
- C. Mallet, F. Bretar. *Full-Waveform Topographic Lidar: State-of-the-Art.* ISPRS Journal of Photogrammetry and Remote Sensing, vol. 64(1), pp. 1–16, 2009.
- F. Bretar, A. Chauve, J.S. Bailly, C. Mallet, A. Jacome. *Terrain surfaces and 3D land-cover classification from small footprint full-waveform LiDAR data: Application to Badlands.* Hydrology and Earth System Sciences, vol. 13(8), pp. 1531–1544, 2009.
- C. Mallet, F. Bretar, U. Soergel. *Analysis of Full-Waveform Lidar Data for Classification of Urban areas.* Photogrammetrie - Fernerkundung - Geoinformation, vol. 5, pp. 337-349, 2008.
- C. Mallet, F. Bretar. *Le Lidar Topographique à Retour d'Onde Complète: État de l'art.* Traitement du Signal, vol. 24 (6), pp. 441–465, 2007 (*in French*).

Conference papers

- C. Mallet, F. Lafarge, F. Bretar, U. Soergel, C. Heipke. *Lidar Waveform Modeling using a Marked Point Process.* IEEE International Conference on Image Processing, pp. 1713–1716, Cairo, Egypt, November 2009.

- N. Chehata, L. Guo, C. Mallet. *Contribution of Airborne Full-Waveform Lidar and Image Data for Urban Scene Classification.* IEEE International Conference on Image Processing, pp. 1669–1672, Cairo, Egypt, November 2009.

- C. Mallet, F. Lafarge, F. Bretar, M. Roux, U. Soergel, C. Heipke. *A stochastic approach for modelling airborne lidar waveforms.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 38 (Part 3/W8), pp. 201–206, Paris, France, September 2009.

- N. Chehata, L. Guo, C. Mallet. *Airborne Lidar feature Selection for urban classification using Random Forests.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 38 (Part 3/W8), pp. 207–212, Paris, France, September 2009.

- N. David, C. Mallet, T. Pons, A. Chauve, F. Bretar. *Pathway detection and geometrical description from ALS data in forested mountaneous areas.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 38 (Part 3/W8), pp. 242–247, Paris, France, September 2009.

(†) F. Bretar, A. Chauve, C. Mallet, B. Jutzi. *Managing Full Waveform Lidar Data: A Challenging Task for the Forthcoming Years.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 37 (Part 1), pp. 415–420, Beijing, China, July 2008.

- C. Mallet, U. Soergel, F. Bretar. *Analysis of full-waveform lidar data for classification of urban areas.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 37 (Part 3A), pp. 85–92, Beijing, China, July 2008.

- C. Mallet, A. Chauve, F. Bretar. *Analyse et traitement d'ondes lidar pour la cartographie et la reconnaissance de formes: Application au milieu urbain.* RFIA (Reconnaissance des Formes et Intelligence Artificielle), pp. 693–702, Amiens, France, January 2008 (*in French*).

- A. Chauve, C. Mallet, F. Bretar, S. Durrieu, M. Pierrot-Deseilligny, W. Puech. *Processing full-waveform lidar data: modelling raw signals.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. 36 (Part 3/W52), pp. 102–107, Espoo, Finland, September 2007.

Bibliography

[AAB99] B. Aiazzi, L. Alparone, and S. Baronti. Estimation based on entropy matching for generalized Gaussian PDF modeling. *IEEE Signal Processing Letters*, 6(6):138–140, 1999. 95

[AC10] S. Arlot and A. Celisse. A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4:40–79, 2010. 172

[AKHS06] E. Ahokas, S. Kaasalainen, J. Hyyppä, and J. Suomalainen. Calibration of the Optech ALTM-3100 laser scanner intensity data using brightness targets. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 1), Paris, France (on CD-ROM), 2006. 71

[AMR05] H.-E. Andersen, R.J. McGaughey, and S. Reutebuch. Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment*, 94(6):441–449, 2005. 36, 75

[AMS+06] J. Anderson, M.E. Martin, M.-L. Smith, R. Dubayah, M.A. Hofton, P. Hyde, B.E. Peterson, J.B. Blair, and R.J. Knox. The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sensing of Environment*, 105(3):248–261, 2006. 77

[Are09] H. Arefi. *From LIDAR Point Clouds to 3D Building Models*. PhD thesis, Universität der Bundeswehr München, Munich, Germany, 2009. 36

[AS64] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. John Wiley & Sons, New-York, NY, USA, 1964. 127

[ATK+10] C. Alexander, K. Tansey, J. Kaduk, D. Holland, and N.J. Tate. Backscatter coefficient as an attribute for the classification of full-waveform airborne laser scanning data in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5):423–432, 2010. 163, 173

[Axe00] P. Axelsson. DEM generation from laser scanner data using adaptative TIN models. In *International Archives of Photogrammetry and Remote Sensing*, volume 33 (Part B4/1), pages 110–117, Amsterdam, The Netherlands, 2000. 75

[Bal99a] E.P. Baltsavias. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3):199–214, 1999. 46

[Bal99b] E.P. Baltsavias. Airborne laser scanning: existing systems and firms and other ressources. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3):164–198, 1999. 45, 60

[Bal99c] E.P. Baltsavias. A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3):83–94, 1999. 36

[BBM07] Y. Bazi, L. Bruzzone, and F. Melgani. Image thresholding based on the EM algorithm and the generalized Gaussian distribution. *Pattern Recognition*, 40(2):619–634, 2007. 91

[BC10] F. Bretar and N. Chehata. Terrain modelling from lidar range data in natural landscapes: a predictive and bayesian framework. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3):1568–1578, 2010. 36

[BCB+09] F. Bretar, A. Chauve, J.-S. Bailly, C. Mallet, and A. Jacome. Terrain surfaces and 3D landcover classification from small footprint full-waveform lidar data: Application to Badlands. *Hydrology and Earth System Sciences*, 13(8):1531–1544, 2009. 71

[BGV92] B. Boser, I. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. In *Fifth Annual Workshop on Computational Learning Theory*, pages 144–152, Pittsburgh, PA, USA, 1992. 164

[BH99] J.B. Blair and M.A. Hofton. Modeling Laser Altimeter Return Waveform Over Complex Vegetation Using High-Resolution Elevation Data. *Geophysical Research Letters*, 26(16):2509–2512, 1999. 80

[BHGR09] A.M. Baldridge, S.J. Hook, C.I. Grove, and G. Rivera. The ASTER spectral library version 2.0. *Remote Sensing of Environment*, 113(4):83–98, 2009. 109

[BHL+08] C. Briese, B. Höfle, H. Lehner, W. Wagner, M. Pfennigbauer, and A. Ullrich. Calibration of full-waveform airborne laser scanning data for object classification. In *SPIE: Laser Radar Technology and Applications XIII*, volume 6950, pages 328–337, Orlando, FL, USA, 2008. 71

[Bis06] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, New-York, NY, USA, 2006. 97

[BL93] A. Baddeley and M.N.M. Van Lieshout. Stochastic geometry models in high-level vision. *Statistics and Images*, 1(2):233–258, 1993. 125, 128

[BL97] A.L. Blum and P. Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1-2):245–271, 1997. 180

[BM07] D.M. Barber and J.P. Mills. Vehicule based waveform laser scanning in a coastal environment. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 5/C55), pages 35–40, Padua, Italy, 2007. 60

[BMWW04] L. Bruzzone, M. Marconcini, U. Wegmuller, and A. Wiesmann. An advanced system for the automatic classification of multitemporal SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1321–1334, 2004. 90

[Bre01] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001. 196

[Bre06] F. Bretar. *Couplage de Données Laser Aéroporté et Photogrammétriques pour l'Analyse de Scènes Tridimensionnelles*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, Paris, France, 2006. 36

[Bré10] M. Brédif. *3D Building Modeling. Automatic Roof Superstructure Reconstruction and Kinetic Topology-Aware Polyhedral Roof Fitting*. PhD thesis, Telecom ParisTech, Paris, France, 2010. 35

[BRH99] J.B. Blair, D.L. Rabine, and M.A. Hofton. The Laser Vegetation Imaging Sensor: a medium-altitude digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3):115–122, 1999. 38, 59

[BSB+96] V. Blanz, B. Schölkopf, H. Bülthoff, C. Burges, V. Vapnik, and T. Vetter. Comparison of view-based object recognition algorithms using realistic 3D models. In *International Conference on Artificial Neural Networks*, volume 1112 of *Lecture Notes in Computer Science*, pages 251–256, Bochum, Germany, 1996. 170

[BTB05] S. Boughorbel, J.-P. Tarel, and N. Boujemaa. The LCCP for Optimizing Kernel Parameters for SVM. In *International Conference on Artificial Neural Networks*, pages 589–594, Warsaw, Poland, 2005. 172

[BV03] Y. Bengio and P. Vincent. Manifold Parzen Windows. In *Advances in Neural Information Processing Systems (NIPS)*, volume 15, pages 825–832, Vancouver, Canada, 2003. 90

[BZB+03] A. Brenner, H.J. Zwally, C. Bentley, B. Csathó, D.J. Harding, M. Hofton, B. Minster, L. Roberts, J. Saba, R. Thomas, and D. Yi. Derivation of Range and Range Distributions From Laser Pulse Waveform Analysis for Surface Elevations, Roughness, Slope, and Vegetation Heights. Technical report, Geoscience Laser Altimeter System (GLAS) - Algorithm Theoretical Basis Document Version 4.1, Goddard Space Flight Center, Greenbelt, MD, USA, available at `http://www.csr.utexas.edu/glas/atbd.html`, 2003. 59, 67, 68

[Car09] M.J. Carlotto. Effect of errors in ground truth on classification accuracy. *International Journal of Remote Sensing*, 30(18):4831–4849, 2009. 185

[CBD+09] A. Chauve, F. Bretar, S. Durrieu, M. Pierrot-Deseilligny, and W. Puech. FullAnalyze: a Research Tool for Handling Processing and Analyzing Full-waveform Lidar Data. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 1669–1672, Cape Town, South Africa, 2009. 56, 78

[CCD96] G. Celeux, D. Chauveau, and J. Diebolt. Some stochastic versions of the EM algorithm. *Journal of Statistical Computation and Simulation*, 55(4):287–314, 1996. 91

[CD85] G. Celeux and J. Diebolt. The SEM algorithm : a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Computational Statistics Quarterly*, 2(1):73–82, 1985. 135

[CDZ09] F. Chatelain, X. Descombes, and J. Zerubia. Parameter estimation for marked point processes. application to object extraction from remote sensing images. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 5681 of *Lecture Notes in Computer Science*, pages 221–234, Berlin-Heidelberg, Germany, 2009. Springer-Verlag. 157

[CFL06] P.-H. Chen, R.-E. Fan, and C.-J. Lin. A study on SMO-type decomposition methods for support vector machines. *IEEE Transactions on Neural Networks*, 17(4):893–908, 2006. 170

[CGCZ09] M. Carlberg, P. Gao, G. Chen, and A. Zakhor. Classifying Urban Landscape in Aerial LiDAR Using 3D Shape Analysis. In *IEEE International Conference on Image Processing*, pages 1701–1704, Cairo, Egypt, 2009. 163, 174

[CGM09] N. Chehata, L. Guo, and C. Mallet. Airborne lidar feature selection for urban classification using random forests. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 3/W8), pages 207–212, Paris, France, 2009. 194, 195

[CL06] Y.-W. Chen and C.-J. Lin. *Feature Extraction. Foundations and Applications*, chapter Combining SVMs with Various Feature Selection Strategies. Springer, USA, 2006. 181

[CML04] A. Charaniya, R. Manduchi, and S. Lodha. Supervised parametric classification of aerial lidar data. In *Workshop on Real-Time 3D Sensors and their use, in conjunction with IEEE CVPR*, Washington, D.C., USA, 2004. 163

[Cra46] H. Cramer. *Signal Processing with Alpha-Stable Distributions and Applications*. Princeton University Press, Princeton, NJ, USA, 1946. 91

[CS02] S.-H. Cha and S.N. Srihari. On measuring the distance between histograms. *Pattern Recognition*, 35(6):1355–1370, 2002. 132

[CS06] F. Coren and P. Sterzai. Radiometric correction in laser scanning. *International Journal of Remote Sensing*, 27(15-16):3097–3104, 2006. 71

[CSL01] T. Carlsson, O. Steinvall, and D. Letalick. Signature Simulation and Signal Analysis for 3-D Laser Radar. Technical Report FOI-R-0163-SE, FOI, Linköping, Sweden, 2001. 47, 69

[CSZ06] O. Chappelle, B. Schölkopf, and A. Zien. *Semi-supervised learning*. The MIT Press, Cambridge, MA, USA, 2006. 201

[CV95] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995. 164

[CVB+09] A. Chauve, C. Vega, F. Bretar, S. Durrieu, T. Allouis, M. Pierrot-Deseilligny, and W. Puech. Processing full-waveform lidar data in an alpine coniferous forest: assessing terrain and tree height quality. *International Journal of Remote Sensing*, 30(19):5211–5228, 2009. 66, 92

[CVBM02] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee. Choosing multiple parameters for support vector machines. *Machine Learning*, 46:131–159, 2002. 172

[DB00] R. Dubayah and J.B. Blair. Lidar Remote Sensing for Forestry Applications. *Journal of Forestry*, 98(6):44–46, 2000. 51, 76

[DBFJ08] M. Doneus, C. Briese, M. Fera, and M. Janner. Archaeological prospection of forested areas using full-waveform airborne laser scanning. *Journal of Archaeological Science*, 35(4):882–893, 2008. 36, 74

[DBS10] M. Doneus, C. Briese, and N. Studnicka. Analysis of full-waveform ALS data by simultaniously acquired TLS data: Towards an advanced DTM generation in wooded areas. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 7B), pages 193–198, Vienna, Austria, 2010. 60

[DDC+02] J.B. Drake, R.O. Dubayah, D.B. Clark, R.G. Knox, J.B. Blair, M.A. Hofton, R.L. Chazdon, J.F. Weishampel, and S.D. Prince. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, 79(2-3):305–319, 2002. 76

[Dem10] J. Demantké. Analyse géométrique de donnés lidar aéroportées: applications à la segmentation et à la classification. Master's thesis, Université Paris V - René Descartes, 2010. (in French). 179, 195

[DES10] DESDynI. http://desdyni.jpl.nasa.gov/. DESDynI sensor website, 2010. Reference: June 30, 2010. 59

[Des11] X. Descombes. *Applications de la Géometrie stochastique à l'analyse d'images*. Hermès, Paris, France, 2011. 21

[DHU+06] V. Ducic, M. Hollaus, A. Ullrich, W. Wagner, and T. Melzer. 3D Vegetation mapping and classification using full-waveform laser scanning. In *EARSeL and ISPRS Workshop on 3D Remote Sensing in Forestry*, pages 211–217, Vienna, Austria, 2006. 73, 163

[DKP03] K. Duan, S.S. Keerthi, and A.N. Poo. Evaluation of simple performance measures for tuning SVM hyperparameters. *Neurocomputing*, 51:41–59, 2003. 172

[DLR77] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977. 68, 91, 135

[DMB08] N. David, C. Mallet, and F. Bretar. Library concept and design for lidar data processing. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 4/C1) (on CD-ROM), Calgary, Canada, 2008. 55

[DMZ09] X. Descombes, R. Minlos, and Z. Zhizhina. Object extraction using a stochastic birth-and-death dynamics in continuum. *Journal of Mathematical Imaging and Vision*, 33(3):347–359, 2009. 140

[Don06] D.L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006. 158

[DPL06] H. Duong, N. Pfeifer, and R. Lindenbergh. Full waveform analysis: ICESat laser data for land cover classification. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 7), pages 30–35, Enschede, The Netherlands, 2006. 73

[DPLV08] H. Duong, N. Pfeifer, R. Lindenbergh, and G. Vosselman. Single and two epoch analysis of ICESat full-waveform data over forested areas. *International Journal of Remote Sensing*, 29(5):1453–1473, 2008. 66, 68, 77, 90

[DRC97] S. Der, B. Redman, and R. Chellapa. Simulation of error in optical radar range measurements. *Applied Optics*, 36(27):6869–6874, 1997. 47

[DS04] M. Dorigo and T. Stützle. *Ant Colony Optimization*. The MIT Press, Cambridge, MA, USA, 2004. 125

[DTC04] A. Dick, P. Torr, and R. Cipolla. Modelling and interpretation of architecture from several images. *International Journal of Computer Vision*, 60(2):111–134, 2004. 137

[DV02] M.N. Do and M. Vetterli. Wavelet-Based Texture Retrieval using generalized Gaussian density and Kullback-Leibler Distance. *IEEE Transactions on Image Processing*, 11(2):146–158, 2002. 95

[dWSD99] G. Van de Wouwer, P. Scheunders, and D. Van Dyck. Statistical texture characterization from discrete wavelet representations. *IEEE Transactions on Image Processing*, 8(4):592–598, 1999. 95

[EJL+01] M. Elmqvist, E. Jungert, F. Lantz, Å. Persson, and U. Söderman. Terrain modelling and analysis using laser scanner data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, volume 34 (Part 3/W4), pages 229–236, Annapolis, MD, USA, 2001. 175

[Elb10] S. Oude Elberink. *Acquisition of 3D topography. Automated 3D road and building reconstruction using airborne laser scanner data and topographic maps*. PhD thesis, ITC, University of Delft, The Netherlands, 2010. 36

[EM00] S. Oude Elberink and H.-G. Maas. The use of anisotropic height texture measures for the segmentation of airborne laser. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 33 (part B3), pages 678–684, 2000. Amsterdam, The Netherlands. 179

[Fau07] M. Fauvel. *Spectral and spatial methods for the classification of urban remote sensing data*. PhD thesis, Grenoble Institute of Technology, France and University of Iceland, Iceland, 2007. 35, 171, 172

[FCH04] A. Fidera, M. Chapman, and J. Hong. Terrestrial lidar for industrial metrology applications: modelling, enhancement and reconstruction. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 35 (Part B5), pages 880–883, Istanbul, Turkey, 2004. 36

[FJ02] M. Figueiredo and A.K. Jain. Unsupervised Learning of Finite Mixture Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):381–396, 2002. 91

[FL07] S.-K.S. Fan and Y. Lin. A multi-level thresholding approach using a hybrid optimal estimation algorithm. *Pattern Recognition Letters*, 28(S1):662–669, 2007. 91

[FN01] M. Figueiredo and R. Nowak. Wavelet-based image estimation: an empirical Bayes approach using Jeffreys' non informative prior. *IEEE Transactions on Image Processing*, 10(9):1322–1331, 2001. 64

[FNSZ06] G. Forlani, C. Nardinocchi, M. Scaioni, and P. Zingaretti. Complete classification of raw lidar data and 3d reconstruction of buildings. *Pattern Analysis & Applications*, 8(4):357–374, 2006. 163

[Foo10] G. Foody. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, 114(10):2271–2285, 2010. 185

[FP05] S. Filin and N. Pfeifer. Neighborhood systems for airborne laser scanner data. *Photogrammetric Engineering & Remote Sensing*, 71(6):743–755, 2005. 179

[GBS06] P.O. Gislason, J.A. Benediktsson, and J.R. Sveinsson. Random Forests for land cover classification. *Pattern Recognition Letters*, 27(4):294–300, 2006. 196

[GC09] W. Ge and R. Collins. Marked point processes for crowd counting. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2913–2920, Miami, FL, USA, 2009. 130

[GCMB10] L. Guo, N. Chehata, C. Mallet, and S. Boukir. Relevance of airborne lidar and multispectral image data for urban scene classification using random forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2010. In press. 163, 194, 196

[GCRR00] G. Guenther, A. Cunningham, P.L. Rocque, and D. Reid. Meeting the accuracy challenge in airborne lidar bathymetry. In 20^{th} *EARSeL Workshop on Lidar Remote Sensing of Land and Sea*, Dresden, Germany (on CD-ROM), 2000. 38, 58

[GE03] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3(7-8):1157–1182, 2003. 180

[GG84] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984. 129

[GGNZ06] I. Guyon, S. Gunn, M. Nikravesh, and L.A. Zadeh. *Feature Extraction. Foundations and Applications*. Springer, USA, 2006. 180

[GJT07] H. Gross, B. Jutzi, and U. Thoenessen. Segmentation of tree regions using data of a full-waveform laser. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W49A), pages 57–62, Munich, Germany, 2007. 73, 173

[GLA10] GLAS. http://glas.gsfc.nasa.gov/. Website of GLAS system associated with ICESAT mission, 2010. Reference: June 30, 2010. 59

[GM88] G.C. Guenther and H.C. Mesick. Analysis of airborne lidar bathymetric waveforms. In *SPIE Ocean Optics IX*, volume 925, pages 232–241, Orlando, FL, USA, 1988. 38, 58

[GM94] C. Geyer and J. Møller. Simulation and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics*, 21:359–373, 1994. 137

[Gre95] P.J. Green. Reversible Jump Markov Chain Monte-Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, 1995. 125, 126, 137, 138

[GT06] H. Gross and U. Thoennessen. Extraction of lines from laser point clouds. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3), pages 86–91, Bonn, Germany, 2006. 174

[GWBV02] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1-3):389–422, 2002. 180, 182

[Har00] D.J. Harding. Boreas Scanning Lidar Imager of Canopies by Echo Recovery (SLICER): Level-3 Data. Available by special arrangement with Oak Ridge National Laboratory Distribued Active Archive Center, Oak Ridge, TN, USA. http://www.daac.ornl.gov, (on CD-ROM), 2000. 147

[Has70] W.K. Hastings. Monte Carlo sampling using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. 125, 137

[HCCG05] J. Ham, Y. Chen, M.M. Crawford, and J. Ghosh. Investigation of the Random Forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):492–501, 2005. 196

[HCDZ10] S. Ben Hadj, F. Chatelain, X. Descombes, and J. Zerubia. Parameter estimation for a marked point process within a framework of multidimensional shape extraction from remote sensing images. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 3A), pages 1–6, Saint-Mandé, France, 2010. 157

[HD02] C.C. Holmes and D.G.T. Denison. Perfect sampling for the wavelet reconstruction of signals. *IEEE Transactions on Signal Processing*, 50(2):237–244, 2002. 90

[HDP+05] P. Hyde, R. Dubayah, B. Peterson, J.B. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment*, 96(3-4):427–437, 2005. 76

[HFH+09] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I.H. Witten. The WEKA data mining software: An update; SIGKDD Explorations 11(1). Website of the Weka software, 2009. 182

[HGCL99] D.J. Harding, D.B. Gesch, C.C. Carabajal, and S.B. Luthcke. Application of the Shuttle Laser Altimeter in an Accuracy Assessment of GTOPO30, a Global 1-kilometer Digital Elevation Model. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 32 (Part 3/W14), pages 81–85, La Jolla, CA, USA, 1999. 59

[HH10a] B. Höfle and M. Hollaus. Urban vegetation detection using high density full-waveform airborne lidar data - combination of object-based image and point cloud analysis. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 7B), pages 281–286, Vienna, Austria, 2010. 73

[HH10b] M. Hollaus and B. Höfle. Terrain roughness parameters from full-waveform airborne lidar data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 7B), pages 287–292, Vienna, Austria, 2010. 69

[HHL+04] J. Hyyppä, H. Hyyppä, P. Litkey, X. Yu, H. Haggrén, P. Rönnholm, U. Pyysalo, J. Pitkänen, and M. Maltamo. Algorithms and methods of airborne laser scanning for forest measurements. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 8/W2), pages 82–89, Freiburg, Germany, 2004. 36

[HHL+08] B. Höfle, M. Hollaus, H. Lehrner, N. Pfeifer, and W. Wagner. Area-based parameterization of forest structure using full-waveform airborne laser scanning data. In *Silvilaser*, Edinburg, UK (on CD-ROM), 2008. 77

[HJW02] C. Heipke, K. Jacobsen, and H. Wegmann. Analysis of the results of the OEEPE test 'Integrated sensor orientation'. Technical Report 43, Technical Report, OEEPE Official Publications, 2002. 45

[HL02] C.-W. Hsu and C.-J. Lin. A comparison of methods for multi-class Support Vector Machines. *IEEE Transactions on Neural Networks*, 13(6):415–425, 2002. 170

[HLP01] D.J. Harding, M.A. Lefsky, and G.G. Parker. Laser altimeter canopy heigh profiles. Methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, 76(9):283–297, 2001. 76

[HMB00] M.A. Hofton, J.B. Minster, and J.B. Blair. Decomposition of Laser Altimeter Waveforms. *IEEE Transactions on Geoscience and Remote Sensing*, 38(4):1989–1996, 2000. 68, 90, 91, 96, 97

[HMH+09] M. Hollaus, W. Mücke, B. Höfle, W. Dorigo, N. Pfeifer, W. Wagner, C. Bauerhansl, and B. Regner. Tree species classification based on full-waveform airborne laser scanning data. In *Silvilaser*, College Station, TX, USA (on CD-ROM), 2009. 77, 80

[HMWG07] S. Hernández-Marín, A. Wallace, and G. Gibson. Bayesian Analysis of Lidar Signals with Multiple Returns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2170–2180, 2007. 125, 126

[HP04] J. Holmgren and Å. Persson. Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90(4):415–423, 2004. 75

[HP07] B. Höfle and N. Pfeifer. Correction of laser scanning intensity data: Data and model-driven approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(6):1415–1433, 2007. 48, 71

[HRBD02] M. Hofton, L. Rocchio, J. Blair, and R. Dubayah. Validation of Vegetation Canopy Lidar sub-canopy topography measurements for a dense tropical forest. *Journal of Geodynamics*, 34(3-4):491–502, 2002. 147

[HSB03] L. Holden, S. Sannan, and H. Bungum. A stochastic marked point process model for earthquakes. *Natural Hazards and Earth System Sciences*, 3(1/2):95–101, 2003. 125

[HSSV03] D. Hasler, L. Sbaiz, S. Süsstrunk, and M. Vetterli. Outlier modelling in image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):301–315, 2003. 95

[HT98] T. Hastie and R. Tibshirani. Classification by pairwise coupling. *The Annals of Statistics*, 26(2):451–(471, 1998. 200

[HTZ04] F. Han, Z. Tu, and S.C. Zhu. Range image segmentation by an effective Jump-Diffusion method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):1138–1153, 2004. 136

[HUG04] C. Hug, A. Ullrich, and A. Grimm. Litemapper-5600 - A Waveform-Digitizing LIDAR Terrain and Vegetation Mapping System. volume 36 (Part 8/W2), pages 24–29, Freiburg, Germany, 2004. 38, 60

[HZ00] R. Hartley and A. Zisserman. *Multiple View Geometry in computer vision.* Cambridge University Press, Cambridge, UK, 2000. 98

[ICE10] ICESAT. http://icesat.gsfc.nasa.gov/. Website of the ICESAT mission with in particular technical information and GLAS system data, 2010. Reference: June 30, 2010. 59

[IL99] J. Irish and W. Lillycrop. Scanning laser mapping of the coastal zone: the SHOALS system. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3):123–129, 1999. 36

[Jel92] R.V. Jelalian. *Laser radar systems.* Artech House, London, UK, 1992. 47

[JG09a] B. Jutzi and H. Gross. Nearest neighbour classification on laser point clouds to gain object structures from buildings. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 1-4-7/W5), Hannover, Germany. (on CD-ROM), 2009. 174

[JG09b] B. Jutzi and H. Gross. Normalization of lidar intensity data based on range and surface incidence angle. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W8), pages 213–218, Paris, France, 2009. 72

[JNS05] B. Jutzi, J. Neulist, and U. Stilla. Sub-pixel edge localization based on laser waveform analysis. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W19), pages 109–114, Enschede, The Netherlands, 2005. 81

[JS03] B. Jutzi and U. Stilla. Laser pulse analysis for reconstruction and classification of urban objects. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 34 (Part 3/W8), pages 151–156, Munich, Germany, 2003. 52, 69, 81

[JS05] B. Jutzi and U. Stilla. Waveform processing of laser pulses for reconstruction of surfaces in urban areas. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 8/W27), Tempe, AZ, USA. (on CD-ROM), 2005. 81, 97

[JS06] B. Jutzi and U. Stilla. Range determination with waveform recording laser systems using a Wiener Filter. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(2):95–107, 2006. 64, 68, 90, 93, 123

[JS10] Y. Jwa and G. Sohn. A multi-span level analysis for improving 3D power-line reconstruction performance using airborne laser scanning data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 3A), pages 97–102, Saint-Mandé, France, 2010. 36

[JSE02] B. Jutzi, U. Stilla, and B. Eberle. Estimation and measurement of backscattered signals from pulsed laser radar. In *SPIE – Image and Signal Processing for Remote Sensing VIII*, volume 4885, pages 256–267, Agia Pelagia, Greece, 2002. 48

[Jut07] B. Jutzi. *Analyse der zeitlichen Signalform von rückgestreuten Laserpulsen.* PhD thesis, Technical University of Munich, Germany (in German), 2007. 8, 38

[KAHS05] S. Kaasalainen, E. Ahokas, J. Hyyppä, and J. Suomalainen. Study of Surface Brighteness From Backscattered Laser Intensity: Calibration of Laser Data. *IEEE Geoscience and Remote Sensing Letters*, 2(3):255–259, 2005. 70

[Kee01] S.S. Keerthi. Efficient tuning of SVM hyperparameters using radius/margin bound and iterative algorithms. Technical Report CD-01-02, National University of Singapore, Singapore, 2001. 172

[KHM05] S. Kaasalainen, J. Hyyppä, and T. Mielonen. Laboratory calibration of backscattered intensity for laser scanning land targets. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W19), pages 13–17, Enschede, The Netherlands, 2005. 70

[KJ97] R. Kohavi and G.H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273–324, 1997. 180

[KJS08] M. Kirchhof, B. Jutzi, and U. Stilla. Iterative processing of laser scanning data by full waveform analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(1):99–114, 2008. 66, 71

[KKL07] A. Kukko, S. Kaasalainen, and P. Litkey. Effect of incidence angle on laser scanner intensity and surface data. *Applied Optics*, 47(7):986–992, 2007. 69

[KMS+06] B. Koetz, F. Morsdorf, G. Sun, K.J. Ranson, K. Itten, and B. Allgöwer. Inversion of a Lidar Waveform Model for Forest Biophysical Parameter Estimation. *IEEE Geoscience and Remote Sensing Letters*, 3(1):49–53, 2006. 80

[KMSZ08] V. Krylov, G. Moser, S.B. Serpico, and J. Zerubia. Modeling the statistics of high resolution SAR images. Technical Report 6722, INRIA Sophia-Antipolis, France, 2008. 91, 128

[Kot05] S. Kotchenova. *New approaches to retrieval and application of lidar-measured vegetation vertical canopy profiles*. PhD thesis, University of Boston, MA, USA, 2005. 81

[KP98] K. Kraus and N. Pfeifer. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(4):193–203, 1998. 75

[KPO+07] A. Kobler, N. Pfeifer, P. Ogrinc, L. Todorovski, K. Ostir, and S. Dzeroski. Repetitive interpolation: A robust algorithm for dtm generation from aerial laser scanner data in forested terrain. *Remote Sensing of Environment*, 108(1):9–23, 2007. 36

[KR92] K. Kira and L. Rendell. A practical approach to feature selection. In *International Conference on Machine Learning*, volume 1, pages 249–256, Aberdeen, Scotland, UK, 1992. 181

[KRSB06] D.S. Kimes, K.J. Ranson, G. Sun, and J.B. Blair. Predicting lidar measured forest vertical structure from multi-angle spectral data. *Remote Sensing of Environment*, 100(4):503–511, 2006. 76

[KZ04] E.E. Kuruoğlu and J. Zerubia. Modeling SAR Images with a Generalization of the Rayleigh Distribution. *IEEE Transactions on Image Processing*, 13(4):527–533, 2004. 91

[LCA+99] M.A. Lefsky, W.B. Cohen, S.A. Acker, G.G. Parker, T.A. Spies, and D.J. Harding. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, 70(3):339–361, 1999. 76

[LCPH02] M.A. Lefsky, W.B. Cohen, G.G. Parker, and D.J. Harding. Lidar Remote Sensing for Ecosystem Studies. *BioScience*, 52(1):19–30, 2002. 76

[LDG10] F. Lafarge, X. Descombes, and G. Gimel'farb. Geometric Feature Extraction by a Multi-Marked Point Process. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1597–1609, 2010. 129, 130, 138

[LDZ05] C. Lacoste, X. Descombes, and J. Zerubia. Point processes for unsupervised line network extraction in remote sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):1568–1579, 2005. 125, 130

[LDZPD10] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny. Structural approach for building reconstruction from a single DSM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1):135–147, 2010. 125

[Let04] Geophysical Research Letters. Special issue on results from Ice, Cloud, and land Elevation Satellite (ICESat) mission. *Geophysical Research Letters*, 57(2-3-4), 2004. 59

[Lev44] K. Levenberg. A method for the solution of certain problems in least squares. *The Quarterly of Applied Mathematics*, 2:164–168, 1944. 97

[LHC+99] M.A. Lefsky, D.J. Harding, W.B. Cohen, G.G. Parker, and H.H. Shugart. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sensing of Environment*, 68(1):83–98, 1999. 59, 76

[Lie00] M.N.M. Van Lieshout. *Markov point processes and their applications*. Imperial College Press, London, UK, 2000. 21, 125, 126, 129, 130

[Lin97] C.S. Lin. Waveform sampling lidar applications in complex terrain. *International Journal of Remote Sensing*, 18(10):2087–2104, 1997. 38

[Lin09] Y.-C. Lin. *Full waverform airborne laser scanning for DTM generation*. PhD thesis, Newcastle University, UK, 2009. 8, 38

[LKHF06] S. Lodha, E. Kreps, D. Helmbold, and D. Fitzpatrick. Aerial LiDAR data classification using Support Vector Machines. In *International Symposium on 3D Data Processing, Visualization, and Transmission*, pages 567–574, Chapel Hill, NC, USA, 2006. 163

[LM09a] R. Lasaponara and N. Masini. Full-waveform airborne laser scanning for the detection of medieval archaeological microtopographic relief. *Journal of Cultural Heritage*, 10(S1):78–82, 2009. 75

[LM09b] Y.-C. Lin and J.P. Mills. Integration of full-waveform information into the airborne laser scanning data filtering process. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W8), pages 224–229, Paris, France, 2009. 75

[LM10] Y.-C. Lin and J.P. Mills. Factors influencing pulse width of small footprint, full waveform airborne laser scanning data. *Photogrammetric Engineering & Remote Sensing*, 76(1):49–59, 2010. 75

[LMBM05] S. Lachérade, C. Miesch, X. Briottet, and H. Le Men. Spectral variability and bidirectional reflectance behaviour of urban materials at a 20 cm spatial resolution in the visible and near infrared wavelengths. A case study over Toulouse (France). *International Journal of Remote Sensing*, 26(17):3859–3866, 2005. 69

[LMSV10] Y.-C. Lin, J.P. Mills, and S. Smith-Voysey. Rigorous pulse detection from full-waveform airborne laser scanning data. *International Journal of Remote Sensing*, 31(5):1303–1324, 2010. 66, 68, 90, 96

[LRLL07] P. Litkey, P. Rönnholm, J. Lumme, and X. Liang. Waveform features for tree identification. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W52), pages 258–263, Espoo, Finland, 2007. 77

[LUVH05] J.-F. Lalonde, R. Unnikrishnan, N. Vandapel, and M. Hebert. Scale selection for classification of point-sampled 3-D surfaces. Technical Report CMU-RI-TR-05-01, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, 2005. 179

[LVI10] LVIS. http://lvis.gsfc.nasa.gov/. LVIS sensor website, 2010. Reference: June 30, 2010. 59

[LZT10] S. Laky, P. Zaletnyik, and C.K. Toth. Land classification of wavelet-compressed full-waveform lidar data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 3A), pages 115–119, Saint-Mandé, France, 2010. 56

[MAH+99] J.E. Means, S.A. Acker, D.J. Harding, J.B. Blair, M.A. Lefsky, W.B. Cohen, M.E. Harmon, and W.A. McKee. Use of large-footprint scanning airborne Lidar to estimate forest stand characteristics in the western cascades of Oregon. *Remote Sensing of Environment*, 67(3):298–308, 1999. 76

[Mar63] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963. 68, 97

[MB04] F. Melgani and L. Bruzzone. Classification of hyperspectral remote sensing images with SVMs. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8):1778–1790, 2004. 170

[MB07] C. Mallet and F. Bretar. Le lidar topographique à retour d'onde complète: Etat de l'art. *Traitement du Signal*, 24(6):441–465, 2007. 13, 14

[MBH10] W. Mücke, C. Briese, and M. Hollaus. Terrain echo probability assignment based on full-waveform airborne laser scanning observables. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 7A), pages 2157–162, Vienna, Austria, 2010. 75

[Mel07] T. Melzer. Non-parametric segmentation of ALS point clouds using Mean Shift. *Journal of Applied Geodesy*, 1(3):159–170, 2007. 163

[Men63] M.V. Menon. Estimation of the shape and scale parameters of the Weibull distribution. *Technometrics*, 5(2):175–182, 1963. 125

[MKH07] L. Matikainen, H. Kaartinen, and J. Hyyppä. Classification tree based building detection from laser scanner and aerial image data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W52), pages 280–287, Espoo, Finland, 2007. 175

[MLB+09] C. Mallet, F. Lafarge, F. Bretar, M. Roux, U. Soergel, and C. Heipke. A stochastic approach for modelling airborne lidar waveforms. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W8), pages 201–206, Paris, France, 2009. 156, 157

[MMR+01] K.-R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf. An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2):181–201, 2001. 166

[MNT04] K. Madsen, H.B. Nielsen, and O. Tingleff. Methods for non-linear least squares problems. Technical report, Technical University of Denmark, 2004. 98

[MOK+09] G. Mandlburger, J. Otepka, W. Karel, W. Wagner, and N. Pfeifer. Orientation and processing of airborne laser scanning data (OPALS) - concept and first results of a comprehensive ALS software. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 3/W8), pages 55–60, Paris, France, 2009. 56

[MP00] G.J. McLachlan and D. Peel. *Finite Mixture Models*. Wiley Series in Probability and Statistic, New-York, NY, USA, 2000. 90

[MRTT53] M. Metropolis, A. Rosenbluth, A. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953. 137, 139

[Müc08] W. Mücke. Analysis of full-waveform airborne laser scanning data for the improvement of DTM generation. Master's thesis, Institut für Photogrammetrie und Fernerkundung, Technische Universität Wien, Austria, 2008. http://publik.tuwien.ac.at/files/PubDat_170289.pdf (accessed July 5, 2010). 75

[MZS06] G. Moser, J. Zerubia, and S.B. Serpico. SAR Amplitude Probability Density Function Estimation Based on a Generalized Gaussian Model. *IEEE Transactions on Image Processing*, 15(6):1429–1444, 2006. 95, 128

[NBWO06] A. Nayegandhi, J. Brock, W. Wright, and M. O'Connell. Evaluating a small footprint, waveform-resolving lidar over coastal vegetation communities. *Photogrammetric Engineering & Remote Sensing*, 72(12):1407–1417, 2006. 58

[NMJD01] W. Ni-Meister, D.L.B. Jupp, and R. Dubayah. Modeling Lidar Waveforms in Heterogeneous and Discrete Canopies. *IEEE Transactions on Geoscience and Remote Sensing*, 39(9):1943–1958, 2001. 80

[NMLW+97] W. Ni-Meister, X. Li, C.E. Woodcock, J.L. Roujean, and R. Davis. Transmission of solar radiation in boreal conifer forests: Measurements and models. *Journal of Geophysical Research*, 102(D24):29555–29566, 1997. 80

[NMT09] A. Neuenschwander, L. Magruder, and M. Tyler. Landcover classification of small-footprint, full-waveform lidar data. *Journal of Applied Remote Sensing*, 3(1), 2009. 73, 76

[Nor06] L. Nordin. Analysis of waveform data from airborne laser scanner systems. Master's thesis, Luleå University of Technology, Sweden, 2006. 64

[NRS05] M. Netzband, C.L. Redman, and W.L. Stefanov. Challenges for applied remote sensing in the urban environment. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 8/W27), Tempe, AZ, USA, (on CD-ROM), 2005. 35

[NRSL10] P. North, J. Rosette, J. Suárez, and S.O. Los. A Monte Carlo radiative transfer model of satellite waveform LiDAR. *International Journal of Remote Sensing*, 31(5):1343–1358, 2010. 77, 81

[NS95] C.L. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications*. John Wiley & Sons, New-York, NY, USA, 1995. 127

[NW06] J. Nocedal and S.J. Wright. *Numerical optimization*. Springer, New-York, NY, USA, 2006. 98

[ODZ07] M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from Digital Elevation Models using marked point processes. *International Journal of Computer Vision*, 72(2):107–132, 2007. 125, 130

[OFG97] E. Osuna, R. Freund, and F. Girosi. Training Support Vector Machines: An application to face detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 130–136, New York, NY, USA, 1997. 170

[PADF02] E. Punskaya, C. Andrieu, A. Doucet, and W. Fitzgerald. Bayesian curve fitting using MCMC with applications to signal segmentation. *IEEE Transactions on Signal Processing*, 50(3):747–758, 2002. 125

[Pal05] M. Pal. Random Forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1):217–222, 2005. 196

[Par07] C. Parrish. *Vertical Object Extraction from Full-Waveform Lidar Data using a 3D Wavelt-based Approach*. PhD thesis, University of Wisconsin-Madison, WI, USA, 2007. 64

[PDHF07] N. Pfeifer, P. Dorninger, A. Haring, and H. Fan. Investigating terrestrial laser scanning intensity data: quality and functional relations. In *Optical 3-D Measurement Techniques*, pages 328–337, Zurich, Switzerland, 2007. 71

[PDZ05] G. Perrin, X. Descombes, and J. Zerubia. Adaptive Simulated Annealing for Energy Minimization Problem in a Marked Point Process Application. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 3757 of *Lecture Notes in Computer Science*, pages 3–17, Berlin-Heidelberg, Germany, 2005. Springer-Verlag. 125, 140

[PG98] A. Pievatolo and P.J. Green. Boundary detection through dynamic polygons. *Journal of the Royal Statistical Society*, 60(3):609–626, 1998. 129

[PHB97] J. Puzicha, T. Hofmann, and J.M. Buhmann. Non-parametric similarity measures for unsupervised texture segmentation and image retrieval. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 267–272, San Juan, Puerto Rico, 1997. 132

[PHS02] Å. Persson, J. Holmgren, and U. Söderman. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering & Remote Sensing*, 68(9):925–932, 2002. 75

[Pla98a] J.C. Platt. *Advances in Kernel Methods - Support Vector Learning*, chapter Making large-scale SVM learning practical. MIT Press, Cambridge, MA, USA, 1998. 170

[Pla98b] J.C. Platt. *Advances in Kernel Methods - Support Vector Learning*, chapter Fast training of support vector machines using Sequential Minimal Optimization. MIT Press, Cambridge, MA, USA, 1998. 171

[PN09] C. Parrish and R. Nowak. Improved approach to lidar airport obstruction surveying using full-waveform data. *Journal of Surveying Engineering*, 135(2):72–82, 2009. 64

[PNMB+01] B. Peterson, W. Ni-Meister, J.B. Blair, M.A. Hofton, P. Hyde, and R. Dubayah. Modeling Lidar Waveforms using a Radiative Transfer Model. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 34 (Part 3/W4), pages 121–124, Annapolis, MD, USA, 2001. 80

[Poi10] Pointools. http://www.pointools.com/. Website of the Pointools software, 2010. Reference: September 28, 2010. 185

[PSTA05] Å. Persson, U. Söderman, J. Töpel, and S. Alhberg. Visualization and Analysis of Full-Waveform Airborne Laser Scanner Data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W19), pages 103–108, Enschede, The Netherlands, 2005. 50, 68, 91, 92, 97

[PV97] M. Pontil and A. Verri. Properties of support vector machines. Technical Report AIM-1612, MIT, Cambridge, MA, USA, 1997. 170

[PY09] C. Poullis and S. You. Automatic reconstruction of cities from remote sensor data. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2775–2782, Miami Beach, FL, USA, 2009. 36

[Rak03] A. Rakotomamonjy. Variable selection using SVM-based criteria. *Journal of Machine Learning Research*, 3:1357–1370, 2003. 182

[RAM05] S. Reutebuch, H-E. Andersen, and R.J. McGaughey. LIDAR: An emerging tool for multiple resource inventory. *Journal of Forestry*, 103(6):286–292, 2005. 76

[RBP10] A. Roncat, G. Bergauer, and N. Pfeifer. Retrieval of the backscatter cross-section in full-waveform lidar data using b-splines. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 3B), pages 137–142, Saint-Mandé, France, 2010. 65, 123

[RC99] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Verlag, New-York, NY, USA, 1999. 136

[Rei10] J. Reitberger. *3D-Segmentierung von Einzelbäumen und Baumartenklassifikation aus Daten flugzeuggetragener Full Waveform Laserscanner*. PhD thesis, Technical University of Munich, Germany (in German), 2010. 8, 38, 79

[RHHP08] M. Rutzinger, B. Höfle, M. Hollaus, and N. Pfeifer. Object-based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification. *Sensors*, 8:4505–4528, 2008. 73, 163, 175

[Rip08] N. Ripperda. Grammar based facades reconstruction using rjMCMC. *Photogrammetrie Fernerkundung Geoinformation*, 2:83–92, 2008. 137

[Ris78] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978. 135

[RKS07] J. Reitberger, P. Krzystek, and U. Stilla. Combined tree segmentation and stem detection using full waveform lidar data. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W52), pages 332–337, Espoo, Finland, 2007. 77

[RKS08] J. Reitberger, P. Krzystek, and U. Stilla. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing*, 29(5):1407–1431, 2008. 66, 68, 76, 77, 90

[RNSL08] J. Rosette, P. North, J. Suárez, and S.O. Los. Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. *International Journal of Remote Sensing*, 29(5):1475—1493, 2008. 76

[Rot09] F. Rottensteiner. Automatic extraction of buildings from airborne laser-scanner data and aerial images. Habilitation of the Technical University of Vienna, Austria, 2009. 36

[RS98] H. Rue and A. Syverseen. Bayesian Object Recognition with Baddeley's Delta Loss. *Advances in Applied Probability*, 30(1):64–84, 1998. 129

[RSKS09] J. Reitberger, C. Schnörr, P. Krzystek, and U. Stilla. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6):561–574, 2009. 77, 79

[RT08] R.B. Roth and J. Thomson. Practical application of Multiple-Pulse in the Air (MPiA) lidar in large area surveys. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 37 (Part 1), pages 183–188, Beijing, China, 2008. 47

[RTCK05] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik. Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. *Information Fusion*, 6(4):283–300, 2005. 179

[RvK03] M. Robnik-Šikonja and I. Kononenko. Theoretical and empirical analysis of ReliefF and RReliefF. *Machine Learning*, 53(1-2):23–69, 2003. 181, 182

[RWMU08] A. Roncat, W. Wagner, T. Melzer, and A. Ullrich. Echo detection and localization in full-waveform airborne laser scanner data using the averaged square difference function estimator. *The Photogrammetric Journal of Finland*, 21(1):62–75, 2008. 65, 90

[SB91] M. Swain and D. Ballard. Color indexing. *International Journal of Computer Vision*, 7(1):11–32, 1991. 132

[SBAT07] C. Strobl, A.-L. Boulesteix, A.Zeileis, and T.Hothorn. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(25), 2007. 196

[SBR10] F. Samadzadegan, B. Bigdeli, and P. Ramzi. A multiple classifier system for classification of LIDAR remote sensing data using multi-class SVM. In *Multiple Classifier Systems*, pages 254–263, Cairo, Egypt, 2010. 163

[SBS98] B. Schölkopf, C.J.C. Burges, and A.J. Smola. *Advances in Kernel Methods. Support Vector Learning*. The MIT Press, Cambridge, MA, USA, 1998. 169

[Sch97] B. Schölkpf. *Support vector learning*. PhD thesis, Technical University of Berlin, Germany, 1997. 166

[SGKR09] D. Salas-González, E.E. Kuruoğlu, and D.P. Ruiz. Finite mixture of α-stable distributions. *Digital Signal Processing*, 19(2):360–369, 2009. 126

[SLI10] SLICER. http://denali.gsfc.nasa.gov/lapf/. Website of SLICER and SLA sensors. Technical specifications and data available, 2010. Reference: June 30, 2010. 59

[Smi07] K. Smith. *Bayesian Methods for Visual Multi-Object Tracking with Applications to Human Activity Recognition*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2007. 157

[Soe10] U. Soergel. *Radar Remote Sensing of Urban Areas*. Springer, 2010. 35

[Soi10] A. Soininen. http://www.terrasolid.fi. Website of the Terrasolid software packages, 2010. Reference: September 30, 2010. 185

[SR00] G. Sun and K.J. Ranson. Modeling Lidar Returns from Forest Canopies. *IEEE Transactions on Geoscience and Remote Sensing*, 38(6):2261–2626, 2000. 80

[SS02] B. Schölkopf and A.J. Smola. *Learning With Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. The MIT Press, Cambridge, MA, USA, 2002. 164, 170

[SSF02] P. Salamon, P. Sibani, and R. Frost. *Facts, Conjectures and Improvements for Simulated Annealing*. SIAM Monographs on Mathematical Modeling and Computation, 2002. 136, 140

[SSZ07] K. Sun, N. Sang, and T. Zhang. Marked Point Process for Vascular Tree Extraction on Angiogram. In Springer-Verlag Lecture Notes in Computer Science, editor, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, number 4679, pages 467–478, Berlin-Heidelberg, Germany, 2007. 125, 130

[ST09] J. Shan and C.K. Toth. *Topographic Laser Ranging and Scanning. Principles and Processing*. CRC Press, Boca Raton, FL, USA, 2009. 36, 47, 49, 58, 96

[Ste00] O. Steinvall. Waveform simulation for 3-D sensing laser radars. Technical Report FOI-R-1530-SE, FOI, Linköping, Sweden, 2000. 48, 69, 93

[SV01] E. Steinle and T. Vögtle. Automated extraction and reconstruction of buildings in laserscanning data for disaster management. In *Workshop on automatic extraction on man-made object from aerial and space images*, pages 309–318, Ascona, Switzerland, 2001. 197

[SV06] G. Sithole and G. Vosselman. Bridge detection in airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(1):33–46, 2006. 36

[SYJ07] U. Stilla, W. Yao, and B. Jutzi. Detection of weak laser pulses by full-waveform stacking. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 3/W49A), pages 25–30, Munich, Germany, 2007. 66

[SZ07] J. Secord and A. Zakhor. Tree Detection in Urban Regions using Aerial Lidar and Image Data. *IEEE Geoscience and Remote Sensing Letters*, 4(2):196–200, 2007. 163

[Sze10] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer, USA, 2010. http://szeliski.org/Book/. 174

[TMMT04] C.K. Tang, G. Medioni, P. Mordohai, and W.S. Tong. First order augmentations to tensor voting for boundary inference and multiscale analysis in 3-D. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5):594–611, 2004. 174

[TMT10] A. Toshev, P. Mordohai, and B. Taskar. Detection and parsing architecture at city scale from range data. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 328–337, San Francisco, CA, USA, 2010. 174

[TNTM04] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maitre. A New Statistical Model for Markovian Classification of Urban Areas in High-Resolution SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 10(42):2046–2057, 2004. 91, 128

[TOP10] BD TOPO. http://professionnels.ign.fr/ficheProduitCMS.do?idDoc=5667214. Webpage of the topographic component of the RGE®, 2010. Reference: September 30, 2010. 38

[Tóv06] D. Tóvári. *Segmentation Based Classification of Airborne Laser Scanner Data*. PhD thesis, Universität Fridericiana zu Karlsruhe (TH), Karlsruhe, Germany, 2006. 163, 175, 197

[TP09] O. Tournaire and N. Paparoditis. A geometric stochastic approach based on marked point processes for road mark detection from high resolution aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6):621–631, 2009. 125, 130, 137

[Tur60] G.L. Turin. An introduction to matched filters. *IRE Transactions on Information Theory*, 6(3):311–329, 1960. 66

[TW04] K.H. Thiel and A. Wehr. Performance and capabilities of laser scanners - an overview and measurement principle analysis. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 36 (Part 8/W2), pages 14–18, Freiburg, Germany, 2004. 46, 49

[TWH05] K.H. Thiel, A. Wehr, and C. Hug. A New Algorithm for Processing Fullwave Laser Scanner Data. In *Earsel Workshop on Remote Sensing 'Use the third dimension for remote sensing purposes'*, Porto, Portugal, 2005. 65

[TZ02] Z. Tu and S.C. Zhu. Image segmentation by data-driven Markov Chain Monte Carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):657–673, 2002. 138

[Uns99] M. Unser. Splines: A Perfect Fit for Signal and Image Processing. *IEEE Signal Processing Magazine*, 16(6):22–38, 1999. 90

[Urb09] *Joint Urban Remote Sensing Event*, Shangai, China, 2009. 35

[USN08] T. Urban, B. Schutz, and A. Neuenschwander. A survey of ICESat coastal altimetry applications: Continental coast, open ocean island, and inland river. *Journal of the Royal Statistical Society*, 19(1-2):1–19, 2008. 59

[UT10] V. Ussyshkin and L. Theriault. ALTM ORION: Bridging conventional LIDAR and full waveform digitizer technology. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 38 (Part 7B), pages 606–610, Vienna, Austria, 2010. 56

[VAM04] N. Vandapel, O. Amidi, and J.R. Miller. Toward Laser Pulse Waveform Analysis for Scene Interpretation. In *IEEE International Conference on Robotics and Automotion*, pages 950–957, New Orleans, LA, USA, 2004. 69

[Vap95] V. Vapnik. *The nature of statistical learning theory*. Springer-Verlag, New-York, NY, USA, 1995. 164

[Vap98] V. Vapnik. *Statistical learning theory*. John Wiley & Sons, New-York, NY, USA, 1998. 165, 170, 201

[Var96] J. Varanelli. *On the Acceleration of the Simulated Annealing*. PhD thesis, University of Virginia, Charlottesville, VA, USA, 1996. 140

[VCL10] VCL. http://www.geog.umd.edu/vcl/. Website of the Vegetation Canopy Lidar (VCL) mission, 2010. Reference: June 30, 2010. 59

[VM10] G. Vosselman and H.-G. Maas. *Airborne and Terrestrial Laser Scanning*. CRC Press, Whittles Publishing, Boca Raton, FL, USA, 2010. 36, 47, 104

[WA91] H. Wong and A. Antoniou. Characterization and Decomposition of Waveforms for LARSEN 500 Airborne System. *IEEE Transactions on Geoscience and Remote Sensing*, 29(6):912–921, 1991. 68

[Wag10] W. Wagner. Radiometric calibration of small-footprint full-waveform airborne laser scanner measurements: Basic physical concepts. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6):505–513, 2010. 173, 175

[WB07] B. Waske and J.A. Benediktsson. Fusion of support vector machines for classification of multisensor data. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12):3858–3866, 2007. 196

[WB09] B. Waske and M. Braun. Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(5):450–457, 2009. 196

[WGS+98] J. Weston, A. Gammerman, M. Stitson, V. Vapnik, V. Vovk, and C. Watkins. *Advances in Kernel Methods - Support Vector Learning*, chapter Support vector density estimation, pages 293–306. MIT Press, Cambridge, CA, USA, 1998. 90

[WHBD08] W. Wagner, M. Hollaus, C. Briese, and V. Ducic. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. *International Journal of Remote Sensing*, 29(5):1433–1452, 2008. 73, 90, 162

[Whi84] S.R. White. Concepts of scale in simulated annealing. In *IEEE International Conference on Computer Design*, volume 122, pages 646–651, New York, NY, USA, 1984. 140, 141

[WHU+08] W. Wagner, J. Hyyppä, A. Ullrich, H. H. Lehner, C. Briese, and S. Kaasalainen. Radiometric calibration of full-waveform small-footprint airborne laser scanners. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 37 (Part 1), pages 163–168, Beijing, China, 2008. 70, 73, 175

[WLW04] T.-F. Wu, C.-J. Lin, and R. C. Weng. Probability estimates for multiclass classification by pairwise coupling. *Journal of Machine Learning Research*, 5:975–1005, 2004. 200

[WUD+06] W. Wagner, A. Ullrich, V. Ducic, T. Melzer, and N. Studnicka. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2):100–112, 2006. 48, 67, 70, 91, 93, 127

[WUM+04] W. Wagner, A. Ullrich, T. Melzer, C. Briese, and K. Kraus. From single-pulse to full-waveform airborne laser scanners: Potential and practical challenges. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35 (Part B3):201–206, 2004. 49, 50, 96

[WvdLB+10] B. Waske, S. van der Linden, J.A. Benediktsson, A. Rabe, and P. Hostert. Sensitivity of support vector machines to random feature selection in classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 48(7):2880—2889, 2010. 201

[WW98] J. Weston and C. Watkins. Multi-class support vector machines. Technical Report CSD-TR-98-04, University of London, Egham, UK, 1998. 169, 170

[WZR+09] Y. Wang, J. Zhang, A. Roncat, C. Künzer, and W. Wagner. Regularizing method for the determination of the backscatter cross section in lidar data. *Journal of the Optical Society of America, A*, 26(5):1071–1079, 2009. 65

[XZ96] G. Xu and Z. Zhang. *Epipolar geometry in stereo, motion, and object recognition.* Kluwer Academic Publishers, Boston, MA, USA, 1996. 174

[YHKM04] X. Yu, J. Hyyppä, H. Kaartinen, and M. Maltamo. Automatic Detection of Harvested Trees and Determination of Forest Growth Using Airborne Laser Scanning. *Remote Sensing of Environment*, 90(4):451–462, 2004. 76

[YS94] L. Yao and W.A. Sethares. Nonlinear parameter estimation via the genetic algorithm. *IEEE Transactions on Signal Processing*, 42(4):927–935, 1994. 125

[Zhu08a] M. Zhu. Kernels and ensembles: Perspectives on statistical learning. *The American Statistician*, 62(2):97–109, 2008. 196

[Zhu08b] X. Zhu. Semi-supervised learning literature survey. Technical Report TR 1530, University of Wisconsin – Madison, WS, USA, 2008. 201

[ZN10] Q. Zhou and U. Neumann. 2.5D dual contouring: A robust approach to creating building models from aerial LiDAR point clouds. In *European Conference on Computer Vision*, Heraklion, Greece, 2010. 36

[ZSA+02] H. Zwally, B. Schutz, W. Abdalati, J. Abshire, C. Bentley, A A. Brenner, J. Bufton, J. Dezio, D. Hancock, D. Harding, T. Herring, B. Minster, K. Quinn, S. Palm, J. Spinhirne, and R. Thomas. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics*, 34(3):405–445, 2002. 66, 67

Analyse de données lidar à Retour d'Onde Complète pour la classification en milieu urbain

Avec l'émergence récente des systèmes lidar aéroportés à retour d'onde complète, capables de fournir plus qu'une représentation topographique en trois dimensions, se pose la question, entre autres, de son utilité pour l'analyse du milieu urbain. Nous souhaitons en particulier comparer ses performances aux systèmes lidar multi-échos traditionnels. Les signaux lidar fournis sont en effet porteurs d'informations supplémentaires sur les objets atteints. L'objectif final visé dans cette thèse est une cartographie automatique 3D améliorée des zones d'occupation du sol, comme socle de grands nombres d'applications déjà existantes, en partant des données brutes enregistrées.

L'approche proposée se compose de deux grandes phases.
La première étape consiste à traiter les signaux enregistrés pour générer des nuages de points 3D de qualité maîtrisée, ainsi que pour extraire des informations sur la morphologie de ces derniers. Deux méthodes distinctes sont présentées. L'une cherche à améliorer la méthode standard consistant à supposer que tous les échos suivent un modèle gaussien. La deuxième permet d'explorer l'hypothèse d'un mélange de modèles, donc de caractériser chaque écho séparément, tout en proposant une formulation physique simple et flexible du problème.
Dans un second temps, les nuages 3D ainsi générés sont classés en se servant d'attributs spatiaux mais également des attributs morphologiques extraits lors de l'étape précédente. Une approche supervisée utilisant les Séparateurs à Vaste Marge est adoptée pour séparer les zones de sol, de bâtiments, et de végétation. Elle est couplée à un processus de sélection des attributs les plus pertinents. En plus de l'obtention d'une classification de bonne qualité, cette étape met en évidence l'apport des données à retour d'onde complète dans un cadre de cartographie automatique des paysages urbains.

Analysis of Full-Waveform lidar data for urban area mapping

The recent development of airborne full-waveform lidar systems, that are able to provide more than a simple 3D representation of the topographic surfaces, has arisen many questions with respect to the analysis of urban areas, and in relation with traditional multiple pulse sensors. Indeed, the recorded lidar waveforms provide additional information about the geometry and the radiometry of the targets. The final aim of this thesis is to process the raw data in order to achieve a 3D automatic land-cover classification, as a basis for various subsequent applications.

The proposed workflow is composed of two main stages.
The first step consists in processing the waveform in order to generate an improved 3D point cloud, and retrieve morphological information about these points. Two distinct methods have been proposed. The first one aims to improve the standard approach, which assumes that all echoes can be modelled by the Gaussian curve. The second method hypothesizes a mixture of several models – each echo is then characterized separately – and allows to propose a simple physical and flexible solution.
Afterwards, the generated 3D points are classified using both spatial features and the morphological ones retrieved from the previous processing step. A supervised approach using Support Vector Machines is adopted in order to label ground, building, and vegetation areas. Furthermore, the classifier is coupled with a feature selection step. Eventually, results with high accuracy are achieved, and enhance the relevance of full-waveform data for automatically mapping urban areas.